UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

CLASSIFYING LOW PROBABILITY OF INTERCEPT RADAR USING FUZZY ARTMAP

by

Pieter Frederick Potgieter

Submitted in partial fulfilment of the requirements for the degree
Master of Engineering (Electronic Engineering)

in the

Faculty of Engineering, Built Environment and Information Technology
Department of Electrical, Electronic and Computer Engineering

UNIVERSITY OF PRETORIA

February 2012

SUMMARY

CLASSIFYING LOW PROBABILITY OF INTERCEPT RADAR USING FUZZY ARTMAP

by

Pieter Frederick Potgieter

Supervisor:	Prof J. C. Olivier
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Master of Engineering (Electronic Engineering)
Keywords:	Classification, detection, electronic warfare, electronic support, estimation, fuzzy ARTMAP, intercept receiver, low probability of intercept, parameters, performance, radar.

Electronic Support (ES) operations concern themselves with the ability to search for, intercept, track and classify threat emitters. Modern radar systems in turn aim to operate undetected by intercept receivers. These radar systems maintain Low Probability of Intercept (LPI) by utilizing low power emissions, coded waveforms, wideband operation, narrow beamwidths and evasive scan patterns without compromising accuracy and resolution. The term LPI refers to the small chance or likelihood of intercept actually occurring. The complexity and degrees of freedom available to modern radar place a high demand on ES systems to provide detailed and accurate real-time information. Intercept alone is not sufficient and this study focusses on the detection, feature extraction (parameter estimation) and classification (using Fuzzy ARTMAP), of the Pilot Mk3 LPI radar.

Fuzzy ARTMAP is a cognitive neural method combining fuzzy logic and Adaptive Resonance Theory (ART) to create categories of class prototypes to be classified. Fuzzy ARTMAP systems are formed by self-organizing neural architectures that are able to rapidly learn and classify both discreet and continuous input patterns.

To evaluate the suitability of a given ES intercept receiver against a particular LPI radar, the LPI performance factor is defined by combining the radar range, intercept receiver range and sensitivity equations. The radar wants to force an opposing intercept receiver into its range envelope. On the contrary, the intercept receiver would ideally want to operate outside the specified radar detection range to avoid being detected by the radar. The Maximum Likelihood (ML) detector developed for this study is capable of detecting the Pilot Mk3 radar, as it allows sufficient integration gain for

detection beyond the radar maximum range.

The accuracy of parameter estimation in an intercept receiver is of great importance, as it has a direct impact on the accuracy of the classification stage. Among the various potentially useful radar parameters, antenna rotation rate, transmit frequency, frequency sweep and sweep repetition frequency were used to classify the Pilot Mk3 radar. Estimation of these parameters resulted in very clear clustering of parameter data that distinguish the Pilot Mk3 radar. The estimated radar signal parameters are well separated to the point that there is no overlap of features. If the detector is able to detect an intercepted signal it will be able to make accurate estimates of these parameters.

The Fuzzy ARTMAP classifier is capable of classifying the radar modes of the Pilot Mk3 LPI radar. Correct Classification Decisions (CCD) of 100% are easily achieved for a variety of classifier configurations. Classifier training is quite efficient as good generalisation between input and output spaces is achieved from a training dataset comprising only 5% of the total dataset.

If any radar is LPI, there must be a consideration for the radar as well as the opposing intercept receiver. Calculating the LPI performance factor is a useful tool for such an evaluation. The claim that a particular radar is LPI against any intercept receiver is too broad to be insightful. This also holds for an intercept receiver claiming to have 100% Probability of Intercept (POI) against any radar.

OPSOMMING

KLASSIFISERING VAN LAE-WAARSKYNLIKHEID-VAN-ONDERSKEPPING-RADAR MET BEHULP VAN FUZZY ARTMAP

deur

Pieter Frederick Potgieter

Studieleier:	Prof J. C. Olivier
Departement:	Elektriese, Elektroniese en Rekenaaringenieurswese
Universiteit:	Universiteit van Pretoria
Graad:	Magister in Ingenieurswese (Elektroniese Ingenieurswese)
Sleutelwoorde:	Klassifisering, deteksie, elektroniese oorlogvoering, elektroniese ondersteuning, estimasie, fuzzy ARTMAP, onderskeppingsontvanger, lae waarskynlikheid van onderskepping, parameters, werkverrigting, radar.

Elektroniese ondersteuningsoperasies het ten doel om uitsendings van bedreigings te soek, te onderskep, te volg en ook te klassifiseer. Moderne radarstelsels probeer op hulle beurt om hul eie werk te verrig sonder om onderskep te word. Hierdie tipe radarstelsels handhaaf 'n Lae Waarskynlikheid van Onderskepping (LWO) d.m.v. lae senderdrywing, geënkodeerde golfvorms, wyebandfrekwensiegebruik, noue antennabundels en vermydende antennasoekpatrone. Hierdie eienskappe veroorsaak dat 'n LWO radar nie akkuraatheid en resolusie prysgee nie. Die term LWO verwys na die skrale kans of waarskynlikheid van onderskepping deur 'n ontvanger wat die radar se gedrag probeer naspeur. Die komplekse seinomgewing en vele grade van vryheid beskikbaar vir 'n LWO-radar, stel baie hoë eise aan onderskeppingsontvangers om gedetaileerde en akkurate inligting in reële tyd te lewer. Die ondersoek van LWO-radaronderskepping op sy eie is nie voldoende nie. Hierdie studie beskou die deteksie, parameter-estimasie asook klassifikasie (m.b.v. Fuzzy ARTMAP) van die Pilot Mk3 LWO-radar as 'n probleem in die geheel.

Fuzzy ARTMAP is 'n kognitiewe neurale metode wat fuzzy-logika en Aanspasbare Resonante Teorie (ART) kombineer om kategorieë of klassifikasieprototipes te vorm en hulle te klassifiseer. Fuzzy ARTMAP stelsels bestaan uit selfvormende neurale komponente wat diskrete asook kontinue insette vinnig kan leer en klassifiseer.

Om die geskiktheid van enige onderskeppingsontvanger te bepaal word 'n LWO-werkverrigtingsyfer gedefinieer. Hierdie werkverrigtingsyfer kombineer beide radar- en onderskeppings ontvanger

vergelykings vir operasionele reikafstand en sensitiwiteit. Die radar beoog om die onderskeppingsontvanger tot binne sy eie reikafstand in te forseer om die ontvangerplatform op te spoor. Die onderskeppingsontvanger wil daarenteen op ’n veilige afstand (verder as die radarbereik) bly, en nogsteeds die radar se uitsendings onderskep. ’n Maksimale Waarskynlikheid (MW) detektor is ontwikkel wat die Pilot Mk3- radargolfvorms kan opspoor, met voldoende integrasie-aanwins vir betroubare deteksie en wat veel verder strek as die radarreikafstand.

Akkurate radarparameterestimasie is ’n baie belangrike funksie in ’n onderskeppingsontvanger aangesien dit ’n direkte implikasie het vir die akkuraatheid van die klassifikasiefunksie. Vanuit ’n wye verskeidenheid van relevante radar parameters word estimasies van antennadraaitempo, senderfrekwensie, frekwensieveegbandwydte en veegherhalingstempo gebruik om die Pilot Mk3-radar te klassifiseer. Die estimasie van hierdie parameters is duidelik gegroepeer met geen oorvleuling om moontlike verwarring te voorkom. Indien die detektor deteksies verklaar, volg die estimasiefunksie met baie akkurate waardes van radarparameters.

Die Fuzzy ARTMAP-klassifiseerder wat ontwikkel is vir hierdie studie beskik oor die vermoë om die Pilot Mk3 LWO-radar te klassifiseer. Korrekte Klassifikasiebesluite (KKB) van 100% is moontlik vir ’n verskeidenheid klassifiseerderverstellings. Die klassifiseerder behaal ’n goeie veralgemening van in- en uitset ruimtes, en die leer- (of oefen-) roetines is baie effektief met so min as 5% van die volle datastel.

Enige radarstelsel wat roem op LWO moet sowel die radar as ’n moontlike onderskeppingsontvanger in gelyke maat beskou. Die LWO- werkverrigtingsyfer verskaf ’n handige maatstaf vir sulke evaluasies. Om bloot te eis dat ’n radar LWO-eienskappe teenoor enige onderskeppingsontvanger het, is te algemeen en nie insiggewend nie. Dieselfde geld vir ’n onderskeppingsontvanger wat 100% (of totale) onderskepping kan verrig teenoor enige radar.

ACKNOWLEDGEMENTS

All the honour and praise goes to the Lord Jesus Christ who gave me the ability, strength and will to perform this work in order to *equip you [me] with everything good for doing Your [His] will* (Hebrews 13:21 NIV).

I would like to acknowledge the support of the following organisations:

The Armaments Corporation of South Africa (Armscor): For supporting this research and granting permission to publish the results of contributing articles flowing from this study.

Council of Scientific and Industrial Research (CSIR): For supporting this research and providing the opportunity to study this relevant topic.

I would further like to sincerely thank the following people for their help and support, without which this dissertation would not have been possible:

Lili, Henri and Jana Potgieter: I am eternally grateful for your support and encouragement during this exceptionally long process. There is no way to make up for the time that I spend away from you. Thank you for your unselfish consideration. I love you.

My parents: I am entirely blessed by your encouragement, support and love throughout my life. Although you were not physically involved with this work, I trust that you will enjoy the result with me.

My parents in-law: We are very privileged to have you in our lives and around us. Thank you for your encouragement and patience.

Warren du Plessis: Thank you for the encouragement and motivation that you provided especially towards the end of this journey.

Jacques Cilliers: For acting as a technical advisor and a sounding board during the initial LPI study conducted at the CSIR.

Barbara Bradley, Anria Breytenbach and Annamarie Stanton: I am very grateful that you took the time to proofread the most revised portion of this dissertation, the Afrikaans summary.

LIST OF ABBREVIATIONS

ART	Adaptive Resonance Theory
ARTMAP	Adaptive Resonance Theory-like Predictive Mapping
CCD	Correct Classification Decisions
CDF	Cumulative Distribution Function
COMINT	Communications Intelligence
CVR	Crystal Video Receiver
CW	Continuous Wave
DOA	Direction of Arrival
EA	Electronic Attack
ECCM	Electronic Counter-Countermeasures
ECM	Electronic Countermeasures
ELINT	Electronic Intelligence
EM	Electromagnetic
ENR	Energy-to-Noise Ratio
EP	Electronic Protection
ES	Electronic Support
ESM	Electronic Support Measures
EW	Electronic Warfare
FMCW	Frequency Modulated Continuous Wave
FSK	Frequency Shift Keying
HMM	Hidden Markov Model
IF	Intermediate Frequency
LDA	Linear Discriminant Analysis
LPI	Low Probability of Intercept
ML	Maximum Likelihood
MP	Modulation Period
MSE	Mean Square Error
NRL	Naval Research Laboratory
OVA	Output Value Accuracy
PDW	Pulse Descriptor Word
PG	Processing Gain

POI	Probability of Intercept
PRI	Pulse Repetition Interval
PSK	Phase Shift Keying
PW	Pulse Width
RBFNN	Radial Basis Function Neural Network
RF	Radio Frequency
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
SEI	Specific Emitter Identification
SIGINT	Signal intelligence
SNR	Signal-to-Noise Ratio

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 PROBLEM STATEMENT

Electronic Support (ES) operations in the Radio Frequency (RF) domain concerns itself with the ability to search for, intercept, track and classify threat emitters. The resulting information is typically used to perform Electronic Intelligence (ELINT) gathering, mission planning, threat avoidance and countermeasures [1–3]. Modern radar systems in turn aim to operate undetected by intercept receivers (or ES receivers). These radar systems maintain Low Probability of Intercept (LPI) by utilizing low power emissions, coded waveforms, wideband operation, narrow beamwidths and evasive scan patterns without compromising accuracy and resolution. The complexity and degrees of freedom available to modern radar place a high demand on ES systems to provide detailed and accurate real-time information. A typical intercept receiver system is shown in Figure 1.1 indicating the possible stages of operation for both analogue and digital ES receivers. Input signals are intercepted from the antenna (or antenna array) and are mixed to a pre-defined Intermediate Frequency (IF) or base band frequency by the receiver RF front end. Emitters are detected and their properties or features extracted into structures called Pulse Descriptor Words (PDWs) [4–6]. These PDWs contain the values for Pulse Repetition Interval (PRI), Direction of Arrival (DOA), carrier frequency, Pulse Width (PW), Instantaneous Bandwidth (BW) and waveform coding. Detection of radar emitters in intercept receivers is achieved depending on the required PDW parameters needed by a classification system.

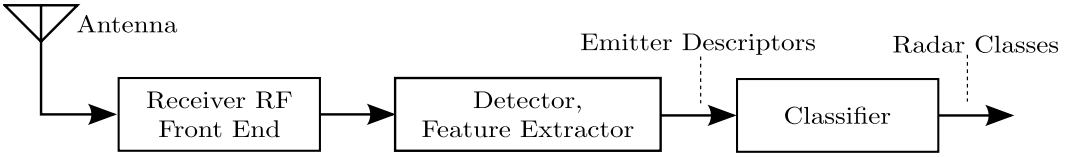


Figure 1.1: Functional diagram of a typical intercept receiver system.

1.2 OVERVIEW OF CURRENT LITERATURE

Correlation based detection methods attempts to model the radar's matched filter as closely as possible by correlating a data stream with a template or correlation filter. The correlation filter typically consists of a stored portion of the input data stream. Time frequency methods like the Wiegner-Ville distribution, Wavelet transform [6] and the Hough transform [7] are capable of detecting modulation parameters such as instantaneous bandwidth and Modulation Period (MP). The outputs of an ES receiver show classification decisions on emitters based on PDW inputs. The classification methods used to classify radar emitters especially in LPI scenarios are either of associative or statistical nature. Associative classifiers make use of membership functions such as fuzzy logic [8, 9] or hyperbox clusters [10–12] to associate radar classes. Associative classification performance is limited when dealing with incomplete data and overlapping class definitions. Statistical classification is achieved with Radial Basis Function Neural Networks (RBFNNs) to estimate likelihood distributions [13] of input data. Also Hidden Markov Models (HMMs) have the ability of evaluating the likelihood of a sequence of events based on a trained model of each class. Both RBFNNs and HMMs train continuous densities and the modelling of these probability densities require parametric descriptions of each density function [14–16]. Both are well suited to classify non stationary behaviour of data on complex decision surfaces. Incomplete data again limits parametric statistical classification methods especially when dealing with noise contaminated data. Fuzzy ARTMAP is a cognitive neural method combining fuzzy logic and Adaptive Resonance Theory (ART) to create categories of class prototypes to be classified. Fuzzy ARTMAP systems are formed by self-organizing neural architectures that are able to rapidly learn, recognize (or classify), test hypothesis and predict consequences of both discreet and continuous input patterns [17–21].

1.3 APPROACH

This study aims to classify radar emitters that employ techniques that make their intercept as unlikely as possible. These cases are commonly referred to as Low Probability of Intercept (LPI) scenarios. The classification of LPI radar emitters is done after the successful intercept, detection and feature extraction done by an intercept receiver. In this study, LPI radar is introduced along with the techniques that distinguish them as LPI. A typical intercept receiver is defined (shown in Figure 1.1) and characterised in order to classify LPI radar emitters. The main focus of this study is to classify LPI radar emitters using Fuzzy ARTMAP classification. This study is approached as follows;

- Define a radar intercept scenario. This includes the radar under investigation, the environment

and the intercept receiver.

- Simulate radar emissions (or signals) and their intercept.
- Detect and extract radar features (or estimate their parameters).
- Classify the radar under investigation.
- Analyse the classification results.

1.4 RESEARCH QUESTIONS

The hypothesis of this study is: *What causes intra-class category selection in a Fuzzy ARTMAP classifier for a given radar emitter operating in a LPI scenario?*
The following assumptions of the hypothesis are made:

1. Successful intercept of radar emitters.
2. A known detector used for detection and the estimation of radar features (or PDW values).

Other research questions flowing from the hypothesis and also addressed in this study are:

- *To what accuracy can LPI radar be classified?*
- *What features provide the greatest classification accuracy?*

1.4.1 Evaluation of topic

Evaluation of the topic will be done according the criteria for evaluating theories outlined in [22]. The criteria are described as follows;

Logically consistent. It is logical that in Fuzzy ARTMAP some input patterns do not activate the current resonating output. The Fuzzy ARTMAP architecture allows hypothesis testing according to the vigilance parameter to either continue to activate an output or to search for another output that will match input patterns better.

Consistent with facts. It is a well known fact that radar operating in LPI scenarios are not easily classified due to their LPI attributes. The understanding of radar behaviour when classifying them quantifies the requirements for improved classifier design.

Testable. The modelling of radar emitters in LPI scenarios as well as a Fuzzy ARTMAP classifier will enable the testing and evaluation of different types of radar behaviour for both simulated and real world data.

Parsimonious. Parsimony is the simplest explanation that explains a great number of observations and properties to be referred to more complex explanations. The hypothesis can be tested for parsimony based on the assumptions made.

Consistent with related theories. The implementation of Fuzzy ARTMAP is consistent with the classification of radar emitters operating in LPI scenarios.

Interpretable. Evaluation of a successful hypothesis will be able the explain the causes of intra-class category searches and be able to predict the Fuzzy ARTMAP classification response to radar emitters operating in LPI scenarios.

Useful. Understanding the response of a Fuzzy ARTMAP classifier to radar emitters operating in LPI scenarios has a direct implication on pre-processing of input data before it is presented to the classifier. Insights will be gained into the grouping of inputs that belong to a radar prototype.

1.5 RESEARCH OBJECTIVES

In order to study the effects of input patterns on a Fuzzy ARTMAP classifier during the learning and selection of categories, the following tasks must be performed. Firstly, LPI scenarios and LPI radar classes will be defined. Low SNR signals qualify as LPI signals due to the overwhelming presence of noise in an intercept receiver system. Partial frequency band intercept will cause received signal PDWs to contain parameter values that are incomplete. For example, an LPI radar transmits a continuously swept frequency modulated signal over a total band of 50 MHz at a certain rate. To the intercept receiver, with a 35 MHz instantaneous bandwidth, only a 35 MHz portion will be covered with the initial phase depending on the start time of reception. Each radar class and an ES receiver will be modelled using a known detector [23, 24]. A Fuzzy ARTMAP classifier will be modelled to investigate radar feature behaviour in the LPI scenarios defined. The Fuzzy ARTMAP architecture is used as a vehicle to determine what factors influence a Fuzzy ARTMAP classifier's decision to search for better categories within a radar class. Models for the detector and Fuzzy ARTMAP classifier will be constructed such that results from real world data can be compared to theoretical (simulated) results. The combination of models will form an intercept system similar to the one shown in Figure 1.1.

1.6 CONTRIBUTION

The modelling of radar, ES receivers (Figure 1.1) and Fuzzy ARTMAP systems (Figure 3.1) will lay the foundation for experimental design and evaluation of hypotheses similar to the one proposed in Section 1.4. Validation of classification results from real world radar emitter data versus simulated radar emitter data will add to the knowledge of determining the requirements for classifying LPI radar. The reasons for a Fuzzy ARTMAP classifier to search for new categories when evaluating radar PDW inputs from an LPI scenario will allow the proposal of possible pre-processing methods to limit potentially redundant intra-class categories. The implication would be faster throughput rates and reduced Fuzzy ARTMAP system complexity. The following articles were published during the course of this research.

1. *Inferring Radar Mode Changes From Elementary Pulse Features Using Fuzzy Artmap Classification*, co-authored with Prof. J. C. Olivier [25]. This paper was presented at the International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR) 2007. The paper details a method for radar mode inference using Fuzzy ARTMAP classification. Using this method elementary radar parameters, Pulse Width (PW) and Pulse Repetition Interval (PRI) originating from a radar operating in a certain mode is input to a Fuzzy ARTMAP classifier. Radar parameters were simulated at different Signal-to-Noise Ratios (SNRs) to train and evaluate the Fuzzy ARTMAP classifier without prior knowledge of radar operating modes. Thus, Fuzzy ARTMAP classification is used in the analysis of radar mode behaviour. Knowledge about the amount of radar modes and mode transition can also be gained by an initial training and evaluation (analysis) process to assign pseudo modes to a particular radar. The resultant modes can then be included into a Fuzzy ARTMAP classifier by increasing the dimension of the predicted output to classify both the radar class and radar operating mode.[1]

2. *Analysis of Measured Radar Data for Specific Emitter Identification*, co-authored with M. Conning [27]. This paper was presented at the IEEE International Radar Conference 2010. It presents findings of analysis of measured radar data for the purposes of Specific Emitter Identification (SEI). Here, the aim was the accurate identification (classification) of a specific emitter within a single class of emitters. Dimensionality reduction of input features was done to obtain a minimal set of input features for peak classification. Fuzzy ARTMAP classification was used to demonstrate the improvement in Correct Classification Decisions (CCD) using various

[1] This paper has subsequently been cited in [26].

reduced feature vectors.[2]

1.7 DISSERTATION OUTLINE

This dissertation is organised in 6 chapters covering the research approach of Sections 1.3 to 1.5 and 2 addenda that present work that are relevant and supporting of the work presented in the main body.

Chapter 1 introduces the topic of study, which is the classification of LPI radar using Fuzzy ARTMAP classification. The LPI radar problem is presented from an Electronic Warfare (EW) point of view. The approach, questions, objectives and contribution of the study are also defined here.

Chapter 2 presents the literature study. Here, the problem stated in Chapter 1 is placed into the context of EW and the need to be able to classify intercepted radar emitter signals. The concept of radar is also introduced from where LPI radar originates. The majority of the chapter is dedicated to a survey of critical LPI radar architectures and concepts.

In Chapter 3 Fuzzy ARTMAP classification is covered. Here, the development of Fuzzy ARTMAP from the subjects of fuzzy logic and Adaptive Resonance Theory (ART) is presented. Detailed explanations of the dynamics and behaviour of a Fuzzy ARTMAP classifier is presented and further explained through illustrative classification examples.

The setting for Fuzzy ARTMAP classification that combines concepts of EW, radar, signal processing, parameter estimation is defined in Chapter 4. Here, a scenario is defined whereby a LPI radar is to be intercepted or *targeted* as well as the intercept receiver who aims to intercept, detect, estimate features of and classify the particular radar. The chapter is referred to as the classification scenario.

Chapter 5 presents the study results from the classification scenario. Here, the intercept receiver's ability to detect, estimate features of and classify intercepted radar signals is considered and evaluated. A rather intensive analysis of each of these intercept receiver elements (or processing stages) is performed. Each processing stage is discussed individually to lead on to conclusions made for this study, which are presented in Chapter 6.

Addendum A presents the results of the search for optimal Fuzzy ARTMAP parameters, which supplements the discussion of classification results in Section 5.3.2 on page 72.

Addendum B presents concepts of dimensionality reduction encountered during this study. The reasons to reduce the dimensions of classification data, and review Linear Discriminant Analysis (LDA) and the Gamma Test as techniques to analyse and reduce classification data dimensions are presented here. These techniques were applied to radar classification in a contributing article for this study [27].

[2] This paper has subsequently been cited in [28] and [29].

CHAPTER 2

LITERATURE STUDY

2.1 ELECTRONIC WARFARE

Electronic warfare (EW) is a military action whose aim is to control the electromagnetic (EM) spectrum. The objective is to exploit, reduce or prevent hostile use of the electromagnetic spectrum while still retaining friendly use of the electromagnetic spectrum. EW is extremely dependent on timely, accurate and focused intelligence. Intelligence is the means of observation, sensing and recording of information, conditions and events. Operational success is achieved by intelligence dominance. However it is perishable, therefore gaining and maintaining intelligence is key in successful EW. Signal intelligence (SIGINT) is part of eight primary intelligence sources for EW. Signal intelligence is divided in two subdivisions, Electronic Intelligence (ELINT) and Communications Intelligence (COMINT).

Electronic Warfare comprises of three main disciplines.

1. Electronic Support (ES), previously known as Electronic Support Measures (ESM).
2. Electronic Attack (EA), previously known as Electronic Countermeasures (ECM).
3. Electronic Protection (EP), previously known as Electronic Counter-Countermeasures (ECCM).

These disciplines are shown in Figure 2.1 and are described in detail in the following subsections.

2.1.1 Electronic support

The start of any EW operation is by performing Electronic Support (ES). It provides the necessary information, description and intelligence to enable (or support) effective EA and EP. Electronic Support is defined as the search for, interception, location and classification of sources of intentional and

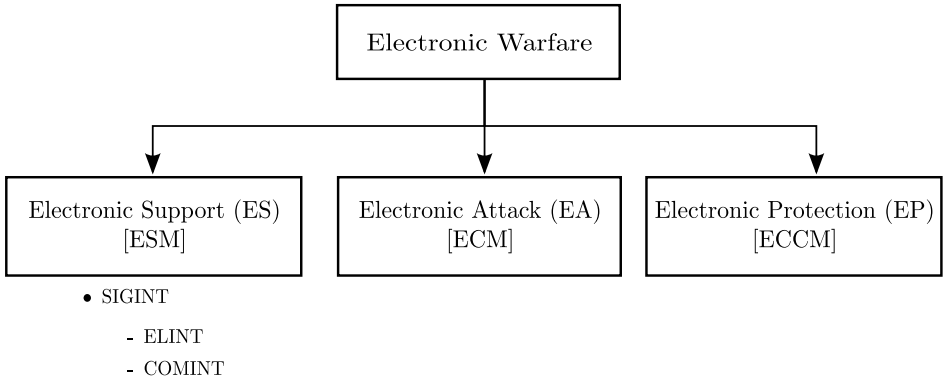


Figure 2.1: Breakdown of the main EW disciplines.

unintentional radiated EM energy.

2.1.2 Electronic attack

Electronic Attack (EA) deals with the deliberate actions taken to radiate or reflect electromagnetic (EM) energy in order to disable or degrade the EM spectrum to enemy capabilities. EA comprises of impairing, disrupting and deceiving enemy assets to gain control in military operations [2, 30]. A typical EA scenario would be an aircraft having to create phase front distortion towards a tracking radar in order to break it's track. A technique commonly known as cross-eye jamming. Other widely used terms to characterise EA are jamming, spoofing and deception jamming. EA has the following the specific sub divided applications:

- Active radiation of EM energy.
- Passive EA. For example, chaff and passive decoys.
- Reductions of radar observations of targets.

2.1.3 Electronic protection

The actions taken to protect facilities, equipment and personnel from any effects of friendly or enemy EW (especially EA) is referred to as Electronic Protection (EP). The following strategies or tactics are regarded as *protecting* facilities against enemy EA:

- Overpowering of jammers.
- Avoiding jamming signals.
- Preventing receiver overload.

- Radar versus jamming signal discrimination.

Examples of EP are emissions control and communication security. By controlling where, when, how often and on which frequency you are operating, the would-be jammer will find it difficult to meet all these conditions for effective EA. Furthermore, securing transmitted data using encryption protects the content of data even if it is intercepted.

2.2 RADAR

The word *radar* was originally an acronym, RADAR, for *radio detection and ranging*. Today, the technology is so common that the word has become a standard English noun. Many people have direct personal experience with radar in such applications as measuring tennis ball speeds or, often to their regret, traffic control. The history of radar extends to the early days of modern electromagnetic theory. In 1886, Hertz demonstrated reflection of radio waves, and in 1900 Tesla described a concept for electromagnetic detection and velocity measurement in an interview. In 1903 and 1904, the German engineer Hulsmeyer experimented with ship detection by radio wave reflection, an idea advocated again by Marconi in 1922. In that same year, Taylor and Young of the U.S. Naval Research Laboratory (NRL) demonstrated ship detection by radar, and in 1930 Hyland, also of the NRL, first detected aircraft (albeit accidentally) by radar, setting off a more substantial investigation that led to a U.S. patent for what would be called a Continuous Wave (CW) radar in 1934.

The development of radar accelerated and spread in the middle and late 1930s, with largely independent developments in the United States, Britain, France, Germany, Russia, Italy, and Japan. In the United States, Page of the NRL began an effort to develop pulsed radar in 1934, with the first successful demonstrations in 1936. The year 1936 also saw the U.S. Army Signal Corps begin active radar work, leading in 1938 to its first operational system, the SCR-268 anti-aircraft fire control system and in 1939 the SCR-270 early warning system, the detections of which were tragically ignored at Pearl Harbour. British development, spurred by the threat of war, began in earnest with work by Watson-Watt in 1935.

Early radar development was driven by military necessity, and the military is still the dominant user and developer of radar technology. Military applications include surveillance, navigation, and weapons guidance for ground, sea, and air vehicles. Military radars span the range from huge ballistic missile defence systems to fist-sized tactical missile seekers [31].

2.2.1 Low probability of intercept radar

To an EW receiver, intercept occurs when all of the following criteria are met;

1. Alignment in the time domain. The intercept receiver has to be accepting or *open* to incident radar signals arriving at the intercept receiver.

2. Alignment in the frequency domain. The operating bandwidth of the intercept receiver should ideally be matched to the signal bandwidth covered by the radar signals arriving at the receiver. Partial matching result in interception of only a portion of the signal whilst the rest of the signal is attenuated by band-limiting receiver components.

3. Spatial alignment. The antennas of both the radar and the intercept receiver must point toward each other to have a radar signal with sufficient power arrive at the receiver.

The term Low Probability of Intercept (LPI) refer to the small chance or likelihood of intercept actually occurring. Any candidate intercept receiver would still have to detect the presence of a radar signal whether it was intercepted or not. This is because interception itself does not reveal the presence of a radar signal. The following quotation from [32] suggests that there is more to the concept of intercepting radar signals.

> *The probability of intercept in itself is not usually the aspect of practical interest. The practical situation requires intercepting a signal of interest within a certain time, or with the expenditure no more than a certain total effort, which may be described as the cost of searching for a signal.*

In the following subsections not much attention to the theoretical formulation of probability of intercept is given, but the focus is rather on the cost of detecting (or searching) intercepted radar signals.

Frequency modulated continuous wave radar

Frequency Modulated Continuous Wave (FMCW) radar transmissions sweep a single frequency sinusoid across a given frequency band, repeating sweeps continually. The received waveform is correlated with the transmitted version so that the result is a waveform located on an intermediate frequency that is related to target range. This is also referred to as homodyne detection. Figure 2.2 shows a block diagram of such a radar. Early versions of FMCW radar used two separate antennas for transmit and receive, but recent systems make use of a single antenna that provides sufficient isolation between both signal paths. FMCW radar has some advantages over conventional high power pulsed radar. The

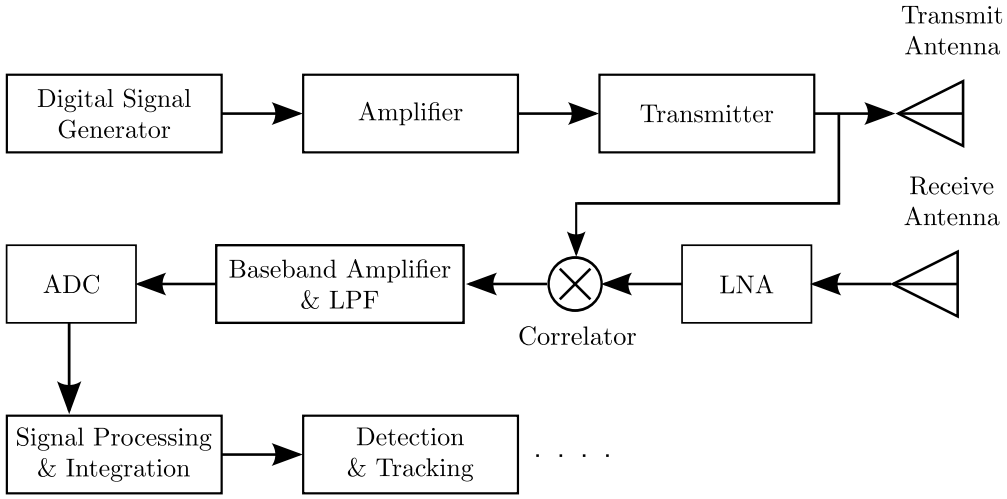


Figure 2.2: System diagram of a FMCW radar. (Reproduced with permission from [6]. © 2004 by Artech House, Inc.)

swept frequency band ΔF increase the range resolution needed to distinguish targets that are closely spaced. The continuous operation (or 100% duty cycle) does not require very high transmit power as it achieves the same amount of average power onto a target. The little amount of return power is easily *recovered* with coherent processing gain on the continually swept waveform. The processing gain of a FMCW radar is defined as the time-bandwidth product as

$$PG_R = \tau_m \Delta F \ . \tag{2.1}$$

Here, τ_m is the modulation period or sweep repetition interval and ΔF the swept frequency band. Furthermore, FMCW transmitters are build using solid–state components and small, light weight units have found their way into man portable and even submarine applications. Interception is rather difficult as the swept bandwidth is sometimes wider than what can be permitted in a non co-operative EW receiver. The transmit waveforms are deterministic and return waveforms can be predicted so as to indicate any intentional interference (or jamming).

Frequency shift keying radar

Frequency Shift Keying (FSK) or frequency hopping techniques build on FMCW techniques to further increase processing gain whilst transmitting continuously. The carrier frequency is changed or *hopped* from a frequency hopping sequence that ultimately reduce matched filter response sidelobes [33]. To an intercept receiver this sequence appears to be random and quite difficult to keep track of. The range resolution is independent of the hopping bandwidth, but the rate at which frequency is hopped.

FSK radar has waveform diversity by using a multitude of hopping sequences to achieve it's objective of target detection and tracking.

Common hopping sequences used in FSK radar are Costas sequences [34]. These codes achieve unambiguous range and Doppler measurements whilst minimising the coupling between range and Doppler (time and frequency). Peak sidelobes in the time-frequency plane are lower by a factor of $1/N$, with N the number of frequencies used in the sequence. The frequency sequence $\{f_1, f_2, \ldots, f_N\}$ of a Costas code must have the property

$$f_{k+i} - f_k \neq f_{j+i} - f_j \tag{2.2}$$

so that for every i, j and k with $1 \leq k < i < i + j \leq N$.

There are some disadvantages to using the FSK technique. The power spectral density of transmitted waveforms is not reduced in FSK radar and consequently increases the probability of detection. Fortunately, the frequency hopping reduces the probability of intercept to an intercept receiver. Spurious responses and phase noise in analogue components of the radar receiver is also difficult to control.

Phase shift keying radar

Low sidelobe levels with respect to target returns are important to any radar design. Phase Shift Keying (PSK) techniques aim to do just that. A block diagram of a typical phase coded radar is shown in Figure 2.3. Phase coded radar still has wideband characteristics and provide the radar with waveform diversity and improved dynamic range (or lower sidelobe levels). Increased dynamic range gives the radar the ability to permit an extended range of signal returns from multiple targets. Range resolution is proportional to the inverse of the waveform sub-period, t_b. During a single code period, T, the CW signal phase is shifted by ϕ_k every sub-code period for N periods that result in $T = t_b N$. There are many different kinds of PSK techniques and each is briefly explained in the following paragraphs.

Binary phase shift keying. Binary phase code sequences has only two possible phase values, $\phi_k \in \{0, \pi\}$. Initially developed for synchronisation in telecommunications applications by R. H. Barker, binary phase codes are commonly referred to as Barker Codes. Allowable Barker code lengths are $N = \{2, 3, 4, 5, 7, 11, 13\}$. Larger codes are generated by formulating a Barker code within itself (known as compound Barker codes) to increase processing gain for this BPSK technique.

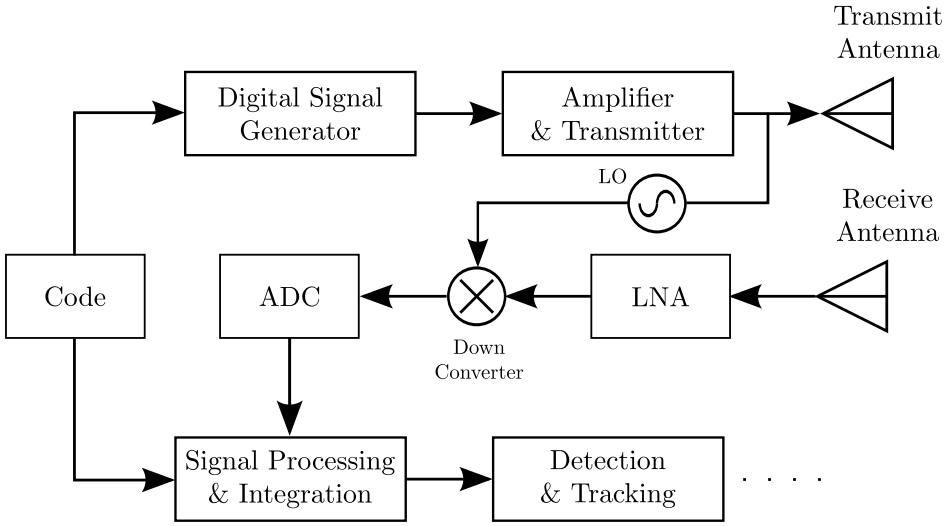


Figure 2.3: System diagram of a phase coding radar. (Reproduced with permission from [6]. © 2004 by Artech House, Inc.)

Polyphase codes. Polyphase codes extend the binary case by using a sequence of multiple discrete phase shifts. The advantages of polyphase sequences are that longer sequences (and thus greater processing gain) can be constructed.

Frank code. Frank codes are a hybrid combination of linear frequency modulation and Barker codes [35, 36]. These codes are created using a step approximation to a linear frequency modulated waveform that contain N frequency steps and n samples per frequency. The total code length and processing gain of such a code is $N_c = Nn$. Usually the frequency steps are chosen to be equal to the number of samples per frequency so that $N_c = N^2$. The phase of a Frank code waveform at sample $i = 1,2,\ldots,N$ during frequency $j = 1,2,\ldots,N$ is

$$\phi_{ij} = \frac{2\pi}{N}(i-1)(j-1)\,. \tag{2.3}$$

A variation of the Frank code that has a double sideband response is the P1 code. It is also a step approximation of a linear frequency modulation waveform and the P1 waveform phase at sample i during frequency j is

$$\phi_{ij} = \frac{-\pi}{N}[N-(j-1)][(j-1)N+(i-1)]\,. \tag{2.4}$$

The P2 code extends the P1 code whereby the initial phase for each frequency step is different from the previous frequency step [37]. For an even $N = 2,4,6,\ldots$ the waveform phase in a P2 code is

$$\phi_{ij} = \frac{-\pi}{2N}(2i-1-N)(2j-1-N)\,. \tag{2.5}$$

Yet another polyphase code used is the P3 code. Detection on this code is performed after complex baseband down conversion. This code is easily distinguished from other polyphase codes as the largest phase increments occur close to the centre of the sequence. The phase of sample i in a P3 code is

$$\phi_i = \frac{\pi}{N}(i-1)^2 \,. \tag{2.6}$$

Lastly, the P4 code is similar to the P3 code and designed for detection at an intermediate frequency. The code is made up out of the discrete phase of a linear frequency modulated waveform taken at specific intervals [38]. It has range-Doppler coupling similar to an unweighted linear chirp waveform but with lower sidelobes [6]. With P4 codes the largest phase increments occur at both ends of the code sequence. The phase at sample i in a P4 code is

$$\phi_i = \frac{\pi(i-1)^2}{N} - \pi(i-1) \,. \tag{2.7}$$

Polytime codes. Thus far, all of the phase coded waveforms considered approximate a linear frequency modulated signal, with a constant dwell time for each state of phase in the waveform sequence. Waveforms that vary the dwell on a particular phase state are termed Polytime codes [39]. Polytime codes consist of two types. The first set of codes are derived from a stepped frequency waveform with two variants, the T1 and T2 codes. The second set are polytime approximations to linear frequency modulated waveforms with another two variants termed T3 and T4 codes. T1 codes have an initial frequency of zero and the expression for signal phase of the T1 code as a function of time is

$$\phi^{T1}(t) = \frac{2\pi}{n}\,\text{floor}\left[\frac{jn}{T}(kt - jT)\right] \,. \tag{2.8}$$

Here, $j = 0, 1, \ldots, k-1$ is the segment number in the stepped frequency waveform, k is the number of segments in the T1 code sequence, n is the number of phase states and T is the total code period. The term enclosed by the square brackets are rounded towards zero with the *floor* operator. The expression of signal phase in a T2 code is

$$\phi^{T2}(t) = \frac{2\pi}{n}\,\text{floor}\left[(kt - jT)\left(\frac{2j-k+1}{T}\right)\frac{n}{2}\right] \,. \tag{2.9}$$

The set of polytime approximations to linear frequency modulated waveforms are as follows. The T3 code with modulation bandwidth, ΔF and modulation period, t_m.

$$\phi^{T3}(t) = \frac{2\pi}{n}\,\text{floor}\left(\frac{n\Delta F t^2}{2t_m}\right) \tag{2.10}$$

Finally, the expression for signal phase in the T4 polytime code is

$$\phi^{T4}(t) = \frac{2\pi}{n}\,\text{floor}\left(\frac{n\Delta F t^2}{2t_m} - \frac{n\Delta F t}{2}\right) \,. \tag{2.11}$$

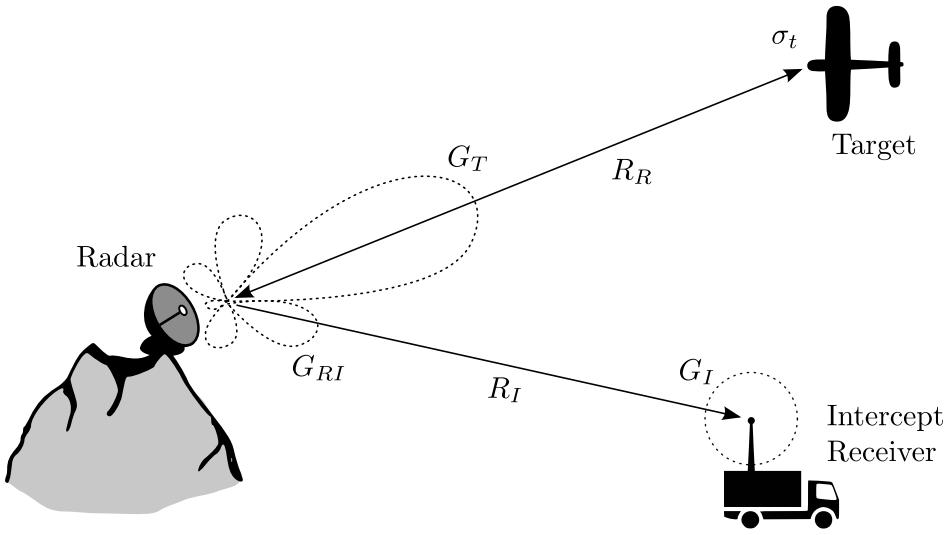


Figure 2.4: Radar intercept scenario.

Until now a significant presentation of typical LPI waveforms and signal processing is given. These are however not the only characteristics that made a radar LPI. From the problem statement in Section 1.1 the combined use of low power emissions, coded waveforms, wideband operation, narrow beamwidths and evasive scan patterns result in LPI. In order to consolidate the aforementioned LPI characteristics radar performance is related with intercept receiver performance using the well known radar range equation [40] and an illustration of a typical intercept scenario is shown in Figure 2.4.

Suppose a radar is located on an elevated position searching for possible airborne targets with an intercept receiver seeking emissions for intelligence purposes. The radar scans in the azimuth plane with a directional antenna. The received signal power, P_R at the radar receiver reflected of a target with cross-section, σ_t, is calculated as

$$P_R = \frac{P_T G_T G_R \lambda^2 \sigma_t}{(4\pi)^3 R_R^4 L_R} \,. \tag{2.12}$$

Here, P_T is the radar transmitter output power, G_T and G_R are the transmit and receive antenna gain respectively. For the remainder of this analysis main beam transmission and reception with $G_T = G_R$ is assumed. In Equation (2.12) λ is the signal wavelength, R_R is the range to the target and L_R accounts for any signal losses within the radar. Radiation patterns in Figure 2.4 are shown as dotted lines. The minimum return power, δ_R, received by the radar to detect a target is

$$\delta_R = kT_0 F_R B_R \eta_R \,. \tag{2.13}$$

Here, k is Boltzmann's constant, F_R is the radar receiver noise figure taken at reference temperature T_0, B_R is the instantaneous bandwidth of the radar receiver and η_R is the minimum input Signal-

to-Noise Ratio (SNR) required by the receiver to detect a target. Target detection involves some processing applied on return signals. The SNR improvement due to this signal processing is referred to as processing gain and is the ratio between the SNR at the receiver input and output respectively. The radar processing gain is

$$PG_R = \frac{\eta_{out}}{\eta_{in}} . \tag{2.14}$$

To the intercept receiver shown in Figure 2.4 the received signal power, P_I at range R_I is

$$P_I = \frac{P_T G_{RI} G_I \lambda^2}{(4\pi)^2 R_I^2 L_I} . \tag{2.15}$$

Here, G_{RI} is the antenna gain from the radar toward the intercept receiver. When the intercept receiver is co-located on the target the antenna gain is $G_{RI} = G_T$ and is dependent on the scenario. G_I is the receiver antenna gain indicated by the dotted pattern around the intercept receiver's antenna in Figure 2.4. Equation (2.15) is very similar to the radar case presented in Equation (2.12) apart from the one-way propagation path indicated by R_I^2. The intercept receiver also has internal signal losses, L_I. Once again, the intercept receiver requires a minimum amount of intercepted signal power to be able to detect the radar. This minimum input signal power, δ_I is very much the same as in Equation (2.13), with variables relating to the receiver with

$$\delta_I = kT_0 F_I B_I \eta_I . \tag{2.16}$$

In order to analyse radar and intercept receiver performance in a combined manner, the range equation (Equation (2.12)) and the minimum required return signal power (Equation (2.13)) are joined to relate radar detection range, R_R to the remaining variables by setting the radar return power equal to the minimum power required to detect a target as follows

$$\delta_R = P_R \tag{2.17}$$

$$kT_0 F_R B_R \eta_R = \frac{P_T G_T^2 \lambda^2 \sigma_t}{(4\pi)^3 R_R^4 L_R} \tag{2.18}$$

$$R_R^4 = \frac{P_T G_T^2 \lambda^2 \sigma_t}{(4\pi)^3 L_R k T_0 F_R B_R \eta_R} . \tag{2.19}$$

The same can be done for the intercept receiver by joining the intercept range equation (Equation (2.15)) and the minimum required intercept signal power (Equation (2.16)) to relate intercept range, R_I to the remaining variables by setting the intercept receiver's return power equal to the min-

imum power required to detect an emitter as follows

$$\delta_I = P_I \tag{2.20}$$

$$kT_0F_IB_I\eta_I = \frac{P_TG_{RI}G_I\lambda^2}{(4\pi)^2R_I^2L_I} \tag{2.21}$$

$$R_I^2 = \frac{P_TG_{RI}G_I\lambda^2}{(4\pi)^2L_IkT_0F_IB_I\eta_I}\,. \tag{2.22}$$

The ratio of intercept receiver range to radar detection range is expressed by using Equations (2.22) and (2.19) as follows

$$\frac{R_I^2}{R_R^2} = \left(\frac{4\pi R_R^2}{\sigma_t}\right)\left(\frac{G_IG_{RI}}{G_T^2}\right)\left(\frac{F_R}{F_I}\right)\left(\frac{B_R}{B_I}\right)\left(\frac{\eta_R}{\eta_I}\right)\left(\frac{L_R}{L_I}\right)\,. \tag{2.23}$$

Using Equation (2.23) analysis may be done of whether the intercept range is sufficient or not when faced with a particular threat radar. Equation (2.23) is further simplified to result in what is referred to as the LPI performance factor, χ [41].

$$\chi^2 = \frac{R_I^2}{R_R^2} = \left(\frac{4\pi R_R^2}{\sigma_t}\right)\left(\frac{G_IG_{RI}}{G_T^2}\right)\left(\frac{F_R}{F_I}\right)\left(\frac{B_R}{B_I}\right)\left(\frac{\eta_R}{\eta_I}\right)\left(\frac{L_R}{L_I}\right) \tag{2.24}$$

The performance factor in Equation (2.24) is a very useful measure, as it provides us with a definition of LPI based on combined radar and intercept receiver performance. LPI radar forces an opposing receiver into it's range envelope (indicated by R_R) when the performance factor is less than one ($\chi < 1$). On the contrary, the intercept receiver would ideally want to operate outside the specified radar detection range ($R_I > R_R$) to avoid being detected by the radar. In this case the LPI performance factor is greater than one ($\chi > 1$). When a radar claims to be LPI (when $\chi < 1$) there exists a range margin from the receiver detection range to the radar range envelope. This margin is called the *quite range* of the LPI radar, $R_{QR} = R_R - R_I$. The quite range is easily related to the radar range envelope and the LPI performance factor as

$$R_{QR} = R_R(\chi - 1) \quad \text{for} \quad \chi \leq 1\,. \tag{2.25}$$

With the LPI performance factor from Equation (2.24) the characteristics that make certain radar LPI can be included and evaluation of the suitability of a given intercept receiver can be performed.

The use of waveform coding as presented earlier provides the radar with improved resolution, wideband operation and most importantly processing gain. The radar can thus afford to transmit using much less power and still perform it's role. Processing gain in the radar has a negative influence on the LPI performance from an intercept receiver point of view. For example, if we assume that the radar and intercept receiver have the same receiver noise figure ($F_R = F_I$) and loss terms ($L_R = L_I$),

the influence of processing gain on the performance factor is

$$\chi^2 = \left(\frac{4\pi R_R^2}{\sigma_t}\right)\left(\frac{G_I G_{RI}}{G_T^2}\right)\left(\frac{B_R}{B_I}\right)\left(\frac{\eta_R^\star}{PG_R\,\eta_I}\right) . \tag{2.26}$$

Here, PG_R is the radar processing gain from Equation (2.14) and $\eta_R^\star$ is the output SNR resulting from radar signal processing. Historically, intercept receivers did not apply any form of processing because radar waveforms were not always known and detection was performed on the energy of individual samples as they were gathered. It is quite straight forward to conclude that if the intercept receiver would turn the performance factor in it's favour some form of gain due to signal processing is needed.

With regard to narrow antenna beams and evasive scan patterns the focus is on the term in the LPI performance factor that relate the radiation patterns of both the radar and intercept receiver. There are two possible situations to consider, that of main beam and sidelobe intercept (inherent in Equation (2.26)). During main beam intercept the gain from the radar antenna to the intercept receiver is equal to the radar antenna gain from its main beam ($G_{RI} = G_R$). The LPI performance factor from Equation (2.26) to include the main beam intercept is expressed as

$$\chi^2 = \left(\frac{4\pi R_R^2}{\sigma_t}\right)\left(\frac{G_I}{G_T}\right)\left(\frac{B_R}{B_I}\right)\left(\frac{\eta_R^\star}{PG_R\,\eta_I}\right) . \tag{2.27}$$

From this equation it is quite intuitive that practical intercept receivers make extensive use of omni-directional antennas, with antenna gain traded for increased loss in signal power due to intercept antenna sidelobes.

The third term in the LPI performance factor to be considered is the difference in radar receiver bandwidth and intercept receiver bandwidth. Suppose a FMCW radar is sweeping a CW waveform that spans 50 MHz at a repetition rate of 1 ms. The receiver front end correlates return signals with the original waveform and result in a de-ramped signal with a bandwidth of 1 KHz (1/0.001 s). To the intercept receiver, wideband Crystal Video Receivers (CVRs) have instantaneous bandwidths of up to 17.5 GHz and wideband super heterodyne receivers have bandwidths up to 500 MHz [42]. This large difference in receiver bandwidths penalises the intercept receiver heavily in the LPI performance factor. If a super heterodyne receiver is used the bandwidth ratio is

$$\frac{B_R}{B_I} = \frac{1\times 10^3}{500\times 10^6} = \frac{1}{500\times 10^3} , \tag{2.28}$$

and the performance factor from Equation (2.26) is

$$\chi^2 = \left(\frac{4\pi R_R^2}{\sigma_t}\right)\left(\frac{G_I G_{RI}}{G_T^2}\right)\left(\frac{1}{500\times 10^3}\right)\left(\frac{\eta_R^\star}{PG_R\,\eta_I}\right) . \tag{2.29}$$

The poor bandwidth ratio in Equation (2.28) is somewhat improved by the use of a channelised intercept receiver that use multiple parallel narrowband receivers, but is still limited to the radar's

transmit signal bandwidth which is 50 MHz. Even if the intercept receiver is exactly matched to the radar transmit bandwidth the performance penalty is only reduced by a factor of 10.

2.3 RADAR CLASSIFICATION

There are many approaches to radar classification, which range from a broad high-level to a very detailed view. Examples of high-level radar classification might be to determine whether a particular set of data belong to a functional group of either tracking, navigation, surveillance or fire control radar systems. In this study a more detailed radar classification approach is taken. Here, the objective is to classify the various operating modes of a chosen radar. Thus, if the classifier is able to reliably recognise such modes, it may be extrapolated as modes belonging to a particular radar type (by make and model), and even further to a functional group of radars.

The detailed view taken here is heavily dependent on a number of critical elements (or factors) that must be considered. Most notably the elements of detection and feature extraction (as shown in Figure 1.1).

To a passive intercept receiver, information about the radar originates from signals emitted by the radar transmitter. A consequence of passive (or non co-operative) interception is that there are other sources of information that compete (or distort) the wanted information, which could be noise, interference, the environment or any combination thereof. Therefore, the intercepted information can not be used directly for classification and has to undergo some form of transformation or processing, such as signal detection and feature extraction. Only then may classification be done.

CHAPTER 3

FUZZY ARTMAP

3.1 THE FUZZY ARTMAP ALGORITHM

A Fuzzy ARTMAP system (shown in Figure 3.1) consists of two adaptive resonance theory modules, ART^a and ART^b. The system aims to create stable categories when presented with arbitrary input sequences of data. Both modules are connected via an associative network called the *Map Field* [18]. Fuzzy ARTMAP control is managed by the map field during learning. The map field ensures that the minimum required recognition categories are formed to meet recognition performance. During supervised learning, input data $a^{(p)}$ is presented to the ART^a module with the ART^b module receiving, $b^{(p)}$ input data with $b^{(p)}$ representing the correct prediction for a given $a^{(p)}$. Fuzzy ARTMAP is capable of fast learning, efficient and accurate as it co-jointly minimizes predictive error and maximizes predictive generalization using independent internal operations. The ART^a vigilance parameter, ρ^a is adjusted to correct predictive errors in ART^b. ρ^a calibrates the minimum confidence that the ART^a module must have in a recognition category so that $a^{(p)}$ is accepted by the ART^a module. If this criteria is not met, a search for another category that will meet the criteria is pursued. This criteria is commonly referred to as the *vigilance criteria*. Small values of ρ^a cause bigger recognition categories to form and implicitly creates more generalization and higher code compression. In the event of a predictive mismatch the interconnecting map field sacrifices just enough generalization for the predictive error to be corrected. This process is called *match tracking*. During match tracking ρ^a is increased by the minimum amount needed to select another ART^a category that will correctly predict $b^{(p)}$ within the ART^b module. Alternatively, another ART^a category is selected which focuses the attention to another set of inputs that is better suited to predict $b^{(p)}$.

3.1.1 Fuzzy ART

Fuzzy ART system learning has the advantage of stable learning as weights are monotonically decreasing. This is a useful attribute but can cause category proliferation as many of the adaptive weight values converge towards zero. Complement coding is a pre-processing step that is applied to input data to prevent category proliferation. This method normalizes input data and preserves amplitude information for each feature dimension. Complement coding also ensures that the degree of absence and the degree of presence of input features are contained in the category weight vector, $\mathbf{w}_j$. Normalisation of a M-dimensional input vector, $\mathbf{I}$ is such that for a given $\gamma > 0$,

$$|\mathbf{I}| \equiv \gamma \tag{3.1}$$

where $|\cdot|$ is the ℓ_1-norm defined as

$$|\mathbf{s}| = \sum_{i}^{M} |s_i| \,. \tag{3.2}$$

The normalised input vector, $\mathbf{I}$ is calculated as

$$\mathbf{I} = \frac{\mathbf{a}}{|\mathbf{a}|} \,. \tag{3.3}$$

As mentioned earlier normalisation preserves amplitude information. Consequently a complement coded vector includes both the present and absent features of the original vector. The absent features are denoted as $\mathbf{a}^c$ where

$$\mathbf{a}^c = 1 - \mathbf{a} \,. \tag{3.4}$$

Thus, an input vector of dimension M is complement coded into a $2M$-dimensional vector so that

$$\mathbf{I} = (\mathbf{a}, \mathbf{a}^c) \equiv (a_1, \ldots, a_M, a_1^c, \ldots, a_M^c) \,. \tag{3.5}$$

Fuzzy ART comprises of three layers (Figure 3.1). The F_0 layer represents the input, $\mathbf{I}$. F_1 receives bottom-up input from F_0 and top down input from F_2, which represents the active category. Each layer has its own activity vectors. The F_0 activity vector is $\mathbf{I} = (I_1, \cdots, I_{2M})$ in complement coded form. The F_1 activity is $\mathbf{x} = (x_1, \cdots, x_{2M})$ and the activity vector for F_2 is $\mathbf{y} = (y_1, \cdots, y_N)$. The initial number of nodes in $\mathbf{y}$ is arbitrary. An adaptive weight vector $\mathbf{w}_j \equiv (w_{j1}, \cdots, w_{j2M})$ is associated with each F_2 category node j $(j = 1, 2, \cdots, N)$. Categories are said to be *committed* when it is chosen for coding, otherwise they are *uncommitted*. Initially the weight values are $w_{j1} = w_{j2} \cdots = w_{j2M} = 1$. The dynamics of a Fuzzy ART system is determined by a *choice parameter* $\alpha > 0$, a *learning rate* parameter $\beta \in [0, 1]$ and a *vigilance* parameter $\rho \in [0, 1]$. Categories are chosen according to a category choice function

$$T_j(\mathbf{I}) = \arg\max \left(\frac{|\mathbf{I} \wedge \mathbf{w}_j|}{\alpha + |\mathbf{w}_j|} \right) \tag{3.6}$$

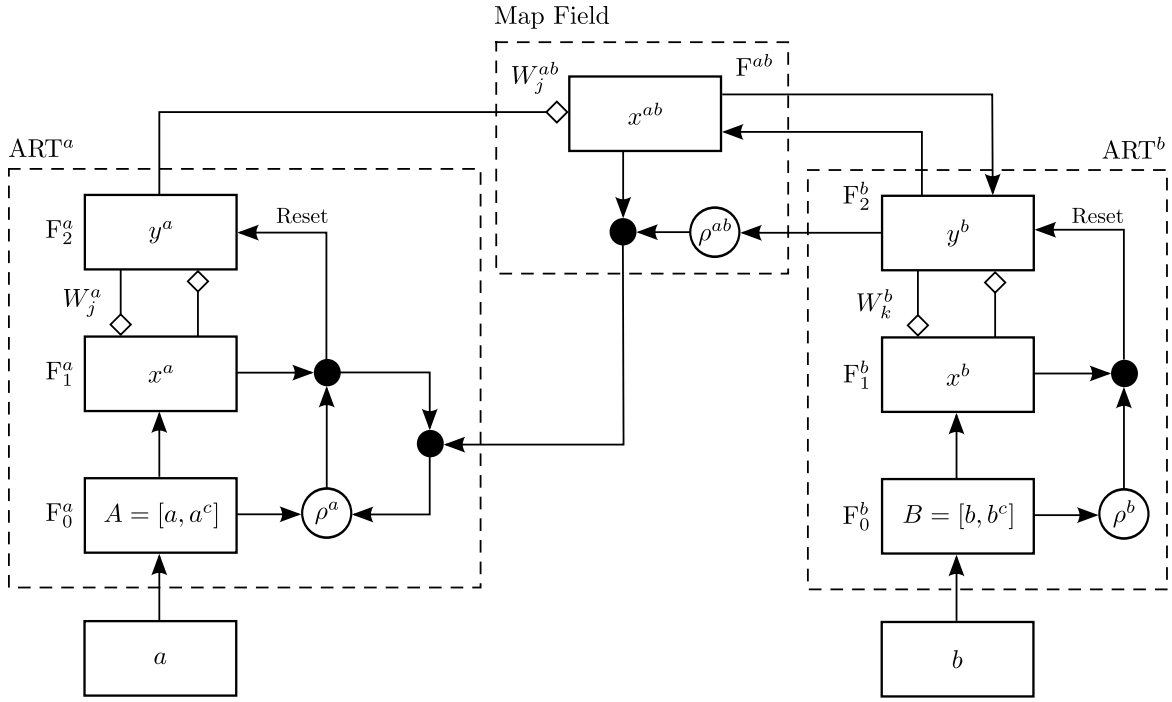


Figure 3.1: Fuzzy ARTMAP architecture and data flow [43].

where $\wedge$ is the Fuzzy AND operator with

$$(\mathbf{s} \wedge \mathbf{q})_i \equiv \min(s_i, q_i) \; . \tag{3.7}$$

If more than one category $T_j(\mathbf{I})$ is maximal, the category with the smallest index j is chosen. Thus categories are committed in order $j = 1, 2, 3, \cdots$. When category J is chosen $y_j = 1$ for $j = J$ and $y_j = 0$ for $j \neq J$. The F_1 activity vector is

$$\mathbf{x} = \begin{cases} \mathbf{I}, & \text{if } F_1 \text{ is inactive} \\ \mathbf{I} \wedge \mathbf{w}_J, & \text{if the } J\text{th node in } F_1 \text{ is active} \; . \end{cases} \tag{3.8}$$

Resonance occurs if the chosen category J meets the following vigilance criteria

$$\frac{|\mathbf{I} \wedge \mathbf{w}_J|}{|\mathbf{w}_J|} \geq \rho \; . \tag{3.9}$$

Mismatch reset and the search for a category that will meet the vigilance criterion occurs when

$$\frac{|\mathbf{I} \wedge \mathbf{w}_J|}{|\mathbf{w}_J|} < \rho \; . \tag{3.10}$$

A new category index is chosen according to Equation (3.6). The previous category T_J is set to zero for the duration of the input presentation so that category J is not chosen repeatedly. Once a category

is chosen that meets the vigilance criteria the weight vector is updated (or trained) with

$$\mathbf{w}_J^{(\text{new})} = \beta\left(\mathbf{I} \wedge \mathbf{w}_J^{(\text{old})}\right) + (1-\beta)\mathbf{w}_J^{(\text{old})} \,. \tag{3.11}$$

3.1.2 Fuzzy ARTMAP

The Fuzzy ARTMAP system depicted in Figure 3.1 shows the interconnection between two ART modules, ART^a and ART^b. The *map field*, F^{ab} links these modules together to form predictive associations between categories and to perform *match tracking* [18]. The interactions between the map field, F^{ab} and other ART modules are as follows. Inputs to the ART^a and ART^b modules are $\mathbf{I} = \mathbf{A} = (\mathbf{a}, \mathbf{a}^c)$ and $\mathbf{I} = \mathbf{B} = (\mathbf{b}, \mathbf{b}^c)$ respectively. Variables and vectors in ART^a and ART^b are denoted by superscript and subscript identifiers, a and b. Then $\mathbf{x}^a \equiv (x_1^a, \cdots, x_{2M_a}^a)$ is the activity vector at F_1^a and $\mathbf{y}^a \equiv (y_1^a, \cdots, y_{N_a}^a)$ is the output vector at F_2^a. The weight vector for category j in F_2^a is $\mathbf{w}_j^a \equiv (w_{j1}^a, w_{j2}^a, \cdots, w_{j2M_a}^a)$. Similarly for ART^b, $\mathbf{x}^b \equiv (x_1^b, \cdots, x_{2M_b}^b)$ is the activity vector at F_1^b and $\mathbf{y}^b \equiv (y_1^b, \cdots, y_{N_b}^b)$ is the output vector at F_2^b. The weight vector for category k in F_2^b is $\mathbf{w}_k^a \equiv (w_{k1}^b, w_{k2}^b, \cdots, w_{k2M_b}^b)$. The map field, F^{ab} output vector is $\mathbf{x}^{ab} = (x_1^{ab}, \cdots, x_{N_b}^{ab})$ and the weight vector from the jth F_2^a node to F^{ab} is $\mathbf{w}_j^{ab} \equiv (w_{j1}^{ab}, \cdots, w_{jN_b}^{ab})$. Vectors $\mathbf{x}^a$, $\mathbf{y}^a$, $\mathbf{x}^b$, $\mathbf{y}^b$ and $\mathbf{x}^{ab}$ are $\mathbf{0}$ in-between input representations. The activation of the map field is subject to activity of either ART^a or ART^b.

When node J is chosen to activate ART^a, the weights $\mathbf{w}_J^{ab}$ activate F^{ab}. When node K activates ART^b, F^{ab} is activated via the ART^b output vector, $\mathbf{y}^b$. When both ART^a and ART^b are active, F^{ab} only becomes active when ART^a correctly predicts the same ART^b category from $\mathbf{w}_J^{ab}$. The F^{ab} activity or output vector, $\mathbf{x}^{ab}$ has the following possible values,

$$\mathbf{x}^{ab} = \begin{cases} |\mathbf{y}^b \wedge \mathbf{w}_J^{ab}|, & \text{if the } J\text{th } F_2^a \text{ node is active and } F_2^b \text{ is active} \\ \mathbf{w}_J^{ab}, & \text{if the } J\text{th } F_2^a \text{ node is active and } F_2^b \text{ is inactive} \\ \mathbf{y}^b, & \text{if } F_2^a \text{ is inactive and } F_2^b \text{ is active} \\ \mathbf{0}, & \text{if } F_2^a \text{ is inactive and } F_2^b \text{ is inactive} \,. \end{cases} \tag{3.12}$$

Predictions are confirmed in Equation (3.12) by producing a single non-zero value in the $\mathbf{x}^{ab}$ vector. Predictions are unconfirmed when all the values in $\mathbf{x}^{ab}$ are zero. Then the map field responds by initiating a search for a better ART^a category with the match tracking process. At the start of each input presentation the ART^a vigilance ρ^a equals a baseline value, $\overline{\rho^a}$ with the map field vigilance parameter, ρ^{ab}. Match tracking occurs when

$$|\mathbf{x}^{ab}| < \rho^{ab}|\mathbf{y}^b| \,. \tag{3.13}$$

During match tracking ρ^a is increased so that it is slightly greater than $|\mathbf{A} \wedge \mathbf{w}_J^a|/|\mathbf{A}|$ and causes the selected Jth ART^a node not to meet the vigilance criteria as in Equation (3.10). This leads to a search for another ART^a category, J so that

$$|\mathbf{x}^a| = |\mathbf{A} \wedge \mathbf{w}_J^a| \geq \rho^a |\mathbf{A}| \tag{3.14}$$

and

$$|\mathbf{x}^{ab}| = |\mathbf{y}^b \wedge \mathbf{w}_J^{ab}| \geq \rho^{ab} |\mathbf{y}^b| \,. \tag{3.15}$$

If no such category exists, a new uncommitted ART^a category is created and F_2^a is shut down (or inactive) for the remainder of the input presentation. Map field learning occurs during resonance when both Equations (3.14) and (3.15) are satisfied [18, 19, 44], and the map field weights, $\mathbf{w}_J^{ab}$ change as follows

$$(\mathbf{w}_{jk}^{ab})^{new} = \begin{cases} (1-\beta_{ab})(\mathbf{w}_{jk}^{ab})^{old} + \beta_{ab} x_k^{ab}, & \text{if } j = J \\ (\mathbf{w}_{jk}^{ab})^{old}, & \text{if } j \neq J \,. \end{cases} \tag{3.16}$$

3.1.3 Fuzzy ARTMAP classification examples

Two examples of Fuzzy ARTMAP classification are presented in the following subsections.

Non-linear signal approximation

The Fuzzy ARTMAP architecture has the ability to approximate output data or functions. This example is a repeat experiment reported on in [45, 46], whereby the classifier is to approximate the values of a function, $f(x)$ in Equation (3.17). The classifier is trained on values of $x \in [0,1]$ sampled at random intervals and then evaluate function approximation for cases where $f(x)$ is error free (no additional noise) and with error $f(x) + \varepsilon$. Classification from noisy data using Fuzzy ARTMAP is reported in [47, 48].

$$\begin{aligned} f(x) = \frac{1}{20} [&\sin(10x) + \sin(20x) + \sin(30x) + \\ &\sin(40x) + \sin(50x) + \sin(60x) + \sin(70x) + 10] \end{aligned} \tag{3.17}$$

To evaluate how good the function approximation results are, the Root Mean Square Error (RMSE) of the classifier output values $\hat{f}(x_n)$ are calculated as

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} |f(x_n) - \hat{f}(x_n)|^2} \,. \tag{3.18}$$

The total dataset contains 10,000 points sampled linearly on the interval $[0,1]$. The training dataset comprises of 2,000 points drawn at random from the greater dataset and used only for training with

the remainder of the data placed into an evaluation dataset. The classifier is configured to train in a *fast commit slow recode* mode with choice parameter $\alpha = 0.001$, learning rate parameter $\beta = 1$ and vigilance parameters, $\rho^a = \rho^b = 0.99$ and $\rho^{ab} = 1$.

Figure 3.2 (a)–(b) show the results from approximating function values from Equation (3.17) out of the evaluation dataset defined above. In the figures the classifier outputs are indicated as predictions along with individual approximation errors and the RMSE for the entire evaluation dataset. In this example fixed parameter values for the Fuzzy ARTMAP classifier were used. It is quite natural to analyse the effect of using other parameter values on classification accuracy for instance. An evaluation of classifier behaviour using variable parameters is presented next.

Circle in the square classification

The next example is a typical classification task in which the Fuzzy ARTMAP classifier has to determine whether 2-dimensional data is located on the inside or the outside of a circle. It is also referred to as the *Circle in the Square* problem [49]. The classifier is trained on 2-dimensional data uniformly distributed in the interval $[0,1]$. Training data is labelled as being either inside or outside the edge of a particular circle. This circle forms the decision boundary when the classifier is trained and when data is evaluated after training. In this example, the influence of the vigilance parameter ρ on the classifier's ability to generalise its own reference (or memory) to the classification task when training, is shown.

The dataset consist of 5,000 points of which 500 points (or 10%) will be used for training and the remaining 4,500 points for evaluation. Training data is drawn at random from the entire set of 5,000 points and training is reiterated until there are no significant change in the weights of the map field, W^{ab}. The classifier is initialised in one of two test cases. In the first test case high vigilance parameters ($\rho^a = \rho^b = 0.8$) are specified to let the classifier include more detail of the training data in its reference (map field). In the second test case vigilance parameters ($\rho^a = \rho^b = 0.1$) are lowered to let the classifier consider a more general reference of the training data. The category choice (α), learning rate (β) and map field vigilance (ρ^{ab}) parameters are held constant for each test case.

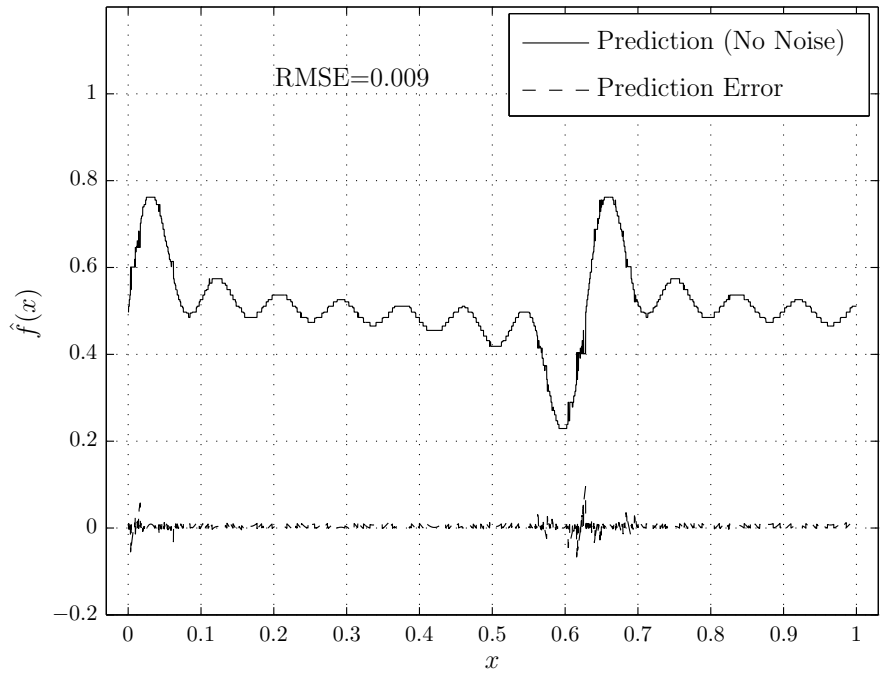
Figure 3.3 shows the results of training in both test cases. To assess what transpired during training the trained ART^a categories (Figure 3.3 (a) and (c)) as well as the map field weights (Figure 3.3 (b) and (d)) are observed graphically.

During training in the first test case 96 ART^a categories were formed. These categories are shown as 2-dimensional rectangles in Figure 3.3 (a) spanning the range of each dimension relative to the

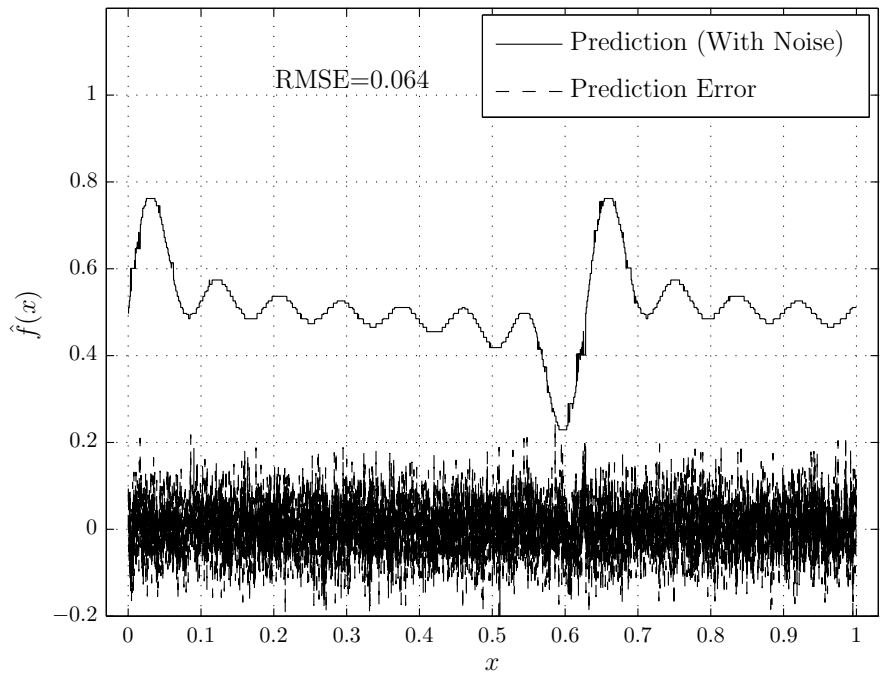
decision boundary (the circle in Figure 3.3 (a)). Here, the size of these rectangles are limited and more categories had to be created to form a mapping of the training data onto labelled output data (or classification decisions). This mapping is defined by the map field and is shown in Figure 3.3 (b). The 96 ART^a categories maps onto 2 ART^b categories. Chosen ART^b categories directly lead to classification decisions. The third ART^b category shown Figure 3.3 (b) contain only zeros. These zeros are attributed to the classifier being limited to only two possible decisions, thus causing the third ART^b category to be redundant. The performance of the classifier when presented with evaluation data is 94.51% Correct Classification Decisions (CCD).

In the second test case only 27 ART^a categories were formed. The categories shown in Figure 3.3 (c) are predominantly larger due to the lower vigilance specified. The map field weights are shown in Figure 3.3 (d). The performance of the classifier when presented with evaluation data in the second test case is 88.62% CCD.

To conclude this example, the Fuzzy ARTMAP classifier's ability to generalise a classification problem is controlled by the vigilance parameter ρ in both the ART modules. A higher vigilance specification forces the classifier to consider every possible detail of training data for resonance to occur (from Equation (3.9)) at the expense of allowing more categories to be formed. The benefit of using high vigilance parameter values is ensure more accuracy during data evaluation. On the contrary, lower vigilance cause greater generalisation, which result in less categories being formed at the expense of classification accuracy. Figure 3.4 shows classifier performance with variable vigilance values. Although the range of classification accuracy is only 6%, it shows improvement when vigilance is larger.

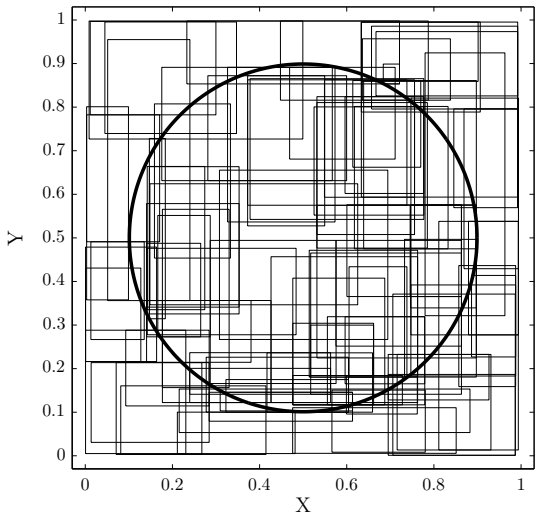


(a)

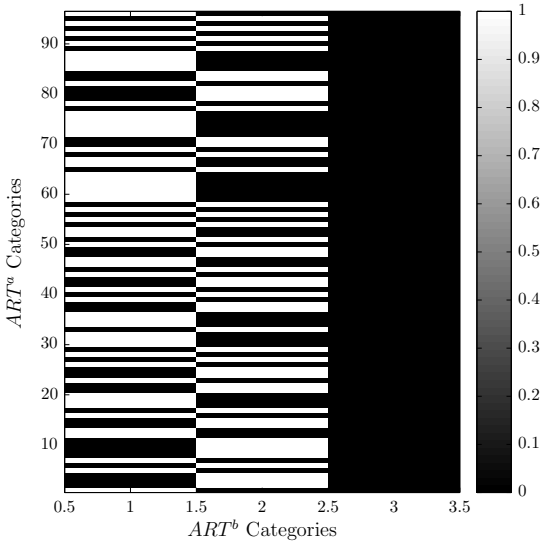


(b)

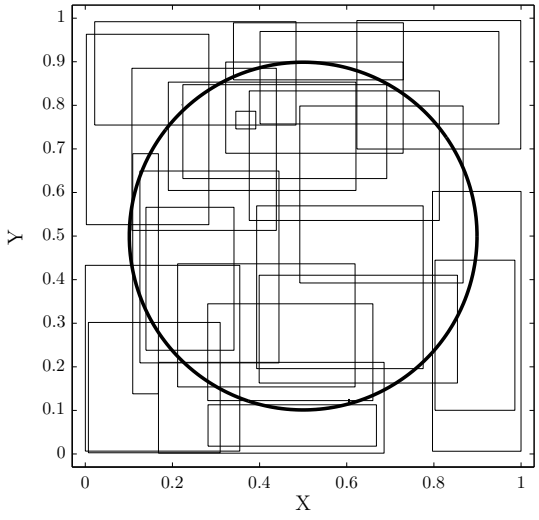
Figure 3.2: Nonlinear function approximation with Fuzzy ARTMAP. (a) Function approximation with no noise $\varepsilon = 0$, and (b) function approximation with added noise $\varepsilon = \mathcal{N}(0, 0.02)$.



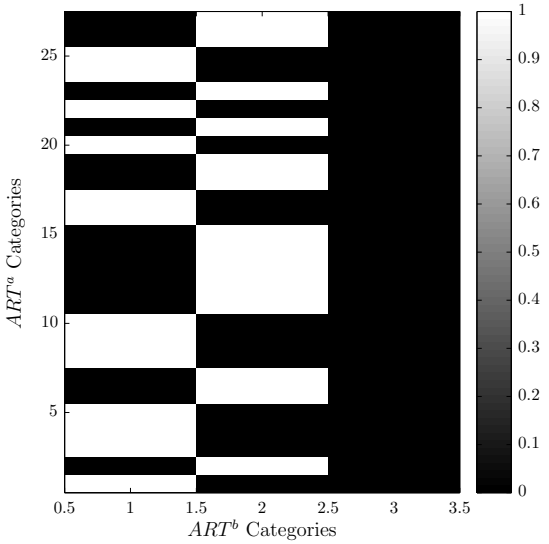
(a) ART^a categories with $\rho^a = 0.8$.



(b) Map field weights with $\rho^a = 0.8$.



(c) ART^a categories with $\rho^a = 0.1$.



(d) Map field weights with $\rho^a = 0.1$.

Figure 3.3: Fuzzy ARTMAP training on circle in the square data. (a) ART^a categories relative to the decision boundary after training with high vigilance, (b) map field weight values after training with high vigilance, (c) ART^a categories relative to the decision boundary after training with low vigilance, and (d) map field weight values after training with low vigilance.

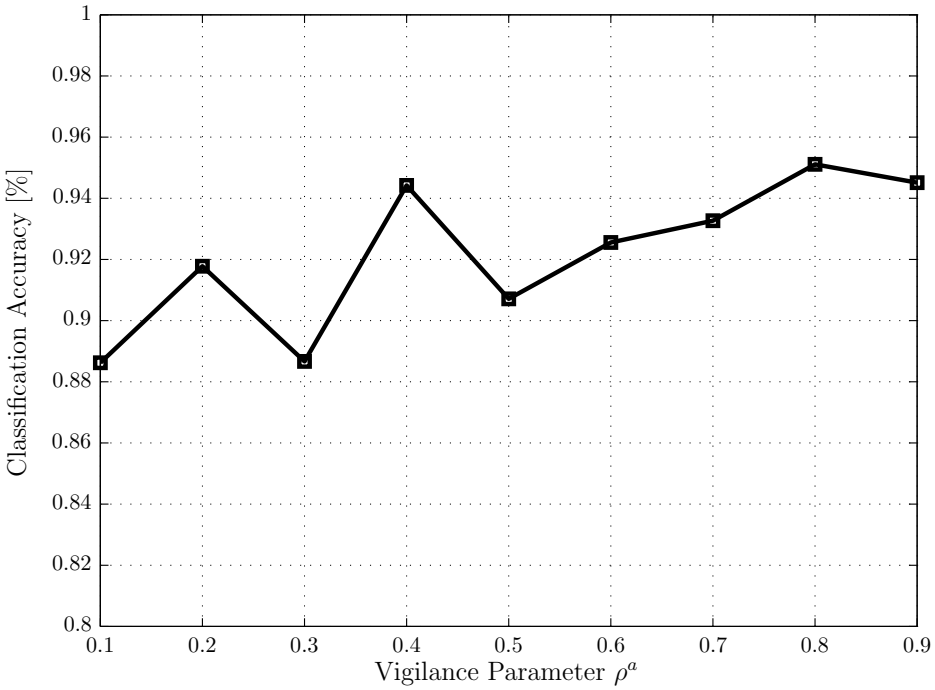


Figure 3.4: Fuzzy ARTMAP classifier performance versus vigilance.

CHAPTER 4

CLASSIFICATION SCENARIO

4.1 TARGET RADAR

In this section a practical LPI radar that can be analysed for the purposes of this study is introduced. On page 56 of [6] there is a list of 25 radar systems that posses LPI characteristics and behaviour. However, because of the tactical advantages of fielding LPI systems there are very little detail information available, at least in the open literature. Fortunately, a suitable candidate is also reported in [6] and known as the Pilot LPI radar. The Pilot is a product of the Swedish-based company SAAB and has been the subject of published analyses with regard to its detectability from an EW point of view [41, 50]. It is therefore a logical subject choice for this work. According to product information from the manufacturer [51],

> *The PILOT is a Low Probability of Intercept (LPI) navigation and threat detection radar with multiple applications. It has better range resolution and close range performance than a pulsed navigation radar with the same detection range.*

Table 4.1 lists the important parameters of the Pilot Mk3 version used in this study. A more complete list of other Pilot radar parameters are presented on page 30 of [6]. From this table two options for rotation rate, four options for transmitter output power and six transmit bandwidth options are identified. The rotation rate will determine the duration or dwell of the antenna main beam to an intercept receiver. Assuming a static land-based intercept scenario, this duration would be 8.33 ms and 4.16 ms for the 24 rpm and 48 rpm rotation rates respectively. The transmitted output power options will limit the range at which the intercept receiver will be able to successfully detect the radar. The six options in transmitted bandwidth provides the greatest diversity and combined with antenna rotation rate, the Pilot radar has twelve configurations from which it can be classified by an intercept receiver within the Pilot radar's 9.1 to 9.5 GHz transmission band.

Table 4.1: Important Pilot Mk3 parameters. (Reproduced with permission from [6]. © 2004 by Artech House, Inc.)

Subsystem	Parameter	Parameter Value
Antenna	Antenna Gain	30 dB
	Peak Sidelobe Level	$<$ -25 dB
	Mean Sidelobe Level	$<$ -30 dB
	Horizontal Beamwidth (3 dB)	1.2 deg
	Vertical Beamwidth (3 dB)	20 deg
	Rotational Speed	24, 48 rpm
	Polarisation	Horizontal
Transmitter	Output Power	1, 0.1, 0.01, 0.001 W
	Frequency Range	9.1–9.5 GHz
	Frequency Sweep	1.7, 3.4, 6.8, 13.75, 27.5, 55 MHz
	Sweep Repetition Frequency	1 kHz
Receiver	IF Bandwidth	512 KHz
	Noise Figure	5 dB

4.2 INTERCEPT RECEIVER

Given the target (or threat) radar specified in Section 4.1 we need to define a suitable intercept receiver to intercept, detect and classify it in order to perform the Electronic Support role of EW as illustrated in Figure 2.1. In order to provide this receiver with the best possible chance of intercepting the Pilot Mk3 radar, the parameters of such an intercept receiver (shown in Table 4.2) is identified in the same fashion as in Section 4.1. In subsequent sections suitable detection, parameter estimation and classification mechanisms are defined to enable the classification of the Pilot LPI radar using Fuzzy ARTMAP.

4.2.1 Signal detection

With reference to the ES system illustrated in Figure 1.1, it needs to detect the presence of any intercepted signals before any classification decisions can be made. In this section a suitable detector to notify a user of the presence of an intercepted emitter signal is developed. The well known Maximum Likelihood (ML) detector is used.

The choice between two hypothesis $\mathscr{H}_0$ and $\mathscr{H}_1$ with regard to the absence or presence of an intercep-

Table 4.2: Intercept receiver parameters.

Subsystem	Parameter	Parameter Value
Antenna	Antenna Type	Omni-directional
	Antenna Gain	0 dBi
	Polarisation	Horizontal
Receiver	Receiver Type	Digital
	Sample Rate	210 MHz
	Sample Resolution	12 bits
	IF Bandwidth	90 MHz
	Noise Figure	8 dB
	Dynamic Range	60 dB

ted emitter is considered. When $\mathscr{H}_0$ is true there is no emitter present and conversely when $\mathscr{H}_1$ is true an emitter is present. Practical systems are always faced with some form of interference. The major source of interference in ES receivers is that of thermal noise. Thermal noise is well described as a Gaussian random variable. The intercept receiver gathers signal samples, $s[n]$, so that each hypotheses can be described as

$$\begin{aligned} \mathscr{H}_0 : &\quad s[n] = w[n] \\ \mathscr{H}_1 : &\quad s[n] = x[n] + w[n] \, . \end{aligned} \tag{4.1}$$

Here, $x[n]$ represents the known signal of interest, $w[n]$ represents the thermal noise interference with mean, μ and variance σ^2. Samples are indexed with the variable $n = [0,1,2,\ldots,N-1]$ so that the signal $s[n]$ is a series of samples with length N. $s[n]$ can be expressed as a column vector, $\mathbf{s}$. The likelihood of the signal for each of the given hypothesis is

$$p(s[n]|\mathscr{H}_0) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1} s^2[n]\right] \tag{4.2}$$

and

$$p(s[n]|\mathscr{H}_1) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1} (s[n]-x[n])^2\right] \tag{4.3}$$

respectively. The detector is faced with a decision if either $\mathscr{H}_0$ or $\mathscr{H}_1$ is true. The hypothesis with the greater likelihood is chosen according a the likelihood ratio as

$$\Lambda(\mathbf{s}) = \frac{P(\mathscr{H}_1)\, p(s[n]|\mathscr{H}_1)}{P(\mathscr{H}_0)\, p(s[n]|\mathscr{H}_0)} \, . \tag{4.4}$$

Here, $P(\mathscr{H}_0)$ and $P(\mathscr{H}_1)$ are prior probabilities of hypotheses $\mathscr{H}_0$ and $\mathscr{H}_1$ respectively. This merely weights the likelihood of each hypothesis according to *a priori* knowledge about each hypothesis.

For example, if for some reason the intercept receiver expects to intercept signals 90% of the time within a given time period the prior probabilities would be $P(\mathscr{H}_0) = 0.1$ and $P(\mathscr{H}_1) = 0.9$. With the inclusion of prior knowledge the same detector can be adjusted in a dynamic situation and still be useful. The likelihood ratio is written as

$$\Lambda(\mathbf{s}) = \frac{P(\mathscr{H}_1)}{P(\mathscr{H}_0)} \exp\left[\frac{1}{\sigma^2}\sum_{n=0}^{N-1}(s[n]x[n]) - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}x^2[n]\right] . \tag{4.5}$$

The optimal decision making occurs where $\Lambda(\mathbf{s}) = 1$, so that the decision for either hypothesis is

$$\Lambda(\mathbf{s}) \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} 1 . \tag{4.6}$$

Equations (4.5) and (4.6) are combined so that

$$\frac{P(\mathscr{H}_1)}{P(\mathscr{H}_0)} \exp\left[\frac{1}{\sigma^2}\sum_{n=0}^{N-1}(s[n]x[n]) - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}x^2[n]\right] \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} 1 . \tag{4.7}$$

From Equation (4.7) the exponent is easily removed by expressing the likelihood ratio equation in logarithmic terms as

$$\sum_{n=0}^{N-1}(s[n]x[n]) \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \sigma^2 \ln\frac{P(\mathscr{H}_0)}{P(\mathscr{H}_1)} + \frac{1}{2}\sum_{n=0}^{N-1}x^2[n] . \tag{4.8}$$

A test function, $D(\mathbf{s})$ is specified that is equal to the left hand side of Equation (4.8) as

$$D(\mathbf{s}) = \sum_{n=0}^{N-1}(s[n]x[n]) . \tag{4.9}$$

The test function suggests a form of correlation between the received signal, $s[n]$ and the expected signal, $x[n]$ in a matched filter sense. Detection of a signal for a given sequence of received samples follows this matched filter process. The probability of detection quantifies the detector's ability to correctly detect signal, $x[n]$ from the result of $D(\mathbf{s})$. The two hypothesis presented in Equation (4.2) and (4.3) has four possible outcomes.

1. The decision that a signal is received when in fact a signal is present ($\mathscr{H}_1$ is true), which leads to a correct detection.

2. The decision that a signal is received when only interference is present ($\mathscr{H}_0$ is true), which leads to a false alarm.

3. The decision that no signal is received when in fact a signal is present ($\mathscr{H}_1$ is true), which leads to a missed detection.

4. The decision that no signal is received when only interference is present ($\mathscr{H}_0$ is true).

By considering the density functions of the test function, $D(\mathbf{s})$ given both hypotheses, the probability of detection is defined as

$$P_D = P(\mathscr{H}_1)\, p(D(\mathbf{s}) > \gamma | \mathscr{H}_1) \tag{4.10}$$

$$= 1 - \int_{-\infty}^{\gamma} P(\mathscr{H}_1)\, p(D(\mathbf{s}) | \mathscr{H}_1) \tag{4.11}$$

and the probability of false alarm is defined as

$$P_{FA} = P(\mathscr{H}_0)\, p(D(\mathbf{s}) > \gamma | \mathscr{H}_0) \tag{4.12}$$

$$= 1 - \int_{-\infty}^{\gamma} P(\mathscr{H}_0)\, p(D(\mathbf{s}) | \mathscr{H}_0)\,. \tag{4.13}$$

The performance of the detector can now be specified starting with the probability of false alarm, P_{FA}. This false alarm probability leads to the threshold γ, which in turn will result in the expected probability of detection, P_D for the detector. The distribution of the test function, $D(\mathbf{s})$ is required to calculate this decision threshold. It is known that each of the original hypotheses are Gaussian and that any function of these Gaussian distributions are again Gaussian [52]. Thus, the distribution of the test function can be described sufficiently by its expected (or mean) value and variance parameters subject to each hypothesis. For hypothesis, $\mathscr{H}_0$ the expected value $\mathrm{E}(D(\mathbf{s})|\mathscr{H}_0)$ and variance $\mathrm{var}(D(\mathbf{s})|\mathscr{H}_0)$ are

$$\mathrm{E}(D(\mathbf{s})|\mathscr{H}_0) = \mathrm{E}\left(\sum_{n=0}^{N-1}(w[n]x[n])\right) = 0 \tag{4.14}$$

$$\mathrm{var}(D(\mathbf{s})|\mathscr{H}_0) = \mathrm{var}\left(\sum_{n=0}^{N-1}(w[n]x[n])\right) \tag{4.15}$$

$$= \sum_{n=0}^{N-1}\mathrm{var}(w[n])x^2[n] \tag{4.16}$$

$$= \sigma^2 \sum_{n=0}^{N-1} x^2[n] \tag{4.17}$$

$$= \sigma^2 \varepsilon\,. \tag{4.18}$$

For hypothesis, $\mathscr{H}_1$ the expected value $\mathrm{E}(D(\mathbf{s})|\mathscr{H}_1)$ and variance $\mathrm{var}(D(\mathbf{s})|\mathscr{H}_1)$ are

$$\mathrm{E}(D(\mathbf{s})|\mathscr{H}_1) = \mathrm{E}\left(\sum_{n=0}^{N-1}(x[n]+w[n])x[n]\right) = \varepsilon \tag{4.19}$$

$$\mathrm{var}(D(\mathbf{s})|\mathscr{H}_1) = \mathrm{var}\left(\sum_{n=0}^{N-1}(x[n]+w[n])x[n]\right) \tag{4.20}$$

$$= \sum_{n=0}^{N-1}\mathrm{var}(w[n])x^2[n] \tag{4.21}$$

$$= \sigma^2\sum_{n=0}^{N-1}x^2[n] \tag{4.22}$$

$$= \sigma^2\varepsilon\,. \tag{4.23}$$

Here, ε refers to the energy in signal $x[n]$. Thus, the test function is a Gaussian distribution under each hypothesis as follows

$$D(\mathbf{s}) = \begin{cases} \mathscr{N}(0,\sigma^2\varepsilon), & \text{under } \mathscr{H}_0 \\ \mathscr{N}(\varepsilon,\sigma^2\varepsilon), & \text{under } \mathscr{H}_1\,. \end{cases} \tag{4.24}$$

From Equation (4.13) the probability of false alarm is related to the Gaussian Cumulative Distribution Function (CDF), which can be used to calculate the threshold γ according to a function Q and its inverse [24]. The Q function is defined as

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt \tag{4.25}$$

$$= 1-\Phi(x)\,. \tag{4.26}$$

Here, $\Phi(x)$ is the CDF of a Gaussian random variable $\mathscr{N}(0,1)$. Using Equation (4.11) and (4.13) the probability of false alarm and detection can be rewritten as

$$P_{FA} = Q\left(\frac{\gamma}{\sqrt{\sigma^2\varepsilon}}\right) \tag{4.27}$$

and

$$P_D = Q\left(\frac{\gamma-\varepsilon}{\sqrt{\sigma^2\varepsilon}}\right) \tag{4.28}$$

respectively. The threshold γ from Equation (4.27) is calculated as

$$\gamma = \sqrt{\sigma^2\varepsilon}\,Q^{-1}(P_{FA})\,. \tag{4.29}$$

By substituting Equation (4.29) into Equation (4.28) the probability of detection is

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{\varepsilon}{\sigma^2}}\right)\,. \tag{4.30}$$

Figure 4.1 (a)–(b) shows the performance curves of the matched filter detector developed above. These curves provide a concise view of the critical detector design parameters and their dependence on each other. The second term in the argument of Equation (4.30) refers to the ratio of signal energy and noise variance. It is referred to as the Energy-to-Noise Ratio (ENR). This ratio supplements design decisions with regard to how much signal energy is required in order to achieve the desired P_D for a given P_{FA}. In Figure 4.1 (b) P_D and P_{FA} are combined in what is commonly referred to as the Receiver Operating Characteristic (ROC).

Using a matched filter detector has advantages compared to a single sample detector. By calculating the Processing Gain (PG) of the matched filter detector relative to a single sample detector this potential advantage is quantified. Processing Gain is the ratio between the ENR of processed data (matched filter) and the ENR of unprocessed data (a single sample). Equation (4.31) shows the ENR of a unit amplitude single sample detector and Equation (4.35) the ENR for the matched filter detector with regard to a unit amplitude sinusoidal signal.

$$\eta_{in} = \frac{\varepsilon}{\sigma^2} \tag{4.31}$$

$$= \frac{1^2}{\sigma^2} \tag{4.32}$$

$$\eta_{out} = \frac{\varepsilon}{\sigma^2} \tag{4.33}$$

$$= \frac{1}{\sigma^2}\sum_{n=0}^{N-1} s^2[n] \tag{4.34}$$

$$= \frac{N}{2\sigma^2} \tag{4.35}$$

The processing gain is

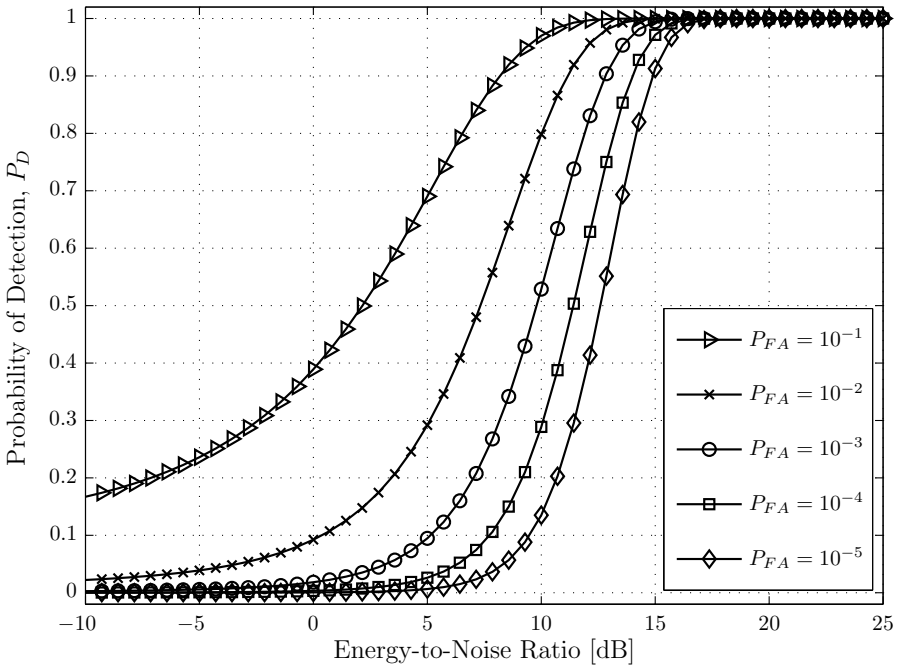
$$PG = \frac{\eta_{out}}{\eta_{in}} \tag{4.36}$$

$$= \frac{N}{2}\,. \tag{4.37}$$

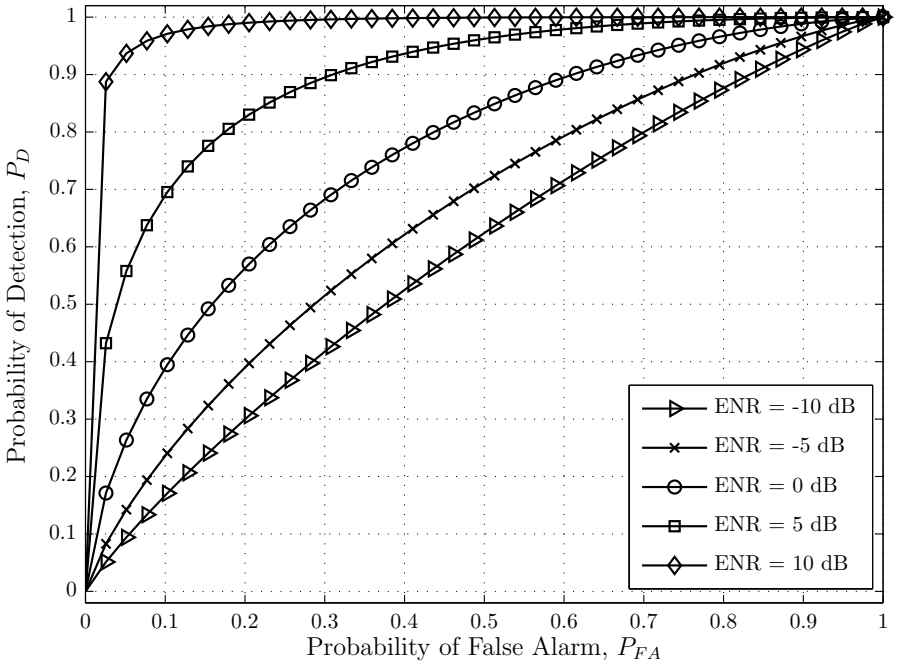
For example, from Figure 4.1 (a) for a $P_{FA} = 10^{-5}$ the detector would require an ENR of 13 dB to achieve a $P_D = 0.55$. If the detector is required to have a $P_D = 0.9$, the ENR must be 15 dB. By increasing the number of samples, N of the matched filter by a factor of 1.58, the ENR is increased by the required 2 dB.

4.2.2 Parameter estimation

Estimation problems are very similar to detection problems. With estimation the aim is to determine parameters (known or unknown) of interest that is uniquely related to a process of interest. The differ-



(a)



(b)

Figure 4.1: Matched filter detector performance curves. (a) Describes the probability of detection P_D versus Energy-to-Noise Ratio for a given probability of false alarm P_{FA}, and (b) describes the Receiver Operating Characteristic for a given Energy-to-Noise ratio.

ence between detection and estimation is that detection is a mechanism to notify us of the existence or the absence of a signal. This is done by decision making based on a given criteria. Estimation usually follows detection and aims to describe a process by determining the values of parameters that describe it. In this study the process of interest is LPI radar signals.

With parameter estimation, results are very seldom exactly correct and therefore the aim is to be nearly correct most of the time [53]. Another distinction between detection and estimation are the performance criteria used for each stage. The detector presented in Section 4.2.1 emphasised performance by selecting a threshold (decision boundary) for a given probability of false alarm, P_{FA}. These false alarms are related to the probability of detection P_D in the detector. In parameter estimation, we start by assuming that a detection has been made and with goal of describing parameters of interest as accurately as possible. Here, the performance criteria could be something like the Mean Square Error (MSE) or the absolute error. The difference (or error) between the true value of a parameter θ and its estimate, $\hat{\theta}$ is

$$\theta_\varepsilon = \hat{\theta} - \theta \, . \tag{4.38}$$

The criteria (or cost) of the MSE is

$$C(\theta_\varepsilon) = \frac{1}{N} \sum^{N} \theta_\varepsilon^2 \, , \tag{4.39}$$

and the criteria for the absolute error is

$$C(\theta_\varepsilon) = |\theta_\varepsilon| \, . \tag{4.40}$$

The cost of estimation may also be defined to fall within a specified range, Δ such that

$$C(\theta_\varepsilon) = \begin{cases} 0, & \text{if} \quad |\theta_\varepsilon| \leq \frac{\Delta}{2} \\ 1, & \text{if} \quad |\theta_\varepsilon| > \frac{\Delta}{2} \, . \end{cases} \tag{4.41}$$

Before a particular parameter estimation technique to describe LPI radar signals is chosen, the feasibility of estimating the parameters of the target radar from Table 4.1 is discussed. The discussion includes a brief evaluation of the need to estimate the values of a particular radar parameter. Selecting appropriate radar parameters are quite important when considering the classification of radar. Some parameters do reveal a particular radar and others are very general and appear to have been the natural (or obvious) choice for radar designers and manufacturers. For example, there are several types of radar operating in the designated X-band (8.5 to 10.5 GHz) of frequencies [54]. If we are to classify a very particular radar such as a LPI one, we will need some means to distinguish estimated frequencies from other in-band navigation, search and tracking radar systems. A sensible approach is to estimate only those parameters that best isolate the target radar from others in order to improve

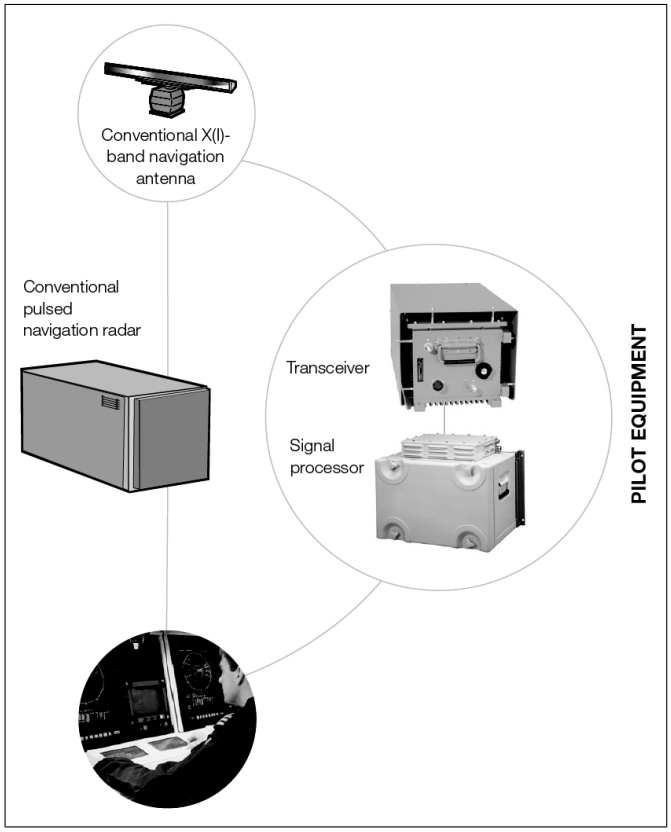


Figure 4.2: Additional Pilot radar subsystem components to add LPI capabilities to a conventional navigation radar. (From [51]. © SAAB Group.)

classification accuracy. The Pilot Mk3 radar is quite interesting in this regard. It is merely a radar add-on system, integrating into a conventional pulsed navigation radar giving it LPI capabilities as shown in Figure 4.2.

Evaluation of suitable parameters originating from the antenna subsystem in the Pilot Mk3 radar are as follows;

Antenna gain. Estimating the actual value of this parameter is not really useful. Even so, the estimator is required to somehow *measure* the entire antenna radiation pattern and find the peak value relative to sidelobes. The reported value of 30 dB is extremely common in radar antennas and thus do not provide us with a clear ability to discriminate the Pilot amongst other radars.

Sidelobe levels. Similarly, the antenna sidelobes are also not useful. It is quite an important parameter to the radar designer, but to the intercept receiver it suggests only difficult detection prospects beyond the antenna main beam.

Antenna beamwidth. Given that the Pilot radar assists an operator with navigation, it scans its an-

tenna repeatedly in the horizontal (azimuth) plane. A surface-based intercept receiver (as in Figure 2.4) will be illuminated by the radar main beam during every rotation. The azimuth beamwidth combined with the rotation rate of the antenna provide information on the duration of main beam illumination an intercept receiver can expect. Given that signal levels are reduced significantly before and after the antenna main beam passes over an intercept receiver, this parameter could be useful. The intercept receiver will need to know the antenna rotation rate a priori in order to estimate the azimuth beamwidth. Estimating the vertical beamwidth of the Pilot is somewhat different to the horizontal dimension. The Pilot only processes return signals in two dimensions (range and azimuth) but illuminate the vertical (elevation) dimension as much as possible to detect elevated as well as surface targets. In many cases the broad vertical coverage (20 degrees) is a consequence of a priority requirement to have as narrow as possible azimuth main beam. For an intercept receiver to estimate the vertical beamwidth of a radar it needs to cover the entire vertical plane (or altitude) while the main beam illuminates the receiver. This is clearly impractical even for airborne intercept receivers.

Rotational speed. As with the antenna azimuth beamwidth, an intercept receiver may estimate the rotation rate of the radar antenna given that observation periods are large enough to include at least two passes of the antenna main beam. This is a potentially useful parameter to estimate.

Polarisation. Polarisation defines the orientation of the electric field of an electromagnetic wave with respect to its propagation direction. The choice of polarisation in radar is somewhat determined by the physical operating environment. Naval radar tend to be horizontally polarised because of better detection performance in sea clutter [55], which does not really distinguish LPI radar from other radars. Radar designers are more concerned with polarisation purity in antennas as it quantifies the leakage of energy into the cross-polarised regions of the antenna [56]. Polarisation mismatch on return signals are severely attenuated and less return power reduce target detectability. Equivalently, intercept receivers also suffer from polarisation mismatch. There is a comprehensive explanation of polarisation mismatch (or loss) given the choice of intercept receiver antenna when faced with dissimilar transmit polarisation in [42]. The main reason why practical intercept receivers use spiral antennas is evident from the aforementioned explanation. The estimation of radar transmit polarisation is therefore not useful unless the intercept receiver can measure multiple polarisations.

Parameter estimation originating from the transmitter subsystem in the Pilot Mk3 radar are as follows;

Output power. The Pilot radar has four possible power settings ranging from 1 mW up to 1 W. The estimation of transmitted power seem useful as there are not many operational radar systems that employ such low transmit power. When they do, it is safe to assume LPI behaviour. Unfortunately the potential of estimating transmit power ends at this point. Accurately estimating radar transmit power is almost always a matter of academic interest. Propagation in realistic environments always influence the amount of signal power intercepted by an intercept receiver. Whether it be multipath, clutter from sea or land or rain, transmitted power does not distinguish any particular radar uniquely. Although the equations with regard to LPI performance (from Section 2.2.1) assume uniform propagation effects, it quantifies the best case expected behaviour.

Frequency range. From the illustrative example earlier the fact that the Pilot also operates in the designated X-band makes it all the more difficult to separate from other radars utilising similar frequencies. The 400 MHz bandwidth available to the Pilot radar also does not help either. If the radar does not transmit at frequencies spanning the full 400 MHz, an intercept receiver has no means of correctly classifying the Pilot from frequency estimates alone. The reciprocal argument to using transmit frequency might be more valuable. Estimated frequencies outside the 400 MHz operating band (from 9.1 to 9.5 GHz) almost certainly do not originate from the Pilot radar. This attribute supports the use of Fuzzy ARTMAP for classification by presenting data in complement coding form so that the classifier may build a reference of the parameters (features) that jointly include and excludes a certain class. Thus, transmit frequency may be a useful parameter after all.

Frequency sweep. The Pilot radar has six options available with regard to swept frequency bandwidth. Given that it is a FMCW radar that control range resolution by adjusting its swept frequency bandwidth, these bandwidths may provide enough diversity to separate the Pilot from other radars. Transmit frequency and swept frequency bandwidth can be estimated by transforming sampled data (post detection) into the frequency domain.

Sweep repetition frequency. The Pilot only has a single option available here, but this parameter is quite unique and is also relatively straight forward to estimate from the joint time-frequency domain.

To an intercept receiver, the parameters that describe the radar receiver is not of any value because its behaviour and inner workings are completely unknown. Thus, the estimation of radar receiver

parameters is not considered.

When viewed individually, the parameters discussed above may never point toward the Pilot radar because of the almost infinite amount of alternatives available. For example, the likelihood of an estimate of sweep repetition frequency alone correctly classifying the Pilot LPI radar is quite small. The situation changes considerably when a combined set of parameters are considered for classification. In this case if a sweep repetition frequency of 1 kHz, a swept frequency band of 27.5 MHz, a rotation rate of 24 rpm and a transmit frequency of 9.25 GHz is estimated, the likelihood of it belonging to the Pilot LPI radar is much higher. Additional parameters add more information relating to classifier outputs at the expense of dealing with more data. There is however a limit as to how much more additional parameters improves classification performance, and if not carefully managed, it may even reduce classification accuracy. Not only does the size of data become unmanageable, but parameters become correlated. Correlated parameters partially include information about other parameters and may render a classifier ineffective. Fortunately, the radar parameters described above are mostly uncorrelated or independent of each other and there is no at risk of presenting the Fuzzy ARTMAP classifier with correlated data.

From the consideration of useful radar parameters, antenna rotation rate, transmit frequency, frequency sweep and sweep repetition frequency parameters of the Pilot radar will be estimated. The following paragraphs present the processing required to estimate these parameters. Ultimately, instances of estimated parameters are combined into vector form as

$$\hat{\theta}^{(p)} = \left[\hat{\theta}_R, \hat{\theta}_F, \hat{\theta}_{\Delta F}, \hat{\theta}_S\right] . \tag{4.42}$$

Here, p is the index value for a particular instance of the estimation vector $\hat{\theta}$. The estimates of the antenna rotation rate are denoted with $\hat{\theta}_R$, transmit frequency estimates with $\hat{\theta}_F$, swept frequency bandwidth with $\hat{\theta}_{\Delta F}$ and sweep repetition frequency with $\hat{\theta}_S$.

Antenna rotation rate

The half-power (3 dB) horizontal beam of the Pilot radar is 1.2 degrees wide and the intercepted signal power rise and fall by 30 dB as the antenna passes over the intercept receiver antenna. By observing the result of the detector test function from Equation (4.9), the antenna rotation rate is easily estimated by calculating the duration of higher test function values. Next, the extent by which the detection threshold in Equation (4.29) must be crossed is specified to indicate that the radar main beam is passing over the intercept receiver. Figure 4.3 shows the threshold values versus signal energy for a given probability of false alarm. Another approach would be to quantify the range of the output

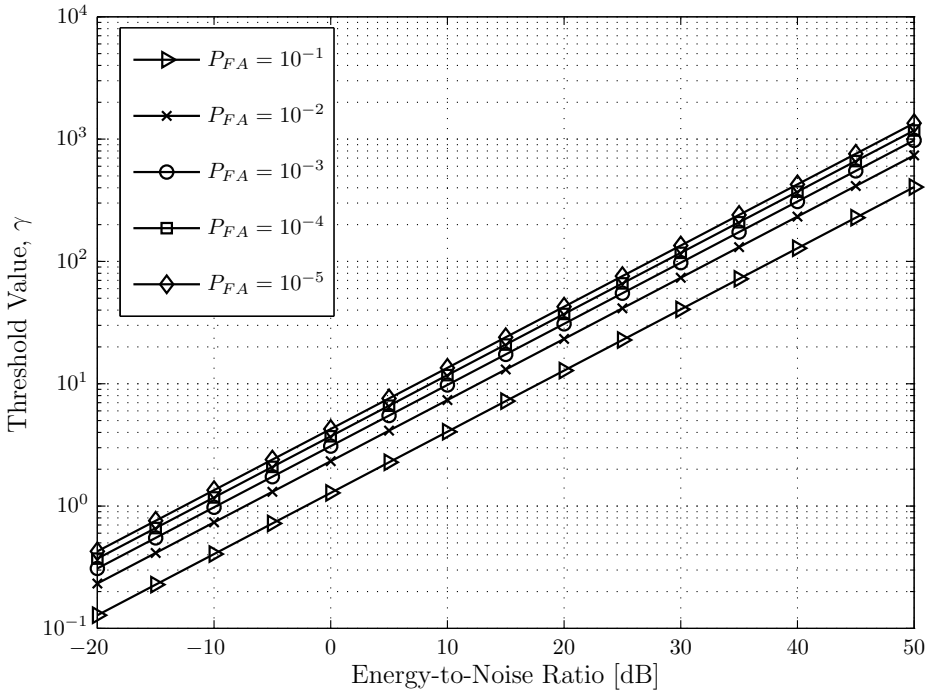


Figure 4.3: Detector threshold values versus Energy-to-Noise ratio for a given probability of false alarm.

values in the detector when the main beam and sidelobes pass across the intercept receiver. This is the difference between the largest expected detector response $D(\mathbf{s})^{max}$ and the threshold value γ that cause detections to be declared. Signal power is calculated as

$$P_x = \frac{1}{N}\sum_{n=0}^{N-1} x^2[n] \tag{4.43}$$

$$= \frac{\varepsilon}{N}\,. \tag{4.44}$$

It is assumed that the reference signal has unit power $P_x = 1$ with $\varepsilon = N$ relative to the mean sidelobe level of the radar antenna and the intercept receiver antenna gain is unity. Thus, the largest expected signal power relative to P_x is equal to G_R, which is the gain in the radar antenna main beam. The gain in terms of the detector test function (using Equation (4.9)) is described as

$$D(\mathbf{s}) = \sum_{n=0}^{N-1}(s[n]x[n]) \tag{4.45}$$

$$= \sum_{n=0}^{N-1}(\tilde{x}[n]+w[n])x[n] \tag{4.46}$$

$$= A\sum_{n=0}^{N-1} x^2[n] + \frac{1}{A}\sum_{n=0}^{N-1} w[n]x[n]\,. \tag{4.47}$$

Here, the received radar signal is denoted by $\tilde{x}[n]$ and the additional gain of the radar main beam is A such that $\tilde{x}[n] = Ax[n]$. The original test function is then modifed by a factor A that can be expressed in terms of the received radar signal power as

$$\begin{aligned} AD(\mathbf{s}) &= A^2 \sum_{n=0}^{N-1} x^2[n] + \sum_{n=0}^{N-1} w[n]x[n] && (4.48)\\ &= P_{\tilde{x}} N + \sum_{n=0}^{N-1} w[n]x[n] && (4.49)\\ &= G_R P_x N + \sum_{n=0}^{N-1} w[n]x[n] && (4.50)\\ &= G_R \sum_{n=0}^{N-1} x^2[n] + \sum_{n=0}^{N-1} w[n]x[n] \, . && (4.51) \end{aligned}$$

Thus, if the original test function is modified by a factor A, the detector has a representation in terms of the radar antenna gain G_R. If the intercept receiver is able to only detect the radar when its main beam is passing over it, the antenna rotation rate is estimated by counting the number of detector output samples greater or equal to $A\gamma$ as

$$\hat{\theta}_R = W_H \left(\frac{60}{360} \right) \left(\frac{1}{T_S N_{A\gamma}} \right) \, . \tag{4.52}$$

Here, $\hat{\theta}_R$ is in units of revolutions per minute, $W_H = 1.2^\circ$ is the antenna horizontal beamwidth in degrees, T_S is the sample interval of the intercept receiver, and $N_{A\gamma}$ is the number of samples above $A\gamma$ at the output of the detector. From Equation (4.52) it should be noted that $\hat{\theta}_R$ is null ($\hat{\theta}_R = \emptyset$) when $N_{A\gamma} = 0$.

Transmit frequency and frequency sweep

The estimation of radar transmit frequency is easily done by transforming the sequence of N samples output from the detector into the frequency domain. The detector from Equation (4.46) operates using signal samples acquired in the time domain. The detector performs a correlation between a template signal $x[n]$ and the received signal $s[n] = \tilde{x}[n] + w[n]$, which contain a modified version of the template signal $\tilde{x}[n]$ with additive noise $w[n]$. The correlation is realised by convolving the template signal with the received signal. Correlation in the time domain is equivalent to multiplication in frequency domain [57]. The estimation of transmit frequency is ideally performed in the frequency domain. For a FMCW radar the theoretical double sided spectrum $X(f)$ is defined to be

$$X(f) = \left[\frac{M}{2} \Pi \left(\frac{f - f_c}{\Delta f} \right) + \frac{M}{2} \Pi \left(\frac{f + f_c}{\Delta f} \right) \right] \exp^{-j2\pi f \tau} \, . \tag{4.53}$$

Here, $\Pi(f)$ is a rectangular function with magnitude $M/2$ centred around frequency f_c having a swept bandwidth Δf. The duration of the signal is denoted by τ. The frequency transformation may be used

to determine the Power Spectral Density (PSD) of a signal. The estimate for transmit frequency (or centre frequency) and frequency sweep is

$$\theta_F = f_c \tag{4.54}$$

and

$$\theta_{\Delta F} = \Delta f \tag{4.55}$$

respectively. Both these parameters are contained in the magnitude spectrum of Equation (4.53), $|X(f)|$.

$$|X(f)| = \frac{M}{2}\Pi\left(\frac{f-f_c}{\Delta f}\right) + \frac{M}{2}\Pi\left(\frac{f+f_c}{\Delta f}\right) \tag{4.56}$$

It is assumed that the response of each frequency f_k is independent from any other component in $|X(f)|$. The joint likelihood function L of these two parameters given each frequency is

$$L \equiv L(\theta_F, \theta_{\Delta F}\,|f_1, f_2, \ldots, f_K) = \prod_{k=1}^{K} |X(f_k; \theta_F, \theta_{\Delta F})|\,. \tag{4.57}$$

The likelihood function can be further reduced to only describe half of the spectrum covered by $X(f)$. The maximum likelihood estimate of each parameter θ_l would be the point where the partial derivative of L with respect to each parameter is zero [13, 58]

$$\frac{\partial L}{\partial \theta_l} = 0\,. \tag{4.58}$$

By observation the magnitude spectrum is equivalent to a uniform distribution of frequency. It is further assumed that the spectrum is double sided and that one half thereof is repeated. Figure 4.4 (a) shows the probability density function $P_F(f)$ of a uniform distribution of the magnitude response with parameters a and b. An estimate of the transmit frequency would be the expected value of the distribution $P_F(f)$,

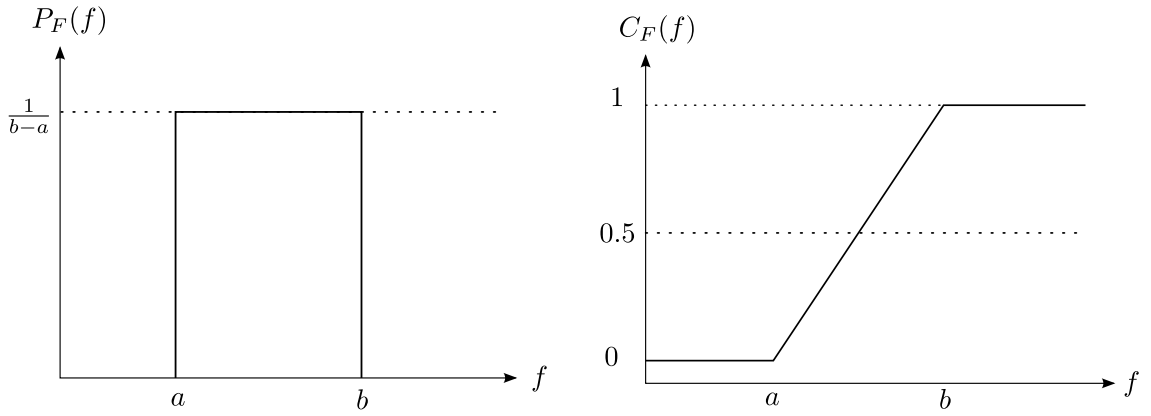
$$\hat{\theta}_F = \mathrm{E}[P_F(f)] \tag{4.59}$$

$$= \frac{a+b}{2}\,, \tag{4.60}$$

with the estimate for frequency sweep being

$$\hat{\theta}_{\Delta F} = b - a\,. \tag{4.61}$$

By estimating the parameters a and b of the uniform distribution in Figure 4.4 (a) estimates for transmit frequency and frequency sweep respectively are obtained. From Equation (4.57) the maximum likelihood estimation of parameters a and b does not seem tractable, because the frequencies f_k may



(a) (b)

Figure 4.4: Uniform distribution with parameters a and b. (a) Probability density function, and (b) probability mass (or cumulative distribution) function of a Uniform distribution.

be correlated. However, by considering the cumulative distribution $C_F(f)$ or probability mass function of $P_F(f)$ (shown in Figure 4.4 (b)), estimates for a and b can be calculated by integrating the density function $P_F(f)$ until a certain point g is reached. This point is determined by upper and lower frequency limits f_u and f_l respectively. The estimate for parameter a would then be $\hat{a} = f_l$,

$$\int_{-\infty}^{f_l} P_F(f)\,\mathrm{d}f \geq g_l\,. \tag{4.62}$$

Similarly, the estimate for b is

$$\int_{-\infty}^{f_u} P_F(f)\,\mathrm{d}f \leq g_h\,, \tag{4.63}$$

with $\hat{b} = f_u$. Figure 4.5 illustrates the principle of estimating parameters a and b from the cumulative probability density $C_F(f)$.

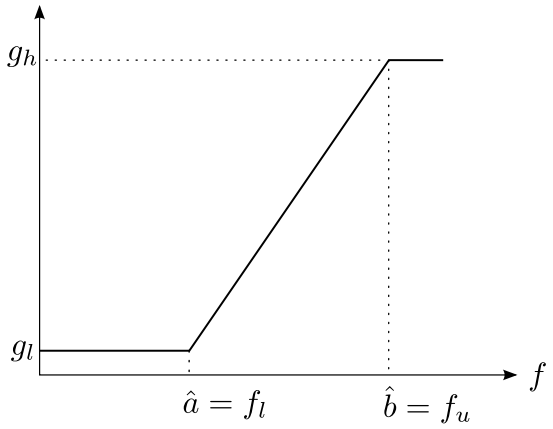


Figure 4.5: Parameter estimation of a uniform distribution.

Sweep repetition frequency

The rate at which frequency sweeps occur may also be estimated from the frequency domain representation of received signals. Here, the correlation delay τ is isolated from the phase spectrum of $X(f)$ as

$$\angle X(f) = -2\pi f \tau \,. \tag{4.64}$$

Here, the phase spectrum is assumed to be linear. Given that the duration of the correlation delay is N samples in length taken at regular known intervals T_s, the value of τ is a fraction of the correlation delay if the correlation sequence $x[n]$ is unmatched to the sweep repetition frequency. The delay is then expressed as

$$\tau = \frac{q}{\theta_S} \,. \tag{4.65}$$

Here, q is the fraction of the correlator duration NT_s to the fundamental signal duration $1/\theta_S$. For the Pilot radar the fundamental duration or sweep repetition frequency (from Table 4.1) is $\theta_S = 1$ kHz. By taking the derivative of Equation (4.64) with respect to frequency and substitute Equation (4.65) into it results in

$$\frac{d}{df}\angle X(f) = -2\pi\tau \tag{4.66}$$

$$= \frac{-2\pi q}{\theta_S} \,. \tag{4.67}$$

Now, the estimate of sweep repetition frequency is expressed as

$$\hat{\theta}_S = \left| -2\pi q \left(\frac{d}{df}\angle X(f) \right)^{-1} \right| \,. \tag{4.68}$$

Here, the derivative term may be bounded by the parameter estimates $\hat{a} = f_l$ and $\hat{b} = f_u$ used to determine transmit frequency and frequency sweep as described earlier. If we are able to obtain the slope of the phase spectrum in a linear least squares sense for example, the estimate for sweep repetition frequency is straight forward given our choice of q.

At this stage, estimated radar parameter vectors $\hat{\theta}^{(p)}$ can be compiled into what is generally referred to as feature vectors. In the next subsection the classification stage, on which results will be generated for this study, is described.

4.2.3 Classification

Given that the Pilot Mk3 LPI radar is the only type of LPI radar for which there is open access to critical technical data, feature vectors estimated earlier will be classified to belong to the Pilot radar.

Thus, the Pilot is indirectly classified by predicting its operating modes. For example, a radar is classified as the Pilot Mk3 when the estimate for antenna rotation rate is 48 rpm, transmit frequency is 9.25 GHz, frequency sweep is 27.5 MHz and the sweep repetition frequency is 1 kHz. Radar operating modes are defined from the antenna rotation rate and frequency sweep parameters, which relate to classifier output values and labels, which are shown in Table 5.25 on page 72.

With pre-defined classifier output values and labels the classifier may be trained via supervised learning. In supervised learning, data is presented as a pair consisting of an input vector and a desired output vector (also called the supervisory signal). During classifier evaluation only input feature vectors are presented with the classifier predicting output values. These output values are then compared to the actual output values.

CHAPTER 5

RESULTS

5.1 LPI PERFORMANCE FACTOR

With reference to the literature study presented in Section 2 there is a need to evaluate if the Pilot radar has low probability of intercept characteristics against the intercept receiver described in Section 4. In order to calculate the LPI performance factor, the Pilot radar parameters from Table 4.1 and the intercept receiver parameters from Table 4.2 are used. The simplified LPI performance factor as derived earlier, in terms of the intercept receiver range R_I and radar detection range R_R is

$$\chi^2 = \frac{R_I^2}{R_R^2} \,. \tag{5.1}$$

The Pilot radar has four maximum detection range selections namely, [0.75, 1.5, 3, 6, 12, 24] nmi. Intercept receiver range R_I is calculated (using Equation (2.22)) and the LPI performance factor χ is evaluated for each radar range selection. Here, the unaccounted loss in the intercept receiver is $L_I = 4$ and the radar operating frequency was chosen from the centre of the operating frequency range at 9.25 GHz. The evaluation also included scenarios of main beam and sidelobe illumination ($G_{RI} = [30, 0]$ dB). Two types of receiver detectors were chosen to illustrate the influence on system design decisions. The first detector is chosen to have a probability of false alarm of $P_{FA} = 10^{-5}$ and a probability of detection of $P_D = 0.55$ and thus requiring an Energy-to-Noise Ratio (ENR) greater or equal to 13 dB to declare detections. The second detector has the same false alarm probability but requires more energy into the receiver (greater or equal to 15 dB) to improve the probability of detection to $P_D = 0.9$.

LPI performance is shown for a chosen receiver against the four transmit power selections in the Pilot radar. Included in the evaluation is a measure of integration gain that the receiver need in order to achieve the required ENR for declaring detections. Larger integration times allow the receiver to be located at further ranges for a more favourable LPI performance factor. To re-iterate, an LPI per-

formance factor greater than 1 causes the intercept receiver to stand-off outside of the radar detection range whilst being able to detect the radar. With a performance factor below 1 the intercept receiver risks being detected by the radar in an attempt to perform its own detection duties within the radar range envelope.

5.1.1 Discussion

The LPI performance factor for each detector selection over all scenarios are presented in Tables 5.2–5.13.

Both intercept receiver detectors prove to be useful in all scenarios with very little difference in LPI performance. The detector with $P_D = 0.9$ is inevitably drawn closer to the radar as it requires a larger ENR. The difference in LPI performance factor really warrants using the $P_D = 0.9$ detector over a $P_D = 0.55$ detector.

The number of samples integrated in the detector provides some interesting insights. A single sample detector with $N = 1$ is equivalent to the greater majority of practical analogue intercept receiver systems, which at best made detection decisions based on single instances (samples) of received signals. Any consideration for more than one sample at the detector requires processing gain. The greater the processing gain of the intercept receiver the more favourable the LPI performance of course. This is true for both detectors in all scenarios. Ideally, correlator size should be an integer multiple of the radar sweep repetition interval. If the correlator size is smaller than a radar sweep interval integration is only effective for the fraction of the signal over which it correlates. The remaining portion of the potentially useful signal is merely ignored and will pass through the detector as an unmatched signal. Correlation sizes of 1,000, 10,000 do provide increasing LPI performance nevertheless. In order for the receiver to match correlation size to the sweep interval, the sample rate would have to be reduced. The sample rates to match a 1 ms sweep repetition interval to 1,000 and 10,000 correlator sizes would be 1 MHz and 10 MHz respectively. With these reduced sample rates the intercept receiver risk an insufficient instantaneous bandwidth when acquiring signals. It might be disabled completely as the smallest frequency sweep of the Pilot radar is 1.7 MHz. The choice of sampling at a rate of 210 MHz is then justified as it fully covers the radar frequency sweep and allowing a large enough correlator to have favourable LPI performance across all radar settings.

The Pilot radar's ability to reduce its transmit power sways the LPI performance in its favour. This is clearly seen in sidelobe intercept scenarios at the smallest maximum radar detection range of 0.75 nmi (Table 5.3). The detector also shows a *breakdown* in LPI performance from Table 5.8 and decreases at further radar ranges. Another interesting phenomenon regarding the transmit power variation in the

Table 5.1: Correlator size to achieve critical LPI performance ($\chi = 1$) with main beam and sidelobe illumination relative to a Pilot range of 24 nmi.

	Main beam Intercept		Sidelobe Intercept	
	$P_D = 0.55$	$P_D = 0.9$	$P_D = 0.55$	$P_D = 0.9$
P_R	ENR=13 dB	ENR=15 dB	ENR=13 dB	ENR=15 dB
0.001 W	30	48	30,556	48,429
0.01 W	3	4	3,055	4,842
0.1 W	1	1	305	484
1 W	1	1	30	48

Pilot radar is that integration gain in the intercept receiver makes up for lower radar transmit power. For example, the LPI performance (for any detector choice) when integrating a single sample when the radar transmits 1 W of power equals integrating 1,000 samples when the radar transmits 0.001 W.

As mentioned earlier traditional analogue intercept receivers are equivalent to a digital receiver integrating a single sample at a time. Ultimately, when choosing an intercept receiver against the Pilot radar consideration of the LPI performance factor from a worst case point of view must be analysed. Here, both the main beam and sidelobe scenarios are evaluated at the furthest radar detection range of 24 nmi in Tables 5.12 and 5.13. Given that the intercept receiver has no means of accurately determining the radar transmit power (which can vary by a factor of 1,000 for the Pilot) a receive detector will be disqualified when the LPI performance factor falls below 1 at any stage, irrespective of radar transmit power. In the main beam scenario there has to be some form of integration to achieve LPI performance greater than 1. Integration is mandatory in the sidelobe scenario, which disqualifies the single sample detector entirely. The amount of integration needed by each detector in order to achieve this critical value of $\chi = 1$ at 24 nmi is shown in Table 5.1. For the intercept receiver to be completely successful against the Pilot radar it must integrate in excess of 30,556 and 48,429 samples for detectors having $P_D = 0.55$ and $P_D = 0.9$ respectively.

Table 5.2: LPI performance factor χ with main beam illumination relative to a Pilot range selection of 0.75 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	5.79	18.31	57.89	183.06
1,000	0.55	183.06	578.89	1830.61	5788.89
10,000	0.55	578.89	1830.61	5788.89	18306.09
210,000	0.55	2652.80	8388.90	26528.04	83889.03
420,000	0.55	3751.63	11863.70	37516.31	118637.00
1	0.90	4.60	14.54	45.98	145.41
1,000	0.90	145.41	459.83	1454.10	4598.28
10,000	0.90	459.83	1454.10	4598.28	14541.04
210,000	0.90	2107.20	6663.54	21071.97	66635.42
420,000	0.90	2980.03	9423.67	29800.27	94236.72

Table 5.3: LPI performance factor χ with sidelobe illumination relative to a Pilot range selection of 0.75 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.18	0.58	1.83	5.79
1,000	0.55	5.79	18.31	57.89	183.06
10,000	0.55	18.31	57.89	183.06	578.89
210,000	0.55	83.89	265.28	838.89	2652.80
420,000	0.55	118.64	375.16	1186.37	3751.63
1	0.90	0.15	0.46	1.45	4.60
1,000	0.90	4.60	14.54	45.98	145.41
10,000	0.90	14.54	45.98	145.41	459.83
210,000	0.90	66.64	210.72	666.35	2107.20
420,000	0.90	94.24	298.00	942.37	2980.03

Table 5.4: LPI performance factor χ with main beam illumination relative to a Pilot range selection of 1.5 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	2.89	9.15	28.94	91.53
1,000	0.55	91.53	289.44	915.30	2894.45
10,000	0.55	289.44	915.30	2894.45	9153.04
210,000	0.55	1326.40	4194.45	13264.02	41944.51
420,000	0.55	1875.82	5931.85	18758.16	59318.50
1	0.90	2.30	7.27	22.99	72.71
1,000	0.90	72.71	229.91	727.05	2299.14
10,000	0.90	229.91	727.05	2299.14	7270.52
210,000	0.90	1053.60	3331.77	10535.99	33317.71
420,000	0.90	1490.01	4711.84	14900.13	47118.36

Table 5.5: LPI performance factor χ with sidelobe illumination relative to a Pilot range selection of 1.5 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.09	0.29	0.92	2.89
1,000	0.55	2.89	9.15	28.94	91.53
10,000	0.55	9.15	28.94	91.53	289.44
210,000	0.55	41.94	132.64	419.45	1326.40
420,000	0.55	59.32	187.58	593.19	1875.82
1	0.90	0.07	0.23	0.73	2.30
1,000	0.90	2.30	7.27	22.99	72.71
10,000	0.90	7.27	22.99	72.71	229.91
210,000	0.90	33.32	105.36	333.18	1053.60
420,000	0.90	47.12	149.00	471.18	1490.01

Table 5.6: LPI performance factor χ with main beam illumination relative to a Pilot range selection of 3 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	1.45	4.58	14.47	45.77
1,000	0.55	45.77	144.72	457.65	1447.22
10,000	0.55	144.72	457.65	1447.22	4576.52
210,000	0.55	663.20	2097.23	6632.01	20972.26
420,000	0.55	937.91	2965.93	9379.08	29659.25
1	0.90	1.15	3.64	11.50	36.35
1,000	0.90	36.35	114.96	363.53	1149.57
10,000	0.90	114.96	363.53	1149.57	3635.26
210,000	0.90	526.80	1665.89	5267.99	16658.86
420,000	0.90	745.01	2355.92	7450.07	23559.18

Table 5.7: LPI performance factor χ with sidelobe illumination relative to a Pilot range selection of 3 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.05	0.14	0.46	1.45
1,000	0.55	1.45	4.58	14.47	45.77
10,000	0.55	4.58	14.47	45.77	144.72
210,000	0.55	20.97	66.32	209.72	663.20
420,000	0.55	29.66	93.79	296.59	937.91
1	0.90	0.04	0.11	0.36	1.15
1,000	0.90	1.15	3.64	11.50	36.35
10,000	0.90	3.64	11.50	36.35	114.96
210,000	0.90	16.66	52.68	166.59	526.80
420,000	0.90	23.56	74.50	235.59	745.01

Table 5.8: LPI performance factor χ with main beam illumination relative to a Pilot range selection of 6 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.72	2.29	7.24	22.88
1,000	0.55	22.88	72.36	228.83	723.61
10,000	0.55	72.36	228.83	723.61	2288.26
210,000	0.55	331.60	1048.61	3316.01	10486.13
420,000	0.55	468.95	1482.96	4689.54	14829.63
1	0.90	0.57	1.82	5.75	18.18
1,000	0.90	18.18	57.48	181.76	574.79
10,000	0.90	57.48	181.76	574.79	1817.63
210,000	0.90	263.40	832.94	2634.00	8329.43
420,000	0.90	372.50	1177.96	3725.03	11779.59

Table 5.9: LPI performance factor χ with sidelobe illumination relative to a Pilot range selection of 6 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.02	0.07	0.23	0.72
1,000	0.55	0.72	2.29	7.24	22.88
10,000	0.55	2.29	7.24	22.88	72.36
210,000	0.55	10.49	33.16	104.86	331.60
420,000	0.55	14.83	46.90	148.30	468.95
1	0.90	0.02	0.06	0.18	0.57
1,000	0.90	0.57	1.82	5.75	18.18
10,000	0.90	1.82	5.75	18.18	57.48
210,000	0.90	8.33	26.34	83.29	263.40
420,000	0.90	11.78	37.25	117.80	372.50

Table 5.10: LPI performance factor χ with main beam illumination relative to a Pilot range selection of 12 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.36	1.14	3.62	11.44
1,000	0.55	11.44	36.18	114.41	361.81
10,000	0.55	36.18	114.41	361.81	1144.13
210,000	0.55	165.80	524.31	1658.00	5243.06
420,000	0.55	234.48	741.48	2344.77	7414.81
1	0.90	0.29	0.91	2.87	9.09
1,000	0.90	9.09	28.74	90.88	287.39
10,000	0.90	28.74	90.88	287.39	908.82
210,000	0.90	131.70	416.47	1317.00	4164.71
420,000	0.90	186.25	588.98	1862.52	5889.80

Table 5.11: LPI performance factor χ with sidelobe illumination relative to a Pilot range selection of 12 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.01	0.04	0.11	0.36
1,000	0.55	0.36	1.14	3.62	11.44
10,000	0.55	1.14	3.62	11.44	36.18
210,000	0.55	5.24	16.58	52.43	165.80
420,000	0.55	7.41	23.45	74.15	234.48
1	0.90	0.01	0.03	0.09	0.29
1,000	0.90	0.29	0.91	2.87	9.09
10,000	0.90	0.91	2.87	9.09	28.74
210,000	0.90	4.16	13.17	41.65	131.70
420,000	0.90	5.89	18.63	58.90	186.25

Table 5.12: LPI performance factor χ with main beam illumination relative to a Pilot range selection of 24 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.18	0.57	1.81	5.72
1,000	0.55	5.72	18.09	57.21	180.90
10,000	0.55	18.09	57.21	180.90	572.07
210,000	0.55	82.90	262.15	829.00	2621.53
420,000	0.55	117.24	370.74	1172.38	3707.41
1	0.90	0.14	0.45	1.44	4.54
1,000	0.90	4.54	14.37	45.44	143.70
10,000	0.90	14.37	45.44	143.70	454.41
210,000	0.90	65.85	208.24	658.50	2082.36
420,000	0.90	93.13	294.49	931.26	2944.90

Table 5.13: LPI performance factor χ with sidelobe illumination relative to a Pilot range selection of 24 nmi.

N	P_D ($P_{FA} = 10^{-5}$)	$P_R = 0.001$W	$P_R = 0.01$W	$P_R = 0.1$W	$P_R = 1$W
1	0.55	0.01	0.02	0.06	0.18
1,000	0.55	0.18	0.57	1.81	5.72
10,000	0.55	0.57	1.81	5.72	18.09
210,000	0.55	2.62	8.29	26.22	82.90
420,000	0.55	3.71	11.72	37.07	117.24
1	0.90	0.00	0.01	0.05	0.14
1,000	0.90	0.14	0.45	1.44	4.54
10,000	0.90	0.45	1.44	4.54	14.37
210,000	0.90	2.08	6.58	20.82	65.85
420,000	0.90	2.94	9.31	29.45	93.13

Table 5.14: The cost of estimating Antenna Rotation Rate versus Energy-to-Noise Ratio.

ENR [dB]	MSE	Percent MSE [%]	ABS Maximum [rpm]
	$\theta_R = 24$ rpm		
10	3.957	16.486	3.00
11	3.906	16.274	3.00
12	3.942	16.426	3.00
13	4.048	16.866	3.00
14	4.021	16.752	3.00
15	3.940	16.416	3.00
	$\theta_R = 48$ rpm		
10	5.486	11.428	4.00
11	5.598	11.662	4.00
12	5.712	11.901	4.00
13	5.630	11.729	4.00
14	5.535	11.530	4.00
15	5.443	11.340	4.00

5.2 RADAR PARAMETER ESTIMATION

Results from estimating the parameters of the Pilot Mk3 radar is presented in this section. The analysis of the parameter estimation stage in the intercept receiver is of great importance from an accuracy (or cost) point of view as it will directly impact the accuracy of the classification stage. Each of the four chosen radar parameters are analysed in the following subsections. The accuracy of estimation of any of the parameters are presented in terms of the Mean Square Error (MSE), the MSE as a percentage of the true parameter value and the absolute maximum estimation error. Each estimation is done at various detector Energy-to-Noise Ratio (ENR) values ranging from 10 to 15 dB.

5.2.1 Antenna rotation rate estimation

The estimation cost of antenna rotation rate for each of the two possible settings in the Pilot radar is given in Table 5.14. From the table there is no significant variation in estimation cost across the range of ENR values tested, with estimation of the 48 rpm rotation rate being slightly more accurate. Another important point to note with regard to antenna rotation rate estimation is that estimates are discrete integer values.

5.2.2 Transmit frequency estimation

The estimation cost of radar transmit frequency for each of the six possible modulation settings in the Pilot radar is given in Table 5.15 and 5.16 respectively. It is assumed that the true X-band transmit frequencies are down converted to Intermediate Frequencies (IFs) sampled at a rate of 210 MHz. Furthermore, down converted signals are relative to an IF start frequency of 10 MHz, where-after the estimation of transmit frequency is performed. Once again there is no significant variation in estimation cost across the range of ENR values tested per modulation setting. Interestingly, the estimation cost differs across modulation settings by factors of between 2 and 3.

5.2.3 Frequency sweep estimation

The estimation cost of radar frequency sweep in the Pilot radar is given in Table 5.17 and 5.18 respectively. There is a slight but gradual improvement in estimation accuracy with increasing ENR for the 1.7, 3.4 and 6.8 MHz frequency sweep settings respectively. This improvement is not evident for the higher valued frequency sweep settings of 13.75, 27.5 and 55 MHz respectively.

Table 5.15: The cost of estimating Transmit Frequency with 1.7, 3.4 and 6.8 MHz Frequency Sweep modes versus Energy-to-Noise Ratio.

ENR [dB]	MSE	Percent MSE [%]	ABS Maximum [MHz]
	$\theta_{\Delta F} = 1.70$ MHz		
10	0.024	0.217	0.298
11	0.022	0.206	0.315
12	0.022	0.206	0.320
13	0.023	0.209	0.323
14	0.023	0.209	0.331
15	0.024	0.219	0.338
	$\theta_{\Delta F} = 3.40$ MHz		
10	0.077	0.655	0.564
11	0.075	0.638	0.568
12	0.077	0.659	0.584
13	0.076	0.652	0.588
14	0.079	0.673	0.596
15	0.079	0.675	0.606
	$\theta_{\Delta F} = 6.80$ MHz		
10	0.230	1.713	1.029
11	0.239	1.781	1.059
12	0.239	1.780	1.070
13	0.249	1.861	1.083
14	0.254	1.892	1.110
15	0.264	1.969	1.125

Table 5.16: The cost of estimating Transmit Frequency with 13.75, 27.5 and 55 MHz Frequency Sweep modes versus Energy-to-Noise Ratio.

ENR [dB]	MSE	Percent MSE [%]	ABS Maximum [MHz]
	$\theta_{\Delta F} = 13.75$ MHz		
10	0.909	5.388	1.974
11	0.919	5.444	1.989
12	0.915	5.422	2.001
13	0.909	5.386	2.017
14	0.919	5.444	2.031
15	0.937	5.550	2.038
	$\theta_{\Delta F} = 27.50$ MHz		
10	3.248	13.675	3.808
11	3.100	13.052	3.808
12	3.126	13.162	3.809
13	3.235	13.621	3.809
14	3.302	13.904	3.812
15	3.233	13.613	3.831
	$\theta_{\Delta F} = 55.00$ MHz		
10	12.161	32.428	7.402
11	12.366	32.975	7.402
12	12.259	32.692	7.402
13	12.211	32.564	7.402
14	12.692	33.845	7.402
15	12.378	33.007	7.405

Table 5.17: The cost of estimating the 1.7, 3.4 and 6.8 MHz Frequency Sweep modes versus Energy-to-Noise Ratio.

ENR [dB]	MSE	Percent MSE [%]	ABS Maximum [MHz]
	$\theta_{\Delta F} = 1.70$ MHz		
10	0.028	1.633	0.420
11	0.026	1.545	0.400
12	0.023	1.358	0.400
13	0.020	1.175	0.390
14	0.019	1.104	0.390
15	0.019	1.132	0.390
	$\theta_{\Delta F} = 3.40$ MHz		
10	0.042	1.224	0.510
11	0.040	1.174	0.510
12	0.034	1.011	0.570
13	0.030	0.870	0.550
14	0.027	0.804	0.540
15	0.026	0.770	0.590
	$\theta_{\Delta F} = 6.80$ MHz		
10	0.105	1.539	0.910
11	0.105	1.547	0.920
12	0.103	1.509	0.970
13	0.104	1.528	0.930
14	0.096	1.409	0.900
15	0.091	1.332	0.990

Table 5.18: The cost of estimating the 13.75, 27.5 and 55 MHz Frequency Sweep modes versus Energy-to-Noise Ratio.

ENR [dB]	MSE	Percent MSE [%]	ABS Maximum [MHz]
	$\theta_{\Delta F} = 13.75$ MHz		
10	0.548	3.989	2.130
11	0.567	4.123	2.130
12	0.564	4.102	2.000
13	0.552	4.018	2.060
14	0.550	4.002	2.110
15	0.535	3.889	2.090
	$\theta_{\Delta F} = 27.50$ MHz		
10	2.880	10.472	4.310
11	2.717	9.879	4.310
12	2.748	9.991	4.310
13	2.817	10.245	4.380
14	2.902	10.554	4.550
15	2.780	10.111	4.560
	$\theta_{\Delta F} = 55.00$ MHz		
10	13.679	24.871	8.460
11	14.109	25.653	8.460
12	14.014	25.480	8.460
13	13.872	25.221	8.790
14	14.423	26.223	8.990
15	14.059	25.562	8.990

Table 5.19: The cost of estimating Sweep Repetition Rate versus Energy-to-Noise Ratio.

ENR [dB]	MSE	Percent MSE [%]	ABS Maximum [Hz]
10	9.685	0.968	31.288
11	9.536	0.954	28.635
12	10.256	1.026	33.286
13	10.464	1.046	32.780
14	11.704	1.170	34.047
15	12.157	1.216	34.284

5.2.4 Sweep repetition rate estimation

The estimation cost of sweep repetition rate in the Pilot radar is given in Table 5.19. Here, there is a gradual decline in estimation accuracy with increasing ENR. Although sweep repetition rate estimates do not actually contribute an additional degree of freedom for classification of the Pilot Mk3 radar it may distinguish the Pilot radar from other radar types. Further analysis of parameter estimation in this subsection is presented graphically. Given that the goal is to classify one of twelve potential operating modes of the Pilot Mk3 radar from four input dimensions (or parameter estimates), two dimensional combinations of the parameters are presented. Table 5.20 shows a cross reference table for each possible two dimensional parameter combination.

On initial visual inspection the twelve radar modes can easily be separated. Figures 5.1–5.3 show a clear distinction of the antenna rotation rate versus any of the other parameters. Rotation rate versus transmit frequency and frequency sweep is tightly separated at low transmit and swept frequencies. This may be a problem for the Fuzzy ARTMAP classifier as it would be prone to errors for a given set of choice and vigilance parameters.

Table 5.20: Cross reference table for two dimensional parameter combinations.

	Transmit Frequency $\hat{\theta}_F$	Frequency Sweep $\hat{\theta}_{\Delta F}$	Sweep Repetition Frequency, $\hat{\theta}_S$
Rotation Rate, $\hat{\theta}_R$	Figure 5.1	Figure 5.2	Figure 5.3
Transmit Frequency, $\hat{\theta}_F$		Figure 5.4	Figure 5.5
Frequency Sweep, $\hat{\theta}_{\Delta F}$			Figure 5.6

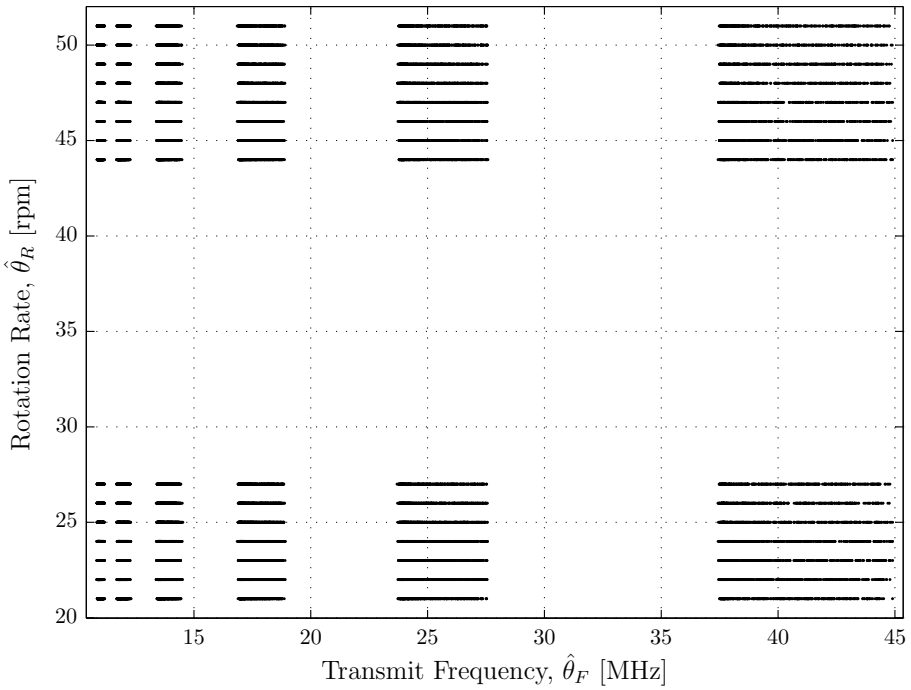


Figure 5.1: Distribution for Antenna Rotation Rate versus Transmit Frequency.

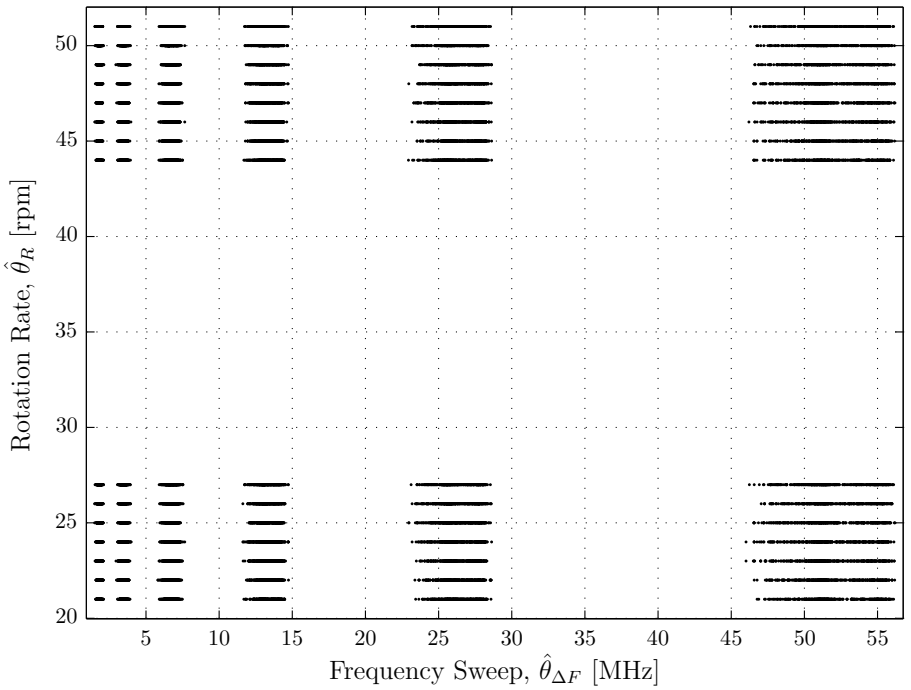


Figure 5.2: Distribution for Antenna Rotation Rate versus Frequency Sweep.

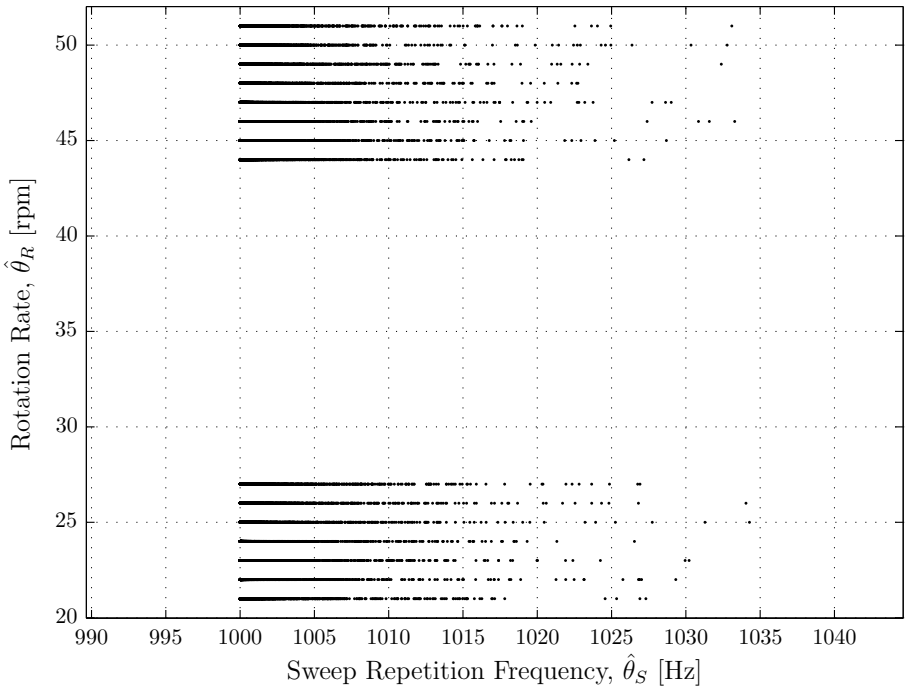


Figure 5.3: Distribution for Antenna Rotation Rate versus Sweep Repetition Frequency.

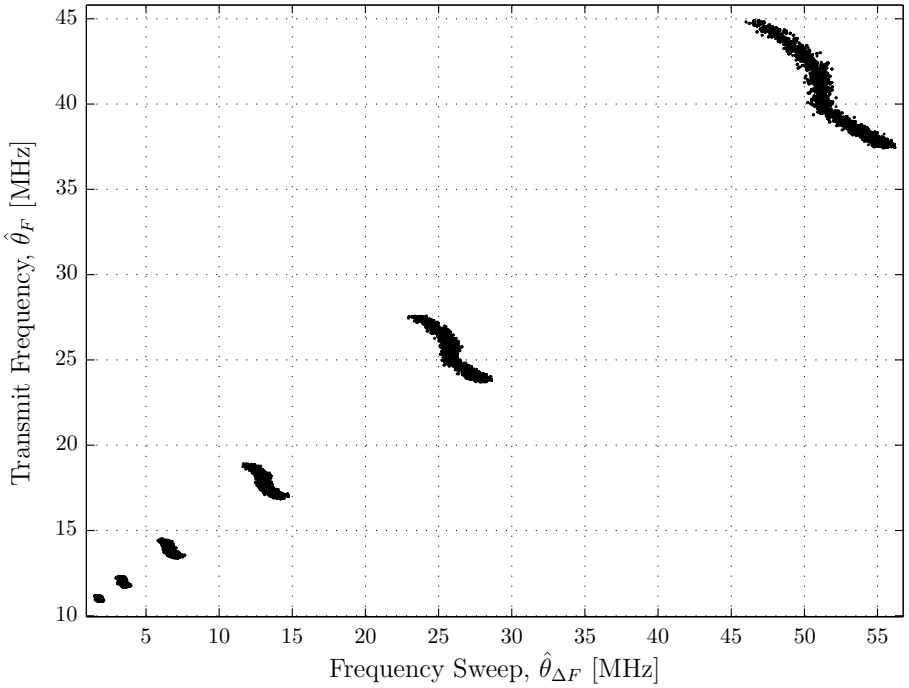


Figure 5.4: Distribution for Transmit Frequency versus Frequency Sweep.

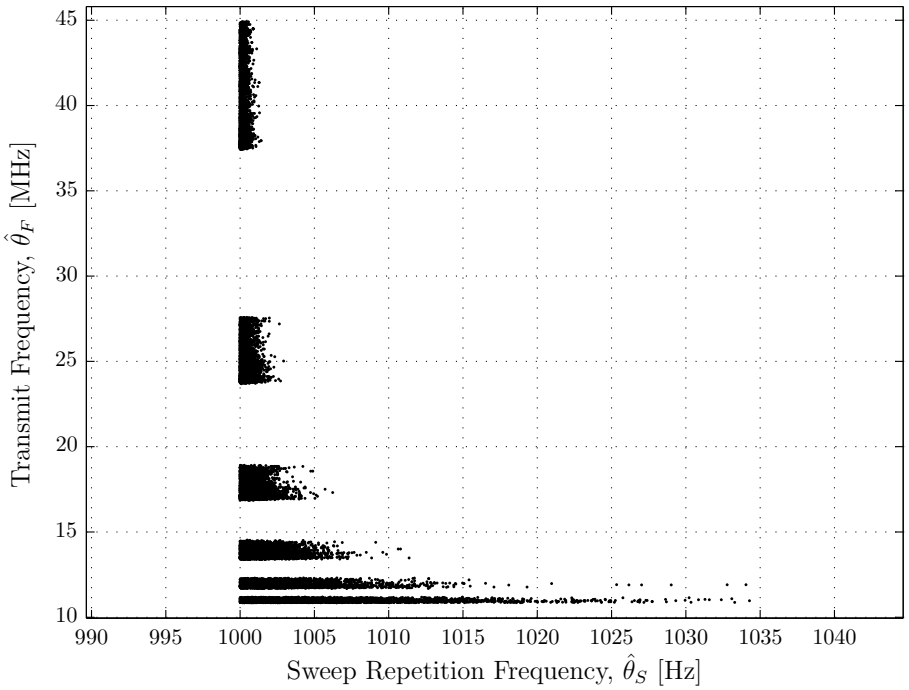


Figure 5.5: Distribution for Transmit Frequency versus Sweep Repetition Frequency.

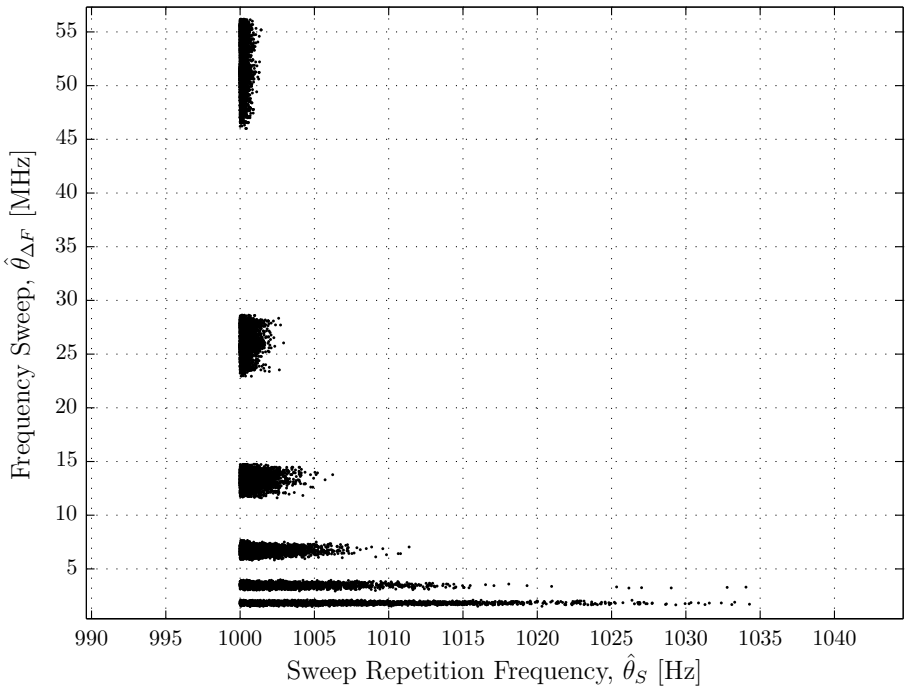


Figure 5.6: Distribution for Frequency Sweep versus Sweep Repetition Frequency.

Table 5.21: Estimation variance of Antenna Rotation Rate, θ_R for each radar mode versus ENR.

θ_R [rpm]	$\theta_{\Delta F}$ [MHz]	ENR [dB]						
		10	11	12	13	14	15	All
24	1.7	4.08	4.32	3.97	4.21	4.13	4.09	4.13
24	3.4	4.03	3.87	4.05	3.85	4.07	3.95	3.97
24	6.8	4.15	3.79	4.03	3.97	4.07	3.89	3.98
24	13.75	3.83	4.06	3.90	3.94	4.06	3.73	3.92
24	27.5	3.87	3.69	3.83	4.12	3.99	3.79	3.88
24	55	3.85	3.87	3.89	4.19	3.85	4.21	3.98
48	1.7	5.20	5.59	5.58	5.55	5.52	5.32	5.47
48	3.4	4.97	5.12	5.64	5.22	5.19	5.42	5.30
48	6.8	5.27	5.08	5.22	5.65	4.94	4.82	5.15
48	13.75	5.43	5.09	5.80	5.13	5.34	5.50	5.38
48	27.5	4.98	5.40	5.00	5.04	5.79	5.27	5.27
48	55	5.37	5.04	5.49	5.50	4.88	5.24	5.25

Tables 5.21 – 5.24 presents the estimation variance of each parameter, for each assigned radar mode (the combination of antenna rotation rate and frequency sweep) versus detector ENR. Estimation variance at each ENR value is compared with the total estimation variance in each mode and is denoted by, *All*. Overall, the estimation variance is quite low and does not vary dramatically with ENR. The only exception is a relatively large sweep repetition frequency estimation variance in the modes that contain frequency sweeps of 1.7 and 3.4 MHz irrespective of antenna rotation rate (Table 5.24).

Table 5.22: Estimation variance of Transmit Frequency, θ_F for each radar mode versus ENR.

θ_R [rpm]	$\theta_{\Delta F}$ [MHz]	ENR [dB]						
		10	11	12	13	14	15	All
24	1.7	0.01	0.01	0.01	0.01	0.01	0.01	0.01
24	3.4	0.02	0.02	0.02	0.02	0.02	0.03	0.02
24	6.8	0.06	0.07	0.08	0.08	0.09	0.09	0.08
24	13.75	0.33	0.32	0.33	0.32	0.32	0.34	0.32
24	27.5	1.08	1.23	1.17	1.22	1.27	1.21	1.21
24	55	4.95	4.74	4.41	4.57	4.68	4.68	4.66
48	1.7	0.01	0.01	0.01	0.01	0.01	0.01	0.01
48	3.4	0.02	0.02	0.02	0.02	0.02	0.02	0.02
48	6.8	0.08	0.08	0.08	0.09	0.08	0.09	0.08
48	13.75	0.29	0.32	0.32	0.33	0.34	0.33	0.32
48	27.5	1.31	1.18	1.27	1.25	1.17	1.19	1.22
48	55	4.79	4.93	5.05	4.78	4.85	4.54	4.82

Table 5.23: Estimation variance of Frequency Sweep, $\theta_{\Delta F}$ for each radar mode versus ENR.

θ_R [rpm]	$\theta_{\Delta F}$ [MHz]	ENR [dB]						
		10	11	12	13	14	15	All
24	1.7	0.01	0.01	0.01	0.01	0.01	0.01	0.01
24	3.4	0.03	0.03	0.03	0.03	0.03	0.02	0.03
24	6.8	0.10	0.10	0.10	0.09	0.09	0.08	0.09
24	13.75	0.39	0.41	0.37	0.35	0.31	0.33	0.36
24	27.5	1.32	1.40	1.37	1.34	1.36	1.22	1.34
24	55	5.58	5.14	5.10	5.09	5.05	5.14	5.16
48	1.7	0.01	0.01	0.01	0.01	0.01	0.01	0.01
48	3.4	0.03	0.03	0.03	0.03	0.02	0.02	0.03
48	6.8	0.11	0.10	0.09	0.10	0.08	0.08	0.09
48	13.75	0.39	0.39	0.41	0.38	0.38	0.32	0.37
48	27.5	1.56	1.43	1.47	1.43	1.29	1.29	1.40
48	55	5.34	5.74	5.60	5.30	5.46	4.93	5.39

Table 5.24: Estimation variance of Sweep Repetition Frequency, θ_S for each radar mode versus ENR.

θ_R [rpm]	$\theta_{\Delta F}$ [MHz]	ENR [dB]						
		10	11	12	13	14	15	All
24	1.7	33.70	26.11	22.94	23.23	24.73	26.98	25.67
24	3.4	6.57	8.07	6.83	7.00	10.30	8.33	8.03
24	6.8	1.67	1.42	2.05	1.65	1.77	1.81	1.74
24	13.75	0.43	0.51	0.48	0.55	0.62	0.46	0.51
24	27.5	0.12	0.13	0.13	0.12	0.12	0.09	0.12
24	55	0.03	0.03	0.04	0.03	0.03	0.03	0.03
48	1.7	22.36	27.49	28.73	23.59	27.68	25.81	26.08
48	3.4	6.04	8.26	8.68	10.31	5.82	6.46	7.57
48	6.8	1.65	1.81	1.61	1.69	1.91	1.96	1.78
48	13.75	0.52	0.49	0.51	0.51	0.69	0.42	0.52
48	27.5	0.11	0.17	0.10	0.11	0.12	0.13	0.12
48	55	0.04	0.03	0.03	0.03	0.03	0.03	0.03

5.2.5 Discussion

In this section results from the parameter estimation stage of the conceptual intercept receiver is presented. For each of the four Pilot Mk3 radar parameters estimated there is very clear clustering of parameter data that distinguish each of the twelve assigned radar modes to be classified. At low detector ENR the cost of estimation is good and estimation variance is quite low, which suggest that if the detector is able to detect an intercepted signal it will be able to make accurate estimates of radar parameters. The parameter estimation stage implemented in this study can easily enable the classification of Pilot Mk3 radar modes.

5.3 RADAR MODE CLASSIFICATION

In this section the classification of the twelve assigned Pilot Mk3 radar modes is presented. Results are analysed with attention to the following;

1. Classifier input and output normalisation. Fuzzy ARTMAP classifiers have an inherent ability to simultaneously consider features that do and don't belong to a given class. This is done by combining data with their respective complement coded versions.

2. Optimal classifier parameter sets. The selection of classifier choice (α), learning rate (β) and vigilance (ρ) parameter values influence classification behaviour during training and evaluation.

3. The ratio of training data to evaluation data required for accurate classification. With too little training data the classifier might not be able to generalise data well enough to be accurate during evaluation.

4. Highest attainable classification accuracy. This is the ultimate goal of the classifier, to be as accurate as possible on a given dataset, for a set of pre-defined parameters.

5.3.1 Data normalisation

The use of complement coding in Fuzzy ARTMAP classification implies that there be some form of data normalisation. With reference to Section 3.1.1 both input and output data are normalised as

$$I = \frac{\mathbf{a}}{|\mathbf{a}|} \,. \tag{5.2}$$

Here, $|\cdot|$ is the ℓ_1-norm of the data vector $\mathbf{a}$. The complement coded form of data vectors is $\mathbf{a}^c = 1 - \mathbf{a}$. There is however some more data conditioning to be done in order to preserve each data distribution as well as the separation of the assigned radar classes. The input data to the classifier is scaled further with the following factor

$$\mathbf{a} = \frac{\theta}{\zeta} \tag{5.3}$$

with

$$\zeta = [75, 110, 110, 1500] \text{ in units of [rpm, MHz, MHz, Hz]} \,. \tag{5.4}$$

Output data ($\mathbf{b}$) is also scaled but the data values are chosen by the user. All data in the dataset are *labelled*, which means that the classifier is only required to produce output values and not define output them. Given the twelve classes, output values that represent each class are defined as follows;

Table 5.25: Output data specification and class assignment.

Mode	Output **b**	θ_R [rpm]	$\theta_{\Delta F}$ [MHz]
1	0.08	24	1.7
2	0.16	24	3.4
3	0.24	24	6.8
4	0.32	24	13.75
5	0.40	24	27.5
6	0.48	24	55
7	0.56	48	1.7
8	0.64	48	3.4
9	0.72	48	6.8
10	0.80	48	13.75
11	0.88	48	27.5
12	0.96	48	55

5.3.2 Classifier parameter selection

As was shown in Section 3.1.3 (Figure 3.4), Fuzzy ARTMAP parameter values do influence the classifier's behaviour, which in turn impacts classification accuracy. The user may attempt an initial guess of the category choice (α), learning rate (β) and vigilance (ρ), but optimal parameter values are found as a result of some kind search. This search process involves training on a set of parameter values followed by evaluating classifier accuracy. The process is repeated with different parameter values until the search is depleted or until a certain criteria is reached. The following paragraphs present analysis of each Fuzzy ARTMAP parameter on the classification of the Pilot Mk3 radar. Extensive graphical results are presented in Addendum A.

Category choice, α

The category choice parameter quantifies the degree of fuzzy subset-hood of data to existing categories in the ART^a and ART^b modules. Small category choice values allow data to belong to a category with increased fuzziness. The category choice function of Equation (3.6) merely selects the best fuzzy set (or category) for the data. From the distribution of input data from Figures 5.1 to 5.6 a large category choice value would help the classifier select ART^a categories with low fuzzy subset-hood, especially when faced with transmit frequency ($\hat{\theta}_F$) and frequency sweep ($\hat{\theta}_{\Delta F}$) data.

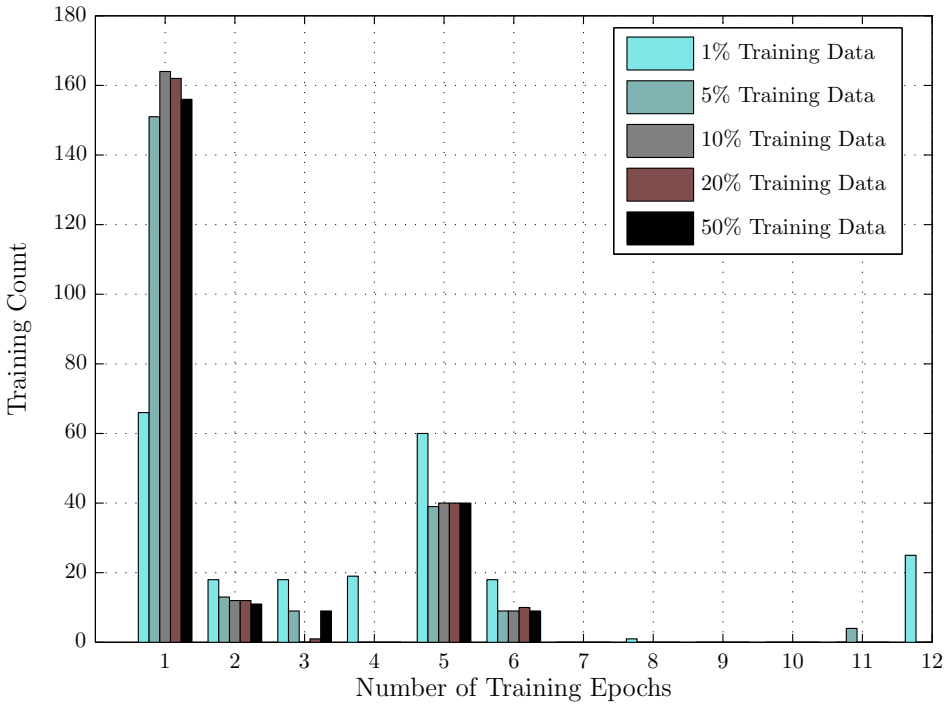


Figure 5.7: Distribution of training epochs for various training dataset sizes.

From Figures A.1 to A.5 in Addendum A, classification accuracy is not dictated by category choice parameter values of 10 and lower. Values greater than 10 up to 100 tend to ensure 100% classification accuracy irrespective of the value of learning rate and vigilance. With these high category choice values the category selection in each ART module is rather strict with data belonging exclusively to a single category. The advantage is that a small amount of categories are formed during training as opposed to having the *category proliferation* problem that forms a new category for each data vector presented.

Learning rate, β

The learning rate parameter only influence the Fuzzy ARTMAP classifier during training as it controls the convergence of weights in both *ART* modules as well as the map field. Learning rate values used during the analysis where $\beta \in [0.3, 0.5, 0.7, 0.9, 1]$. Classification accuracy does not depend on any of the learning rate values used in the analysis. This independence also prevails across different training data set sizes and the number of training epochs (or training repetitions) required for map field weight convergence. Figure 5.7 shows the distribution of training epochs for various training dataset sizes. Very rarely does it take more than 6 epochs for map field weights to converge with the overwhelming

majority of training requiring only a single presentation of training data.

Vigilance, ρ

On visual inspection of the input data distribution for each assigned class, a low the vigilance parameter value will result in poor classification as it allows the formation of larger sized categories. Larger categories will accommodate radar frequency sweeps of 13.75, 27.5 and 55 MHz who are separated well, but not for frequency sweeps of 1.7, 3.4 and 6.8 MHz who are not separated too far apart. Thus, the search for the optimal vigilance parameter was conducted for $\rho \in [0.8, 0.83, 0.85, 0.87, 0.89, 0.9, 0.91, 0.92, 0.93]$. Classification accuracy increases with increasing vigilance values and 100% accuracy achieved at $\rho = 0.93$.

5.3.3 Amount of training data

Another important point to consider is the ratio of training dataset size to evaluation dataset size. Here, the proper dataset size is specified to ensure good generalisation of the classifier during training. When any classifier has a good generalisation of its data space, it is more robust and has a more complete reference of the bounds of valid inputs and outputs. In the analysis of training dataset size, classifier accuracy is evaluated for training dataset sizes of 1%, 5%, 10%, 20% and 50% respectively. Training samples are drawn randomly and represent the percentage of input data samples associated with each class. For example, if there are 234 and 189 input samples for class number 1 and 2 respectively, a 10% training dataset will comprise of 23 and 18 data samples respectively. Training data are not reused during evaluation as it does not contribute to any understanding of classifier behaviour. Interestingly, the 1% training dataset also achieved 100% classification accuracy for certain parameter values. However, the 1% training dataset did have problems with category proliferation and some of the training attempts never resulted in map field weight convergence. From Figures A.1 to A.5 in Addendum A there is very little improvement in classification accuracy when the classifier trains on more data. Thus, the classifier will develop a complete reference of the data space with at least 5% of the data used for training and be able to achieve 100% classification accuracy.

5.3.4 Overall classification accuracy

Analysis of the accuracy of the Fuzzy ARTMAP classifier is imperative to answering the research questions of this study. The classifier is trained to predict the output values defined in Table 5.25. These values are assigned (or linked) to class labels that describe each mode of the Pilot Mk3 radar. The accuracy of the classifier's output predictions and the selection of the correct class labels are of

importance here. Output predication accuracy is defined by a variable termed Output Value Accuracy (OVA) as an average accuracy value of an evaluation set of size P. OVA is defined as

$$\text{OVA} = \frac{1}{P}\sum^{p}\left|\frac{\tilde{\mathbf{b}}^{(p)}}{\mathbf{b}^{(p)}}\right| . \tag{5.5}$$

OVA is calculated as the total number of output predictions, $\tilde{\mathbf{b}}$ that are less than 10%, 20% and 50% of the true output values, $\mathbf{b}$, as an absolute fraction averaged to the evaluation dataset size as

$$\text{OVA}_{10} \equiv 0.9 < \text{OVA} < 1.1 \tag{5.6}$$

$$\text{OVA}_{20} \equiv 0.8 < \text{OVA} < 1.2 \tag{5.7}$$

$$\text{OVA}_{50} \equiv 0.5 < \text{OVA} < 1.5 . \tag{5.8}$$

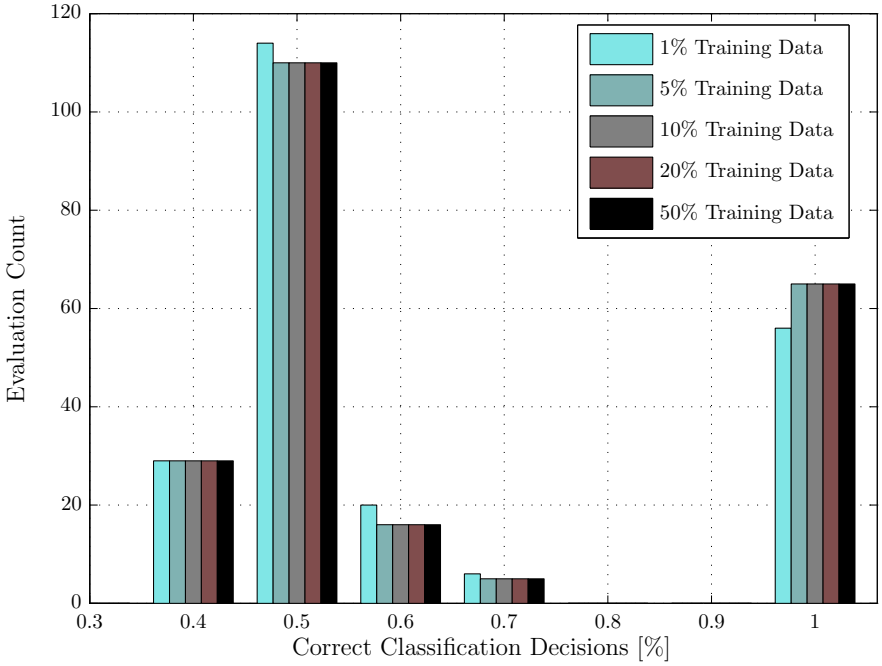
Here, the subscript 10, 20 and 50 represent predicted outputs within 10%, 20% and 50% of the true outputs respectively. Classification decisions are made by choosing the nearest true output value from the set of assigned classes as defined in Table 5.25. Given that the assigned output values are equally spaced by $r = 0.08$, any predicted output from the classifier greater or less than $r/2$ from the true output will result in an incorrect classification decision. Correct Classification Decisions (CCD) are defined as the ratio of the total correct classification decisions to the total size of the evaluation set P as follows;

$$\text{CCD} = \frac{1}{P}\sum^{p}\left|\tilde{\mathbf{b}}^{(p)} - \mathbf{b}^{(p)}\right|, \text{ for } \left|\tilde{\mathbf{b}}^{(p)} - \mathbf{b}^{(p)}\right| < \frac{r}{2} \tag{5.9}$$

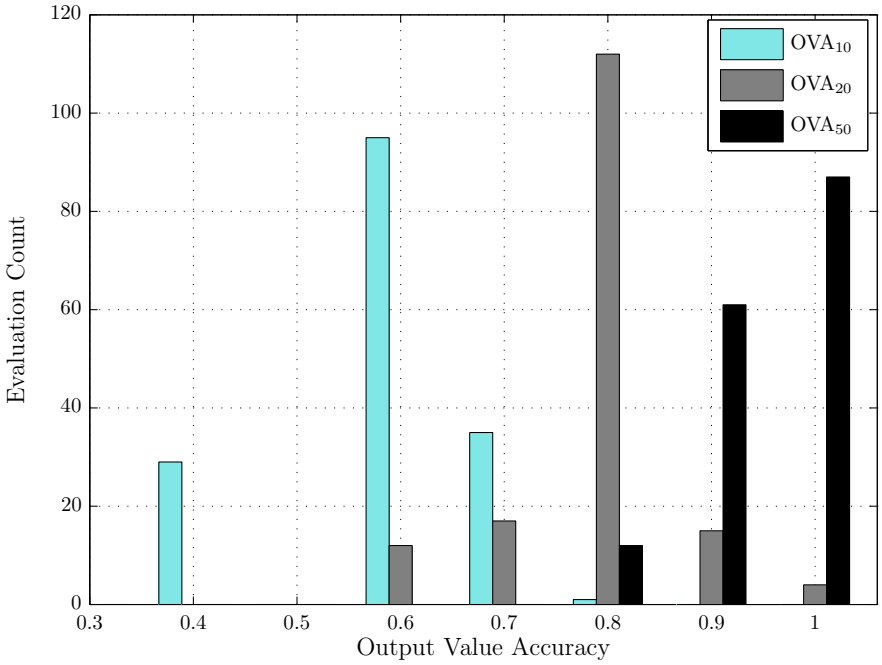
Since there are many classifier parameter combinations that achieve 100% CCD it is worth commenting on the distribution of CCD and OVA as performance indicators. Figure 5.8 show these distributions in histogram form for CCD (Figure 5.8 (a)) and OVA (Figure 5.8 (b)) respectively. In Figure 5.8 (a) 100% CCD is feasible even with 1% training dataset size and that classifier performance is constant with increasing training dataset size. The distribution of OVA when CCD is less than 100% from the 50% training dataset. Here, OVA_{50} may still be achieved during evaluation. This is due to the range of possible classifier output values. For example, 50% relative to the minimum output value of 0.08 is exactly 0.04. This is not enough to result in an erroneous classification, which only occurs at output values greater than 0.12 or less than 0.04. On the contrary 50% of the maximum output value of 0.96 is 0.48, which is enough to cause errors beyond adjacent classes. OVA is thus not related to CCD in any way, as it merely shows the user that the classifier is capable of a certain accuracy when predicting output values during evaluation.

5.3.5 Discussion

The Fuzzy ARTMAP classifier is capable of classifying the radar modes of the Pilot Mk3 LPI radar. Correct classification decisions of 100% is easily achieved for a variety of classifier parameters. Fortunately there is no global optimum classifier parameter combination, which would require a far more exhaustive search. Classifier training is quite efficient as good generalisation between input and output spaces is achieved from a training dataset comprising only 5% of the total dataset. Figure 5.9 shows the map field weights after training on the 5% training dataset. The 5% training dataset contains 1470 samples and 16 categories are formed in the ART^a module, which result in a compression ratio of 16:1470 (or 8:735). This compression ratio is quite significant as the classifier will form only 8 ART^a categories per 735 input samples during training, which confirms the classifier's ability to properly generalise its reference toward the greater evaluation dataset during training. In other words the classifier is able to describe the input data space by having only 16 memory elements (ART^a weights) instead of remembering everyone of the 1470 individual input samples. Compression ratios for each training dataset size are shown in Table 5.26. It must also be mentioned that the compression ratio on the 1% training dataset is the greater of all listed in Table 5.26, but map field convergence during is not guaranteed for the 1% dataset.



(a)



(b)

Figure 5.8: Distribution of Fuzzy ARTMAP classifier performance indicators. (a) A histogram of CCD for various training dataset sizes (bin size is 0.1), and (b) shows a histogram of OVA when CCD is less than 100% (bin size is 0.1) from the 50% training dataset.

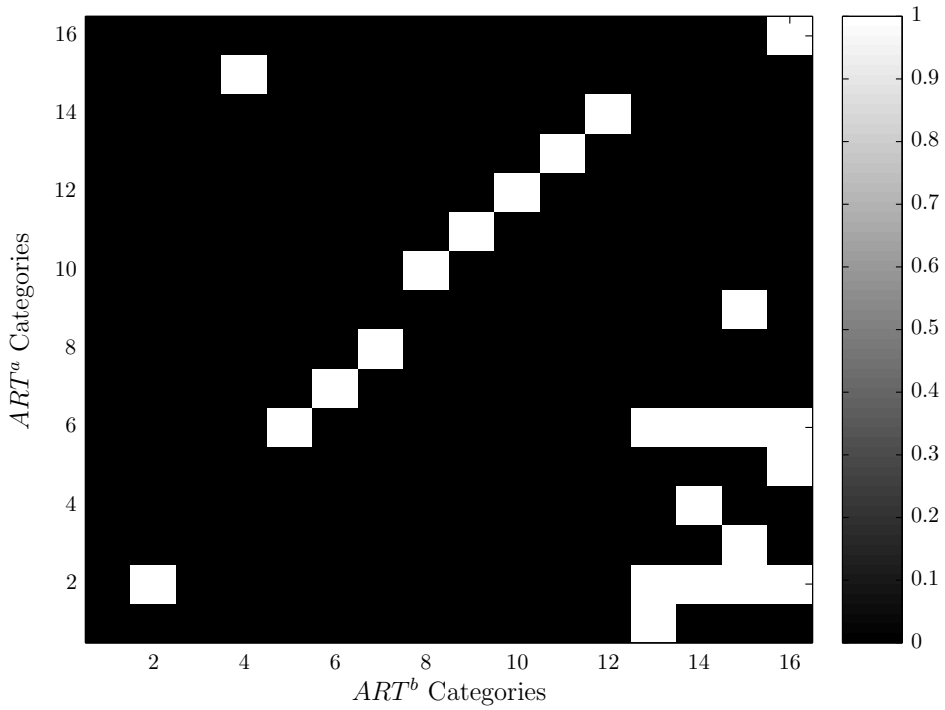


Figure 5.9: Map field weights after training on the 5% dataset with $\alpha = 10$, $\beta = 1$ and $\rho = 0.93$.

Table 5.26: Classifier compression ratio when training on various dataset sizes.

Training Dataset Size [%]	ART^a Categories	Compression Ratio
1	17	17:291
5	16	8:735
10	16	16:2945
20	16	16:5893
50	16	16:14739

CHAPTER 6

CONCLUSIONS

The classification of Low Probability of Intercept (LPI) radar was studied here. The study was conducted from an Electronic Warfare (EW) point of view. With reference to the functional flow of a typical intercept receiver (Figure 1.1), such a receiver aims to *extract* intelligence from intercepted signals. For the purposes of conducting a thorough study a radar intercept scenario is defined, with simulated radar emissions, the detection these emissions (or signals), extracted radar features (estimate signal parameters) and the classification of radar operating modes from the resulting features, was considered.

Classification accuracy of 100% can readily be achieved without being constrained to a unique set of classifier parameters. The features or estimated radar signal parameters are well separated to the point that there is no overlap of features that may cause classification errors. The accuracy of parameter estimation is also very useful to the point that if the detector makes a detection, the parameter estimates would be accurate enough to ensure accurate classification. Four features were selected for classification of the Pilot Mk3 LPI radar. They are;

1. Antenna rotation rate, θ_R,
2. Transmit frequency, θ_F,
3. Frequency sweep, $\theta_{\Delta F}$, and
4. Sweep repetition frequency, θ_S.

This four dimensional input space is mapped to a one dimensional output space in order to classify one of twelve radar operating modes. These operating modes relate directly to the combinations of antenna rotation rate and frequency sweep.

The concept of LPI radar might be somewhat misleading to the point were one might be lead to

believe that intercepting LPI radar is a matter of impossibility. If any radar (or emitter) is LPI, there must be a consideration for the radar as well as the opposing intercept receiver. The claim that a particular radar is LPI against any intercept receiver is too broad to be insightful. Similarly, it also holds for an intercept receiver claiming to have 100% Probability of Intercept (POI) against any radar.

In the following sections conclusions are drawn on the detection, feature extraction (parameter estimation) and classification processing stages in the intercept receiver conceptualised for this study.

6.1 DETECTOR

Given a-priori knowledge of the Pilot Mk3 LPI radar, a Maximum Likelihood (ML) (or Neyman-Pearson) energy detector was developed. The detector was evaluated for various Energy-to-Noise Ratios (ENRs) to account for realistic operating environments where intercepted signal levels are all but constant. The LPI performance factor for various amounts of detector integration gain was also evaluated and the stand-off range for an intercept receiver, using an equal-length matched filter (with integration gain greater than 50 dB), is comfortably beyond the maximum radar range of 24 nmi. This also applies to a sidelobe scenario which result in the detector having to detect radar signals with 30 dB less input power.

6.2 PARAMETER ESTIMATION

Once detections are made, radar signal parameters are estimated quite reliably and accurately. The reason for this is that detector outputs (correlation between noise intercepted samples and the chosen matched filter) are processed backwards in time. N samples (the matched filter length) ahead of the detected sample is recalled from memory and all N samples are used for parameter estimation processing. This results in a feature vector $\hat{\theta} = [\hat{\theta}_R, \hat{\theta}_F, \hat{\theta}_{\Delta F}, \hat{\theta}_S]$. There is also no significant difference in parameter estimation accuracy at low detector ENR values.

6.3 CLASSIFICATION

Due to the accuracy and separability of estimated radar parameters, a Fuzzy ARTMAP classifier has very little difficulty to achieve 100% classification decisions, even when training on a 5% data subset. By choosing category choice values that are strict on fuzzy subset-hood and vigilance values that form relatively few *ART* module categories, the classifier is sure to be accurate. A compression ratio of 16:1470 (or 8:735) is achieved on the 5% training dataset. This compression ratio is quite significant as the classifier will form only 8 ART^a categories per 735 input samples during training. The Fuzzy ARTMAP classifier is capable of fully describing Pilot Mk3 radar modes by having only 16 memory

elements (ART^a weights) on the 5% training dataset with vigilance $\rho = 0.93$ irrespective of category choice and learning rate parameter values.

REFERENCES

[1] D. C. Schleher, *Introduction to Electronic Warfare*. Dedham, MA: Artech House, 1986.

[2] ——, *Electronic Warfare in the Information Age*. Norwood, MA: Artech House, 1999.

[3] J. B. Tsui, *Microwave Receivers with Electronic Warfare Applications*. New York: John Wiley & Sons, 1986.

[4] D. Adamy, *EW 101: A First Course in Electronic Warfare*. New York: Artech House, 2000.

[5] E. Granger, M. A. Rubin, S. Grossberg, and P. Lavoie, "A What-and-Where fusion neural network for recognition and tracking of multiple radar emitters," *Neural Networks*, vol. 14, pp. 325–344, 2001.

[6] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*. Norwood MA: Artech House, 2004.

[7] P. Rattey and A. Lindgren, "Sampling the 2-D Radon transform," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 5, pp. 994–1002, Oct. 1981.

[8] J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition*. Piscataway, NJ: IEEE Press, 1992.

[9] L. Zadeh, "Fuzzy sets," *Information Control*, vol. 8, no. 1, pp. 338–353, 1965.

[10] P. K. Simpson, "Fuzzy min-max neural networks – part 2: Clustering," *Fuzzy Systems, IEEE Transactions on*, vol. 1, no. 1, p. 32, Feb. 1993.

[11] ——, "Fuzzy min-max neural networks – part 1: Classification," *Neural Networks, IEEE Transactions on*, vol. 3, no. 5, pp. 776–786, Sep. 1992.

[12] P. K. Simpson and G. Jahns, "Fuzzy min-max neural networks for function approximation," in *Neural Networks, 1993., IEEE International Conference on*, vol. 3, 1993, pp. 1967–1972.

[13] C. M. Bishop, *Neural Networks for Pattern Recognition*. London: Oxford University Press, 1995.

[14] P. L. Ainsleigh, N. Kehtarnavaz, and R. L. Streit, "Hidden-gauss markov models for signal classification," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1355–1367, 2002.

[15] P. L. Ainsleigh, S. G. Greineder, and N. Kehtarnavaz, "Classification of nonstationary narrowband signals using segmented chirp features and hidden-gauss markov models," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 147–157, 2005.

[16] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[17] G. A. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organising neural pattern recognition machine," *Computer Vision, Graphics and Image Processing*, no. 37, pp. 54–115, 1987.

[18] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 698–713, September 1992.

[19] G. A. Carpenter, S. Grossberg, and J. H. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Networks*, vol. 4, pp. 565–588, 1991.

[20] S. Grossberg, "Adaptive pattern classification and universal recoding part 1: Parallel development and coding of neural feature detectors," *Biological Cybernetics*, vol. 23, no. 1, pp. 121–134, 1976.

[21] ——, "Adaptive pattern classification and universal recoding part 2: Feedback,expectation, olfaction and illusions," *Biological Cybernetics*, vol. 23, no. 1, pp. 187–202, 1976.

[22] G. Dogic-Crnkovic, *Theory of Science*. Sweden: Malardalens Hogskola, 2003.

[23] Z. Wang and P. Willett, "A performance study of some transient detectors," *IEEE Trans. Signal Processing*, vol. 48, no. 9, pp. 2682–2685, 1994.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, ser. Prentice Hall signal processing series, A. V. Oppenheim, Ed. Prentice-Hall PTR, 1998, vol. II.

[25] P. F. Potgieter and J. C. Olivier, "Inferring radar mode changes from elementary pulse features using fuzzy artmap classification," in *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, vol. 1, November 2007, pp. 88–93.

[26] X. Gong, H. Meng, and X. Wang, "A GMM-based Algorithm for Classification of Radar emitters," in *Signal Processing, 2008. ICSP 2008. 9th International Conference on*. IEEE, 2008, pp. 2434–2437.

[27] M. Conning and P. F. Potgieter, "Analysis of measured radar data for specific emitter identification," in *Radar Conference, 2010 IEEE*, May 2010, pp. 35–38.

[28] G. W. Fisher, "Robust Geolocation Techniques for Multiple Receiver Systems," Masters Dissertation, Baylor University, December 2011.

[29] J. Kosinski, W. Palmer, and M. Steer, "Unified understanding of RF remote probing," *Sensors Journal, IEEE*, vol. 11, no. 12, pp. 3055–3063, December 2011.

[30] L. B. Van Brunt, *Applied ECM*. EW Engineering Inc., 1995, vol. 3.

[31] M. A. Richards, *Fundamentals of Radar Signal Processing*, ser. McGraw-Hill Electronic Engineering Series. McGraw-Hill, 2005, ch. 1, pp. 1–3.

[32] R. G. Wiley, *Electronic Intelligence: The Interception of Radar Signals*. Artech House Inc., 1985.

[33] M. Burgos-Garcia, J. Sanmartin-Jara, F. Perez-Martinez, and J. Retamosa, "Radar sensor using low probability of interception ss-fh signals," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 15, no. 4, pp. 23 –28, Apr. 2000.

[34] J. Costas, "A study of a class of detection waveforms having nearly ideal range–Doppler ambiguity properties," *Proceedings of the IEEE*, vol. 72, no. 8, pp. 996–1009, August 1984.

[35] R. Frank, "Polyphase codes with good nonperiodic correlation properties," *Information Theory, IEEE Transactions on*, vol. 9, no. 1, pp. 43 – 45, Jan. 1963.

[36] ——, "Polyphase complementary codes," *Information Theory, IEEE Transactions on*, vol. 26, no. 6, pp. 641–647, Nov. 1980.

[37] G. Painchaud, J. McKenzie, M. Blanchette, and A. Voy, "An experimental adaptive digital pulse compression subsystem for multi-function radar applications," in *Radar Conference, 1990., Record of the IEEE 1990 International*, May 1990, pp. 153 –158.

[38] B. Lewis, "Range-time-sidelobe reduction technique for fm-derived polyphase pc codes," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 29, no. 3, pp. 834 –840, Jul. 1993.

[39] J. Fielding, "Polytime coding as a means of pulse compression," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 35, no. 2, pp. 716 –721, Apr. 1999.

[40] L. V. Blake, *Radar Handbook*, 2nd ed. McGraw-Hill, 1990, ch. 2, pp. 2.1–2.10.

[41] D. C. Schleher, "LPI radar: fact or fiction," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 21, no. 5, pp. 3 – 6, May 2006.

[42] Naval Air Warfare Center, Weapons Division, *EW and Radar Systems Engineering Handbook*, 1999, Revision 2.

[43] G. A. Carpenter and S. Grossberg, "A self-organising neural network for supervised learning, recognition and prediction," *IEEE Communications Magazine*, pp. 38–49, September 1992.

[44] G. A. Carpenter, S. Grossberg, and J. H. Reynolds, "A Fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems," *IEEE Transactions on Neural Networks*, vol. 6, no. 6, pp. 1330–1336, November 1995.

[45] S. Marroitt and R. F. Harrison, "A modified Fuzzy ARTMAP architecture for the approximation of noisy mappings," *Neural Networks*, vol. 8, no. 4, pp. 619–641, 1995.

[46] S. Verzi, G. Heileman, M. Georgiopoulos, and G. Anagnostopoulos, "Universal approximation with Fuzzy ART and Fuzzy ARTMAP," in *Neural Networks, 2003. Proceedings of the International Joint Conference on*, vol. 3, July 2003, pp. 1987–1992.

[47] J. S. Lee, C. G. Yoon, and C. W. Lee, "Learning method for fuzzy ARTMAP in a noisy environment," *Electronics Letters*, vol. 34, no. 1, pp. 95–97, January 1998.

[48] D. Chralampidis, T. Kasparis, and M. Georgiopoulos, "Classification of noisy signals using fuzzy artmap neural networks," *Neural Networks, IEEE Transactions on*, vol. 12, no. 5, pp. 1023–1036, September 2001.

[49] G. A. Carpenter and S. Grossberg, "The ART of adaptive pattern recognition self-organizing by a neural network," *IEEE Computer Magazine*, pp. 77–88, March 1988.

[50] K. Fuller, "To see and not be seen," in *IEE Proceedings Radar, Sonar and Navigation Signal Processing*, vol. 137, no. 1, February 1990, pp. 1–9.

[51] SAAB Group. PILOT LPI Radar. SE-412 89 Gothenburg, Sweden. Product Technical Data and Specifications. [Online]. Available: http://www.saabgroup.com/en/Naval/Situational-Awareness/Navigation-and-Detection/PILOT_LPI_Radar_Low_Probability_of_Intercept/

[52] J. G. Proakis and M. Salehi, *Communcation Systems Engineering*. Prentice–Hall Int., 1994.

[53] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, 1st ed. John Wiley and Sons, Inc., September 1968.

[54] Independent Communications Authority of South Africa, "National frequency plan," *South African Government Gazette*, July 2010.

[55] L. B. Wetzel, *Radar Handbook*, 2nd ed. McGraw-Hill, 1990, ch. 13, pp. 13.1–13.2.

[56] H. E. Schrank, G. E. Evans, and D. Davis, *Radar Handbook*, 2nd ed. McGraw-Hill, 1990, ch. 6, pp. 6.2–6.10.

[57] G. E. Carlson, *Signal and Linear System Analysis*, 2nd ed. John Wiley, 1998, ch. 5, pp. 178–180.

[58] I. Jae Myung, "Tutorial on maximum likelihood estimation," *Journal of Mathematical Psychology*, vol. 47, no. 1, pp. 90–100, 2003.

[59] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego: Academic Press, 1990.

[60] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Human Genetics*, vol. 7, no. 2, pp. 179–188, 1936.

[61] A. Stefánsson, N. Končar, and A. J. Jones, "A note on the gamma test," *Neural Computing & Applications, Springer-Verlag London Limited*, vol. 5, no. 3, pp. 131–133, 1997.

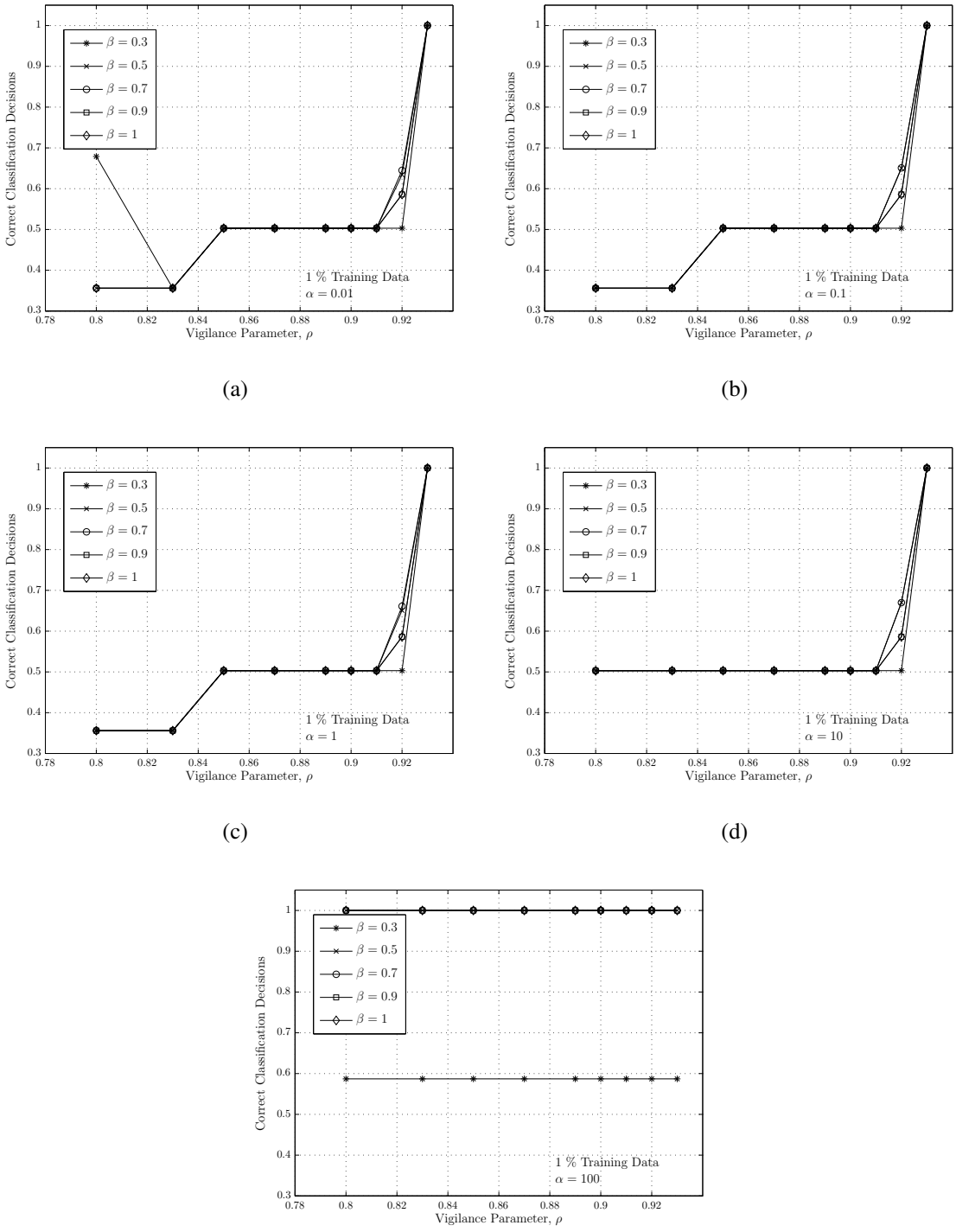
[62] S. E. Kemp, I. D. Wilson, and J. A. Ware, "A tutorial on the gamma test," *International Journal of Simulation*, vol. 6, no. 1–2, pp. 67–75, January 2005.

[63] D. Evans and A. J. Jones, "A proof of the gamma test," *Article to the Royal Society, Department of Computer Science, Cardiff University*, 2002.

ADDENDUM A

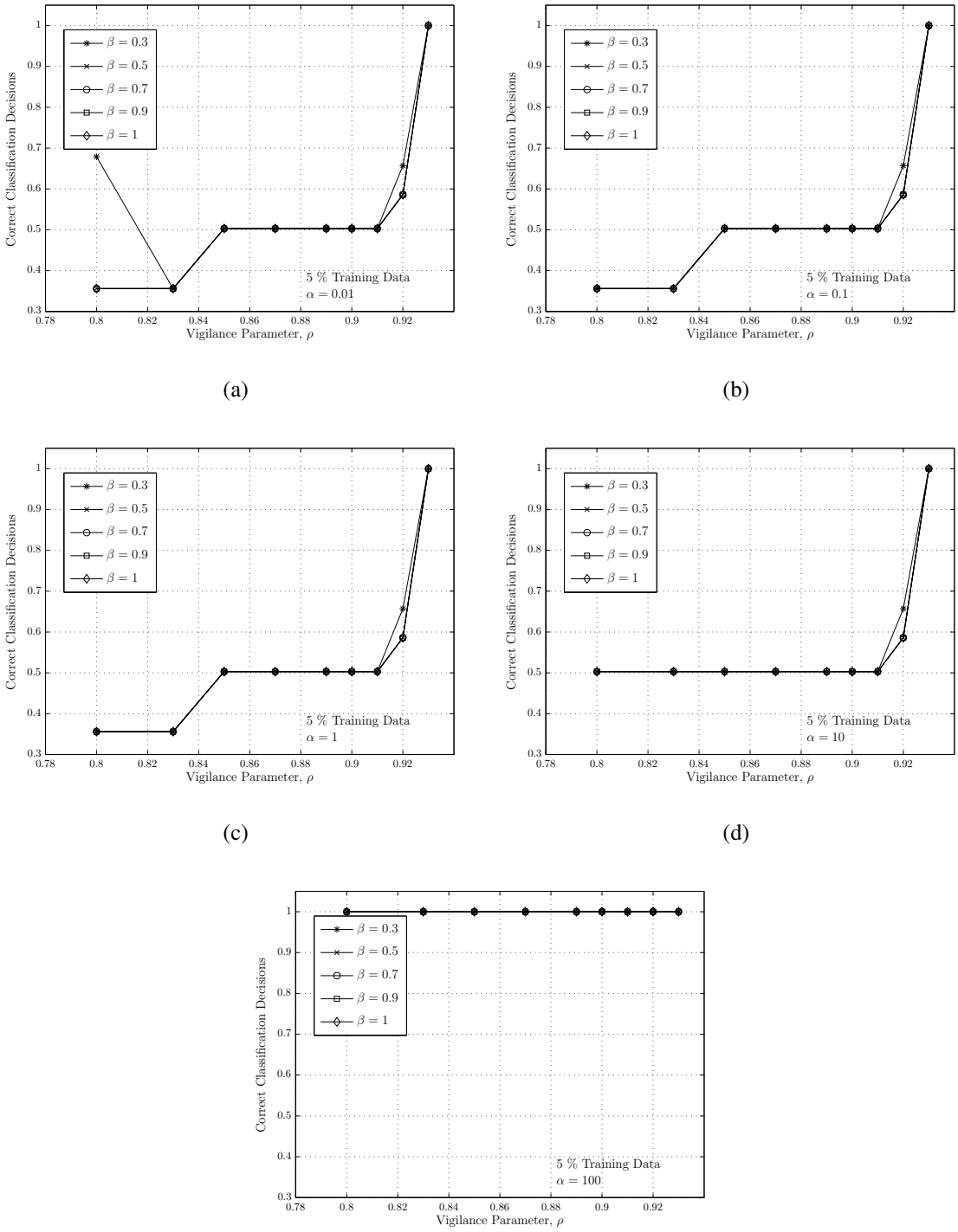
ANALYSIS OF FUZZY ARTMAP PARAMETERS

The results of the search for optimal Fuzzy ARTMAP parameters is presented in this addendum. The search spans across a defined set of parameter values for classifier category choice $\alpha \in [0.01, 0.1, 1, 10, 100]$, learning rate $\beta \in [0.3, 0.5, 0.7, 0.9, 1]$ and vigilance $\rho \in [0.8, 0.83, 0.85, \ldots 0.87, 0.89, 0.9, 0.91, 0.92, 0.93]$ respectively. All these parameter values are used to train the classifier for varying dataset sizes of 1%, 5%, 10%, 20% and 50% respectively. Correct Classification Decision (CCD) performance is used to compare classifier parameter choices for a given training dataset size. Any parameter combination that result in 100% CCD is considered to be optimal. Figures A.1 to A.5 show CCD versus classifier vigilance.



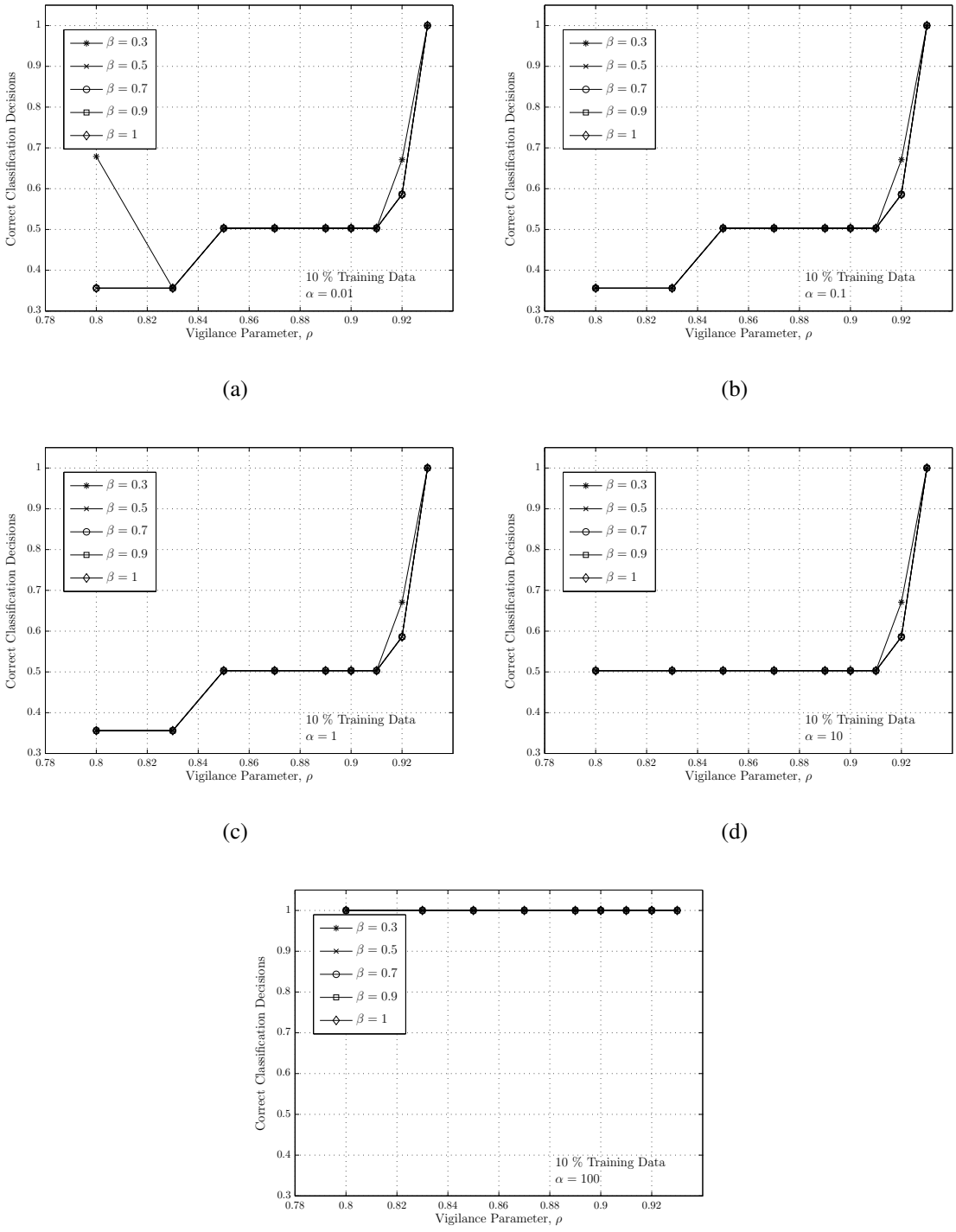
(a) (b) (c) (d) (e)

Figure A.1: The search for Fuzzy ARTMAP parameters that maximise classification accuracy from 1 % training data. Here, CCD is plotted versus the vigilance parameter, ρ across a range of learning rate parameter values β and category choice parameters α. (a) $\alpha = 0.01$, (b) $\alpha = 0.1$, (c) $\alpha = 1$, (d) $\alpha = 10$, and (e) $\alpha = 100$.



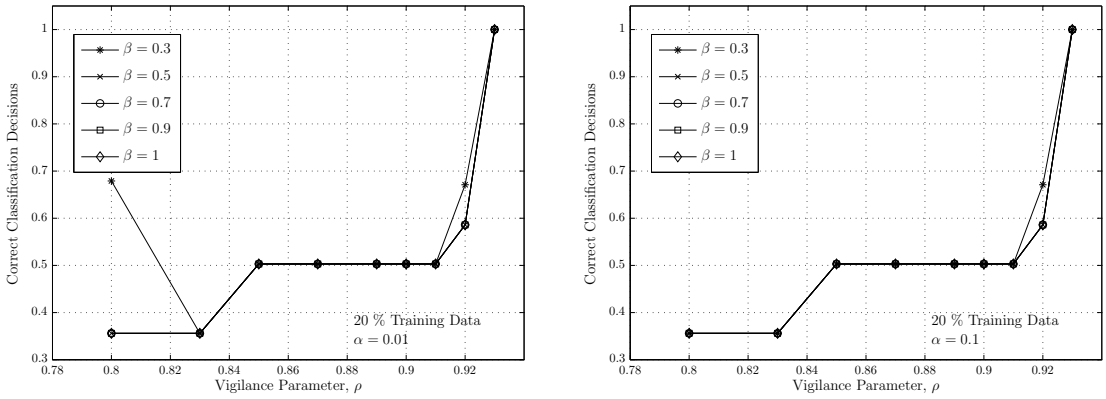
(a) (b) (c) (d) (e)

Figure A.2: The search for Fuzzy ARTMAP parameters that maximise classification accuracy from 5 % training data. Here, CCD is plotted versus the vigilance parameter, ρ across a range of learning rate parameter values β and category choice parameters α. (a) $\alpha = 0.01$, (b) $\alpha = 0.1$, (c) $\alpha = 1$, (d) $\alpha = 10$, and (e) $\alpha = 100$.



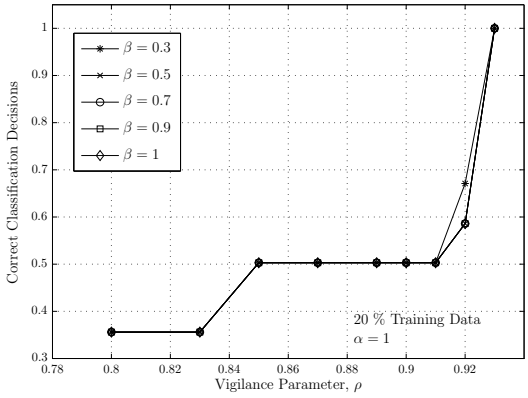
(a) (b) (c) (d) (e)

Figure A.3: The search for Fuzzy ARTMAP parameters that maximise classification accuracy from 10 % training data. Here, CCD is plotted versus the vigilance parameter, ρ across a range of learning rate parameter values β and category choice parameters α. (a) $\alpha = 0.01$, (b) $\alpha = 0.1$, (c) $\alpha = 1$, (d) $\alpha = 10$, and (e) $\alpha = 100$.

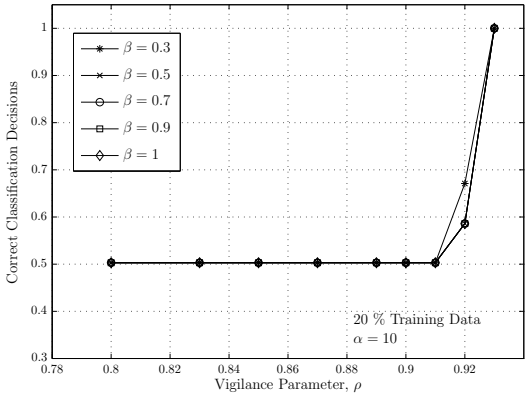


(a)

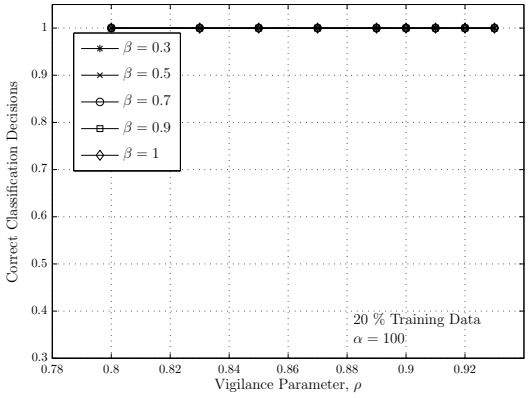
(b)



(c)

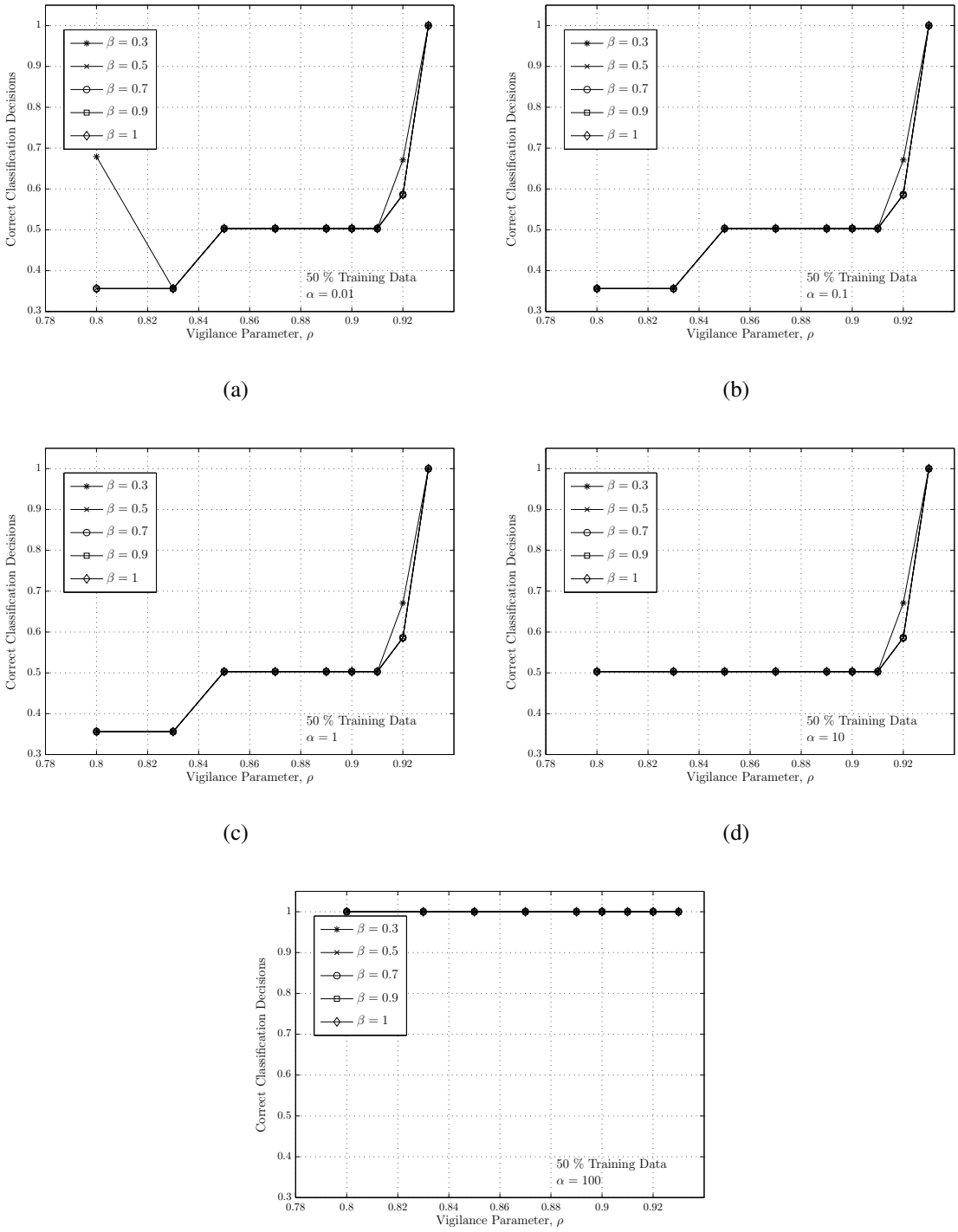


(d)



(e)

Figure A.4: The search for Fuzzy ARTMAP parameters that maximise classification accuracy from 20 % training data. Here, CCD is plotted versus the vigilance parameter, ρ across a range of learning rate parameter values β and category choice parameters α. (a) $\alpha = 0.01$, (b) $\alpha = 0.1$, (c) $\alpha = 1$, (d) $\alpha = 10$, and (e) $\alpha = 100$.



(a) (b) (c) (d) (e)

Figure A.5: The search for Fuzzy ARTMAP parameters that maximise classification accuracy from 50 % training data. Here, CCD is plotted versus the vigilance parameter, ρ across a range of learning rate parameter values β and category choice parameters α. (a) $\alpha = 0.01$, (b) $\alpha = 0.1$, (c) $\alpha = 1$, (d) $\alpha = 10$, and (e) $\alpha = 100$.

ADDENDUM B

DATA PRE PROCESSING AND DIMENSIONALITY REDUCTION

For any classifier the addition of more features in both the input and the output space increases the classification performance as the additional features may contain information that would otherwise not be available. This is however only true up to a certain point where classification performance is reduced when additional features are added. This phenomenon is well known as the *curse of dimensionality* [13,59]. When the majority of input features in a classification problem are correlated with each other, these features have an intrinsic dimensionality. This means that the input data from a subspace of lower dimension can describe the input data just as well as the original amount of input features. Furthermore, there is a strong demand for more data points to be able to train a classifier adequately. For example, M input data points in nineteen dimensions ($d = 19$) are required to properly train the classifier, the input dataset would comprise of M^{19} values. It is clear that the amount of data required for classification quickly grows and may easily render the classification solution impractical.

Thus, in order to explore the intrinsic dimensionality and the large amount of data values of the original amount of input features, original features are pre-processed and only a selected subset thereof is used to train and evaluate the classifier. Here, pre-processing techniques called Linear Discriminant Analysis (LDA) and the Gamma Test are applied. Figure B.1 shows the process by which an original input vector of dimension d, will be reduced to d' (where $d' < d$) and then presented to the classifier.

B.1 LINEAR DISCRIMINANT ANALYSIS

The original feature set data may be reduced through a linear transformation or discriminant. Discriminants are typically used to form decision boundaries for a selection of different classes in typical classification problems. Here, the linear discriminant transformation is applied to the input feature set in order to reduce (discriminate) features and evaluate the effect of a reduced feature set on classifier

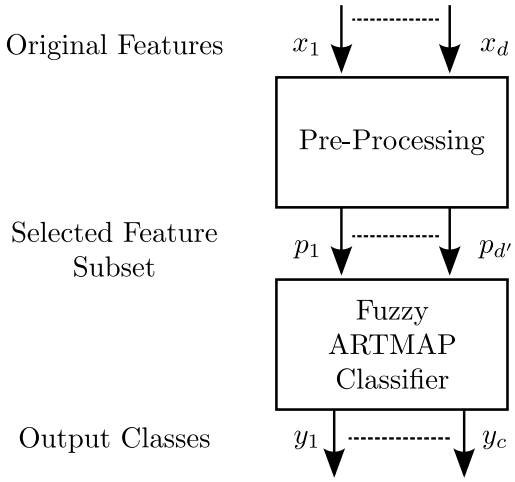


Figure B.1: Process for dimensionality reduction.

performance. If a new feature subset $\mathbf{p}$, of dimension d' is required out of the original feature set $\mathbf{x}$, of dimension d so that $d' < d$, the linear transformation is

$$\mathbf{p} = \mathbf{W}\mathbf{x}\,. \tag{B.1}$$

Thus, $\mathbf{p}$ results by applying a weight matrix $\mathbf{W}$, to the original feature set $\mathbf{x}$. Each of the rows in $\mathbf{W}$ creates an additional dimension in $\mathbf{p}$. After the transformation is applied there must be some criteria for selecting features above others in order to have a feature set with reduced dimensionality. A well known criterion is called the Fisher criterion or Fisher's linear discriminant [13, 60]. This discriminant aims to select features that maximize the between-class covariance matrix, $\mathbf{S}_B$ and minimize the within-class covariance matrix, $\mathbf{S}_W$. This criterion also indicates the separability of data for each class in a classification problem. In the cases where data are not separated adequately, one is forced to consider higher order discriminants (for example quadratic). Thus, the LDA done on a typical classification dataset does not aim to find an optimal decision boundary for classification but the choice of an optimal feature subset. The within-class covariance matrix is calculated for c classes as

$$\mathbf{S}_W = \sum_{k=1}^{c} \sum_{n \in C_k}^{n} (\mathbf{x}^n - \mathbf{m}_k)(\mathbf{x}^n - \mathbf{m}_k)^T\,. \tag{B.2}$$

Here, $\mathbf{x}^n$ represents each of the data vectors (or feature vectors) from class k and $\mathbf{m}_k$ is the mean for class k. For the between-class covariance the only additional variable calculated is the global mean, $\mathbf{m}$. This is the mean across all data vectors irrespective of the class it belongs to. The between-class covariance is calculated as

$$\mathbf{S}_B = \sum_{k=1}^{c} N_k(\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})^T\,. \tag{B.3}$$

Here, N^k is the amount of data vectors in class k. Given the covariance matrices calculated in Equations (B.2) and (B.3) the criterion must be large when $\mathbf{S}_B$ is large and when $\mathbf{S}_W$ is small. In order to determine $\mathbf{W}$, d' eigenvectors are chosen that correspond to the d' greatest eigenvalues of

$$\mathbf{L} = \mathbf{S}_W^{-1}\mathbf{S}_B \, . \tag{B.4}$$

The chosen eigenvectors then form the rows of the weight matrix, $\mathbf{W}$. Thus, feature subsets that are chosen contain the best features based on the LDA criterion described in Equation (B.4). The linear transformation in Equation (B.1) results in a dataset that does not contain any labels. These labels provide the user with an explanation of input features in order to keep track of dimension index values, as it is easier to remember names than numbers.

The results of Fuzzy ARTMAP classifier performance, when evaluated with transformed feature vectors, of dimension, d' is presented in [27]. Here, the classification performance using the transformed (or pre-processed) features is relatively poor. This indicates that the LDA approach reduces the dimensions of the input space at the cost of classification performance. LDA not only reduces the input space dimensions but it also showed that the transformation omits information that the classifier would need to achieve peak performance.

So, what is the reason for the LDA approach countering classifier performance? The eigenvalue decomposition of the Fisher criterion, $(\mathbf{S}_W^{-1}\mathbf{S}_B)$ produces complex valued eigenvectors. This has two implications that ultimately worsen classification performance. Firstly, the transformation in Equation (B.1) results in the original real valued data being converted to complex valued data. The classifier now has to deal with complex values arriving at the input and might not be adequately defined to handle them. Fortunately Fuzzy ARTMAP classification is not influenced by this, as the category choice function, $T_j(\mathbf{I})$, always results in a real valued category index j for an input vector, $\mathbf{I}$. The choice function and its intrinsic equations are given in Equations (3.6) to (3.7) on page 21. Secondly, the complex valued eigenvectors of Equation (B.4) indicate that there exists a higher order surface to separate input features. This surface is of higher dimension than the dimensions of the reduced features after LDA is applied. In other words, the reduced features have an intrinsic dimensionality twice as large after LDA is applied. This is due to the real and imaginary parts that make up complex numbers. The main reason for LDA not providing the desired reduction in dimensionality is due to the feature data being greatly overlapped and some features being highly correlated.

B.2 GAMMA TEST

Similarly with the use of linear discriminants in Section B.1 the Gamma Test [61] is also a tool for finding the optimum dataset for classification. However, the Gamma test is a statistical method used to quantify the underlying uncertainty (or noise) for a given process. In general this process can be described as

$$y = f(x_1,\ldots,x_d)+r \tag{B.5}$$

$$= f(\mathbf{x})+r\,. \tag{B.6}$$

Here, y consists of a function, $f(\mathbf{x})$ of input variables, x_n and an uncertainty term, r. The vector $\mathbf{x}$ represents d dimensional space. The function $f(\mathbf{x})$ represents the classifier. Thus, in order to train a classifier to have a good amount of generalization and accuracy we will evaluate the process in Equation (B.5) for all the dimensional combinations of $\mathbf{x}$ using the Gamma test. The test is not dependent on any form of classifier as long as $f(\mathbf{x})$ is a smooth function. Thus, the Gamma test will point out a subset of dimensions (from the original set) that has the smallest noise variance (uncertainty) with respect to the output, y. The Gamma test can be considered a non-linear equivalent of the sum of squares error used in linear regression [62]. The Gamma test works on the basis that if two points x' and x are close together in the input space then their corresponding outputs y' and y should also be close in the output space. If the outputs are not close together then it is because of noise. The equations required to calculate the Gamma test value, with respect to both the input and output spaces are as

$$\delta_M(k) = \frac{1}{M}\sum_{i=1}^{M}|\mathbf{x}_{N(i,k)}-\mathbf{x}_i|^2\,. \tag{B.7}$$

Here, M is the amount of data points in both the input and output space. $\mathbf{x}_{N(i,k)}$ is the k-th nearest neighbour input vector to the current input vector $\mathbf{x}_i$. $\delta_M(k)$ is known as the input value statistic. The output value statistic, $\gamma_M(k)$ is

$$\gamma_M(k) = \frac{1}{2M}\sum_{i=1}^{M}|y_{N(i,k)}-y_i|^2\,. \tag{B.8}$$

Here, $y_{N(i,k)}$ is the k-th nearest neighbour output value that corresponds with the nearest neighbour input vector, $\mathbf{x}_{N(i,k)}$. $|\cdot|$ represents the Euclidean distance between points in both Equations (B.7) and (B.8). The linear least squares regression between $\delta_M(k)$ and $\gamma_M(k)$ is then calculated so that the Gamma value, Γ is the intersection of the gamma axis, $\gamma_M(k)$ when $\delta_M(k)=0$. This linear relationship is given as

$$\gamma_M(k) = \Delta\delta_M(k)+\Gamma\,. \tag{B.9}$$

It is also shown in [63] that the Gamma value, Γ in probability approaches the variance of r as the number of data points, M approach infinity. Once again with reference to the classification problem of [27] to illustrate the concept. In order to evaluate the Gamma test on subset combinations of the original input features of classifier data, the following steps are performed. The aim is to find all dimensional subset combinations from 19 features as well as the amount of combinations. Next, the Gamma value, Γ is calculated for each subset combination. It is important to note that the order in which the dimensions are arranged does not influence the results of the Gamma test. Then, the feature subset combination that produces the smallest Gamma value and hence the smallest uncertainty, with regard to the process in Equation (B.5), is chosen. As mentioned before, the process is equivalent to the classifier's ability to train on a given dataset. Finally, the Fuzzy ARTMAP classifier is trained on the chosen feature subset and the resulting performance is evaluated. Should the performance improve, there is good enough reason to retain only features from the subset combination that produced the smallest Gamma value. The other features can thus be discarded as the classifier will not loose performance.

B.3 DISCUSSION ON DIMENSIONALITY REDUCTION

Linear Discriminant Analysis (LDA) is quite a useful technique to quantify the separability of features of the input space in almost any classification problem. By calculating the within-class covariance and the between-class covariance along with using a criterion for selection (in our case the Fisher discriminant) one can reduce the dimensions of the feature space quite easily. The main problem with LDA is that it does not provide an effective solution for reducing the dimensions when feature data of different classes are almost totally overlapped. The presence of intrinsic dimensionality where the original features are overlapped neutralise the intended reduction in dimensionality and the aim to retain only those features that cause peak classification performance. Furthermore, the LDA formulation does not have a direct independent link between the input and output spaces. It assumes output classes are linked to input data on account of class labels (which are almost always assigned by the user) and leaves the rest to dimension reduction using a given criterion. The poor results of the LDA application may also be due the choice of criterion function and other criteria could be considered in future.

The Gamma test is an alternative technique to find input features that cause the least uncertainty for a classifier to make accurate class decisions. The result of the Gamma test is the Gamma value, Γ and approaches in probability the noise variance (or uncertainty) as the amount of data points approach infinity. As we are interested in finding a subset of features that will achieve high classification

performance, we used the Gamma test to evaluate dimensional subset combinations of features to find those features that cause the least classification uncertainty. The features not included in the subset are then removed. After performing the Gamma test on 524,286 dimensional subset combinations out of the original 19 dimensions, a feature subset consisting of 7 dimensions was found that achieve peak classification performance. Thus the feature subset that produced the lowest Gamma value can be retained without any loss in classifier performance.

The Gamma test has a significant advantage over LDA in that it directly links output class values with input values belonging to those outputs. The Gamma test is not sensitive to overlapped features, does not alter (transform) datasets that may cause poor classification performance and can generalise almost any type of classifier. The Gamma test is not an *in-line* process as shown in Figure B.1 but requires to be done at least once before the classifier is deployed. The main disadvantage of the Gamma test is that it may take a substantial amount of time to work through all dimensional subset combinations before proceeding with classification.

As mentioned earlier, a feature subset was used to train and evaluate the classifier. Even though the results from the subset did not yield exactly the same classification performance as with the full dataset, it is still very close to that of the full set and very useful [27]. It is definitely worthwhile to use a subset of features if processing speed and memory space are limited. It may also be that the optimum subset was not chosen and another subset may give better results.