Detection of Direct Sequence Spread Spectrum Signals

by

Jacobus David Vlok
B.Eng (Electronic), University of Pretoria, 2003
B.Eng (Hons)(Electronic), University of Pretoria, 2004
M.Eng (Electronic), University of Pretoria, 2006

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

School of Engineering
University of Tasmania
Hobart

October, 2014

Supervisor: Professor J. C. Olivier

I have seen that everything human has its limits and end no matter how extensive, noble and excellent; but Your commandment is exceedingly broad and extends without limits into eternity.

Psalm 119.96 Amplified

STATEMENTS AND DECLARATIONS

1. Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

2. Authority of access

This thesis may be made available for loan and limited copying and communication in accordance with the Australian *Copyright Act 1968* and the South African *Scientific Research Council Act No 46 of 1988*.

3. Statement regarding published work contained in the thesis

The publisher of the papers comprising Chapters 3 to 5 hold the copyright for that content, and access to the material should be sought from the journal. The remaining non-published content of the thesis may be made available for loan and limited copying and communication in accordance with the *Authority of access* statement above.

Signature: J.D. Vlok

Date

4. Statement of co-authorship

The papers comprising Chapters 3 to 5 were co-authored by J.D. Vlok and J.C. Olivier. J.D. Vlok is the primary author of all three papers; he developed and evaluated the algorithms presented therein and wrote the manuscripts. J.C. Olivier contributed in a supervisory role, provided technical guidance and advice in terms of the final presentation of the manuscripts.

Signature: J.D. Vlok	Date
Signature: J.C. Olivier	Date

ABSTRACT

Since early experimentation in the late 1800's, wireless communication has become increasingly important and has been widely adopted by civilian and military markets worldwide. The proliferation of wireless communication systems presents new challenges, threats and opportunities for society and government institutions. Although the possibility of infringing privacy laws exist, electronic surveillance has become an important capability in military, counter-terrorism and law-enforcement operations. Through interception of wireless communication signals, an advantage may be gained by extracting intelligence from, or interfering with, communication signals of an adversary. Interception can only be performed once the presence of the communication signal is detected. However, communication signals are typically not intended for reception by third parties and security mechanisms are often employed to protect communication transmissions from compromise. Sophisticated techniques are therefore required to reliably detect the presence of, and to extract information from, the communication signal of interest.

Due to the ubiquitous use of wireless communication devices, techniques to efficiently use and manage system resources, such as the available radio frequency (RF) spectrum, have been developed and are implemented in these devices to ensure co-existence and to limit interference. Communication systems are also designed to minimise transmission power dynamically, which brings about several advantages, such as enhanced battery life for mobile users and lower detection probability in military applications. Techniques to share resources among several users are also employed in order to increase system capacity and availability. Detecting the presence of a certain communication signal within the resultant dense signal environment is therefore challenging, especially if the intercept receiver does not have accurate knowledge of the parameters being used by the target communication system. The signal of interest will typically be weak, hidden in background noise and among several other competing communication signals.

The detection of communication signals, and specifically weak signals, forms an integral part of modern electronic warfare (EW) in applications of communication surveillance. Signal detection is foundational in extracting parameter values and communications intelligence (COMINT) from radio transmissions, which are important components of com-

munications EW. Knowledge of the communication parameter values of the target radio system must be obtained before further action can be taken to counter potentially hostile communication transmissions. Efficient detection of weak communication signals will therefore enhance the detection capability of communication intercept receivers, and will provide an improved capability to perform interception, direction finding and jamming of these hidden transmissions.

This thesis considers the non-cooperative or blind detection of a specific class of covert communication signals, known as direct sequence spread spectrum (DSSS). DSSS is a low probability of detection (LPD) communications technique, initially developed for military application to hide transmitted messages below the noise floor in order to avoid detection by potential enemy interceptors. DSSS has also become popular in non-military communication systems and is widely implemented in existing wireless communication standards. The popularity of DSSS is due to its interference-rejection, multipath-resistance, co-existence and transmission-security properties, which are desirable for communication in mobile radio channels. As DSSS was designed as a covert communication technique, detecting and demodulating DSSS transmissions present a significant challenge, especially in the non-cooperative context.

The performance of detection algorithms can be expressed in terms of the probability of detection over a range of signal-to-noise ratios (SNRs), although computational complexity should also be taken into account. Sophisticated algorithms which provide high detection probabilities usually also have high computational demands, which will limit their implementation in real-time detection systems. Existing detection techniques are investigated and evaluated in this thesis through mathematical analysis and Monte-Carlo computer simulation, in terms of both detection probability and computational complexity. Most existing detection techniques rely on differentiating between the statistical properties of the signal and the noise in which the signal is potentially hidden, using test statistics based on either energy or correlation characteristics. New and improved detection and estimation techniques, based on similar concepts and eigen analysis, are presented and evaluated in this thesis.

The main body of this thesis consists of three published journal articles, which resulted from the Ph.D. research work, embedded into the text. The first publication presents an approximation to a statistical distribution which can be used to predict the performance of the eigen detection techniques presented here. The second publication presents two new semi-blind DSSS detection techniques, and the third publication considers the blind estimation of the sequence length of DSSS spreading codes. Sequence length estimation is important as several semi-blind DSSS detection and estimation techniques require the sequence length as input parameter.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AAC	aperiodic autocorrelation
AWGN	additive white Gaussian noise
BER	bit error rate
BLAS	basic linear algebra subprograms
BPSK	binary phase shift keying
CDF	cumulative distribution function
CDMA	code division multiple access
COMINT	communications intelligence
DC	direct current
DSSS	direct sequence spread spectrum
ECCM	electronic counter-counter measures
ECM	electronic counter measures
ED	energy detection
EM	electromagnetic
ESM	electronic support measures
ESPRIT	estimation of signal parameters via rotational invariance techniques
EW	electronic warfare
FHSS	frequency hopping spread spectrum
GOE	Gaussian orthogonal ensemble
GPS	global positioning system
GSE	Gaussian symplectic ensemble
GSL	GNU scientific library
GUE	Gaussian unitary ensemble
HOS	higher order statistical
HPC	high-performance computing

i.i.d.	independent and identically distributed
IEEE	Institute of Electrical and Electronic Engineers
ISM	industrial, scientific and medical
LFSR	linear feedback shift register
LLR	log-likelihood ratio
LPD	low probability of detection
LPE	low probability of exploitation
LPI	low probability of intercept
LRT	likelihood ratio test
MIMO	multiple-input multiple-output
MSE	mean square error
NN	neural network
NP	Neyman-Pearson
PAC	periodic autocorrelation
PCA	principal component analysis
PDF	probability density function
PN	pseudo-noise
PSD	power spectral density
PSK	phase shift keying
QPSK	quadrature PSK
RF	radio frequency
ROC	receiver operating characteristic
RV	random variable
SCM	sample covariance matrix
SCvM	Smirnov-Cramér-Von-Mises
SNR	signal-to-noise ratio
SSD	sum of squared difference
SVD	singular value decomposition
TW	Tracy-Widom

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

In certain scenarios where humans or electronic systems interact, it is desirable to hide the fact that communication is taking place. Covert communication can provide an advantage over an enemy in a battle scenario or strategic engagement, but can also be beneficial in a communal setting where neutral entities attempt to communicate over a shared channel without causing undue interference to neighbours. Conversely, knowing that information is being transmitted between two or more enemy entities can also provide an advantage in a military setting. Detecting signal activity is the first step towards intercepting and exploiting potentially valuable information by tapping into enemy communication transmissions. Signal detection is also important in non-military settings, where the presence of a communication signal can serve as an indicator that the channel is currently occupied and cannot therefore be accessed by potential competitors.

1.1.1 Signal detection and transmission security

Covert communication techniques were originally developed to hide communication transmissions for military application in order to prevent detection by enemy receivers [1]. This type of communication is therefore known as low probability of detection (LPD) communication [2]. Closely related to LPD is low probability of intercept (LPI), which refers to communication signals which are difficult to intercept, although not necessarily difficult to detect. Although the terms "detection" and "interception" are sometimes used interchangeably [3, 4], a distinction can be made as follows.

Signal detection can be viewed as the process of merely detecting the presence of a signal without attempting to estimate the signal parameters for the purpose of extracting the message content of the signal [5]. A certain amount of parameter estimation will however be part of the detection process, as certain assumptions need to be made in order to detect the presence of a signal. For example, a receiver needs to tune its centre frequency to a certain band (or scan over the band) where the signal of interest is assumed to be present. If the receiver finds that the signal is indeed present, information regarding the transmission band will therefore also be available, which inherently forms part of estimating the

carrier frequency of the signal of interest. This may however not be sufficiently accurate, and additional information, such as the signal constellation and symbol mapping, may be required to perform demodulation. Furthermore, signal interception encompasses the processes of detecting, receiving and demodulating the signal of interest after the parameters have been estimated, such that the message contained in the signal can be extracted.

LPD and LPI communications are two approaches which can be used to enhance transmission security, which refers to securing radio frequency (RF) or wireless transmissions against compromise, adverse channel effects and jamming signals [2]. LPD schemes to prevent detection include direct sequence spread spectrum (DSSS) signaling, where codes are used to communicate secretly [6]. The code employed by the DSSS system is typically unknown to third parties, especially when military systems are considered. The codes are used to spread the transmitted signal over a wide bandwidth, enabling communication at power levels below the noise floor [7]. Not knowing the spreading code, a third-party receiver cannot take advantage of the processing gain available to the intended DSSS receiver, and must employ alternative strategies to differentiate between the signal hidden in the noise, and the noise itself [5]. In order to prevent interception, LPI schemes such as frequency hopping spread spectrum (FHSS) signaling can be used. FHSS signaling is a technique where the carrier frequency of the narrowband signal is constantly changed or hopped over a much wider bandwidth in a seemingly random pattern. Only the intended receiver that knows the hopping pattern can recover the signal. Although it is difficult to intercept a FHSS signal, it may not be difficult to detect that it is present. A DSSS signal however is difficult to detect, and therefore also difficult to intercept. This concept is illustrated in Fig. 1.1 where LPD is on the top level hierarchy. It is however important to note that LPD and LPI can be combined. For example, the carrier frequency of a DSSS signal can be changed randomly to create a DSSS/FHSS system.

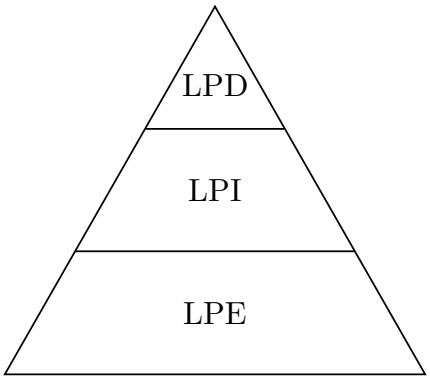


Figure 1.1: Hierarchy of low probability of detection, interception and exploitation.

To protect the actual message being transmitted over the RF channel, another layer of protection can be added using low probability of exploitation (LPE) techniques. One example of such a message security mechanism is encryption, which makes extracting the message from a communication transmission difficult [2]. As illustrated in the hierarchy of

Fig. 1.1, a signal which is difficult to detect or intercept will also be difficult to exploit. It is however possible that a LPD/LPI transmission can be compromised, if the interceptor gains access to the spreading code or hopping pattern. If a wireless communication system employs a separate message protection mechanism in addition, the interceptor may not be able to extract the actual message from the encrypted signal if the signal is successfully detected and intercepted. Nevertheless, detecting the presence of a certain signal (without extracting the message) may still be valuable, as the fact that communication is taking place is information in its own right.

1.1.2 Electronic warfare in wireless communications

Electronic warfare (EW) refers to any action taken to control the electromagnetic (EM) spectrum with the purpose of preserving friendly use and denying enemy use of the spectrum [8]. The focus of EW was initially on radar systems, although applying EW on wireless communication systems gained importance as it became clear that similar needs and opportunities in terms of protection and exploitation exist. The dawn of digital communications and the implementation of techniques such as channel coding and encryption complicated gaining unauthorised access to information-bearing RF signals, which further contributed to the interest of applying EW in wireless communications [1]. As shown in Fig. 1.2, EW consists of electronic attack, protection and support which respectively replaced the older classifications of electronic counter measures (ECM), electronic counter-counter measures (ECCM) and electronic support measures (ESM). Each of these three concepts is subsequently explained with an emphasis on wireless communication transmissions.

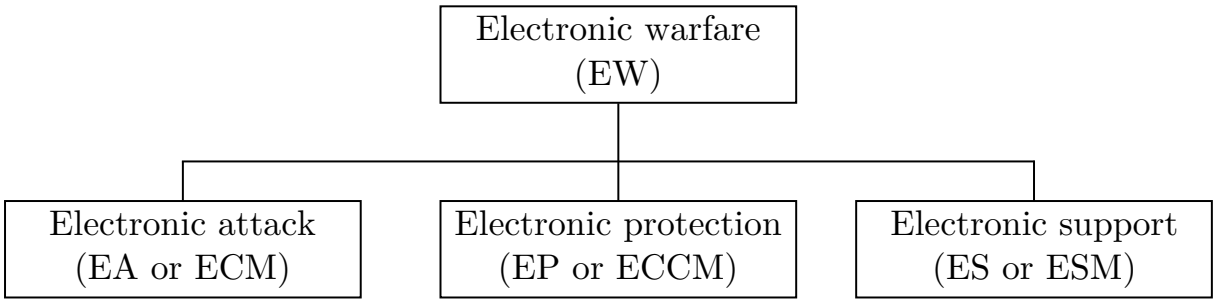


Figure 1.2: The three components of electronic warfare.

1.1.2.1 Electronic attack

ECM refers to methods (e.g. jamming, chaff and flares) to counter (or to interfere with) the operation of radar, communication and threat systems (e.g. heat-seeking missiles). Electronic attack encompasses the older ECM definition, but also includes weapon systems aligned with the purposes of ECM. These weapon systems include anti-radiation

weapons which aim to incapacitate sources of EM radiation using a physical projectile, and directed-energy weapons which transmit energy in a focused direction without using a projectile. An example anti-radiation weapon system is an anti-satellite missile that may be used to destroy a communication or radar satellite, and an example directed-energy weapon application is transmitting high-power RF signals to destroy or saturate the electronics in a receiver system [8].

Electronic attack in communications EW therefore refers to attacking an enemy communication system by either physically destroying elements of the system or by injecting interfering RF (jamming) signals into the enemy communication band with the aim of denying, degrading or disrupting information flow. Jamming involves the injection of noise waveforms, interfering signals or signals masquerading as communication signals (control or information e.g. false messages) into the target receiver [1, 9].

1.1.2.2 Electronic protection

Electronic protection or ECCM refers to measures taken to counter the effects of electronic attack such as jamming [8], and to protect radar or communication signals from compromise, including interception and direction finding. Electronic protection approaches relevant to wireless communications include spread spectrum (see the transmission security mechanisms discussed in Section 1.1.1), channel coding or forward error correction, and null-steering antennas [1]. Protecting communication signals from adverse channel effects is similar to countering enemy jamming attempts and the same techniques (e.g. channel coding) can be applied.

1.1.2.3 Electronic support

Electronic support or ESM refers to the reception of RF transmissions with the aim of supporting electronic attack and also electronic protection. Although electronic support is defined in terms of tactical or quick-response activities (such as providing the carrier frequency of a transmitting enemy communication system to be jammed to the electronic attack system), the electronic support receiver may also monitor wireless communication activity over longer periods to determine the locations and types of enemy transmitters [8].

Closely related to electronic support is communications intelligence (COMINT), which refers to the interception of communication signals with the purpose of extracting intelligence from the information contained in the transmission [8]. The COMINT or intercept receiver employs signal detection and estimation techniques to determine whether a certain type of signal is present from which intelligence may be extracted.

Electronic support is the foundational component of communications EW, and with insufficient electronic support capability, communication electronic attack and protection attempts would be fruitless. Efficient detection of communication signals will enhance communication interception, direction finding and jamming.

1.1.2.4 Tactical and strategic interception

Practical signal detection operations consist of tactical and strategic aspects [2]. Tactical COMINT platforms are typically mobile, have limited processing capabilities, and are usually required to perform real-time analysis in order to interact with and respond to the environment quickly [10]. By contrast, strategic COMINT platforms are typically stationary (for example housed in a building), have greater computational power, and operate over longer time scales to collect and analyse signals [11].

Signals intercepted by tactical platforms can be stored and passed on to strategic platforms for further off-line analysis. Usually the amount of data is large and high-performance computing (HPC) platforms are required to perform analysis. Signal parameters obtained from strategic analysis can then be passed back to tactical platforms to intercept and analyse signals in real time.

Detection algorithms with lower processing requirements are therefore applied in tactical operations, while more computationally complex algorithms are used in strategic analysis. Details regarding which algorithm is used on which platform are proprietary information and not available in the literature, except for declassified information on systems that are no longer in use [12].

Fig. 1.3 illustrates the three elements of EW applied to a COMINT and jamming scenario. The transmitted signal is protected using electronic protection techniques such as spread spectrum signaling or channel coding. The electronic support receiver attempts to extract signal parameters from the transmitted signal of the target communication system, and the communication receiver is the target of the electronic attack system (the use of anti-radiation weapons is not considered here).

In a tactical scenario, real-time analysis of the intercepted signal will be attempted, such that the estimated signal parameter values will be fed to the jammer or electronic attack system to attack the receiver of the target communication system (using jamming waveforms) while the target communication system is still transmitting. The electronic support receiver may also perform direction finding in order to estimate the location of the target transmitter.

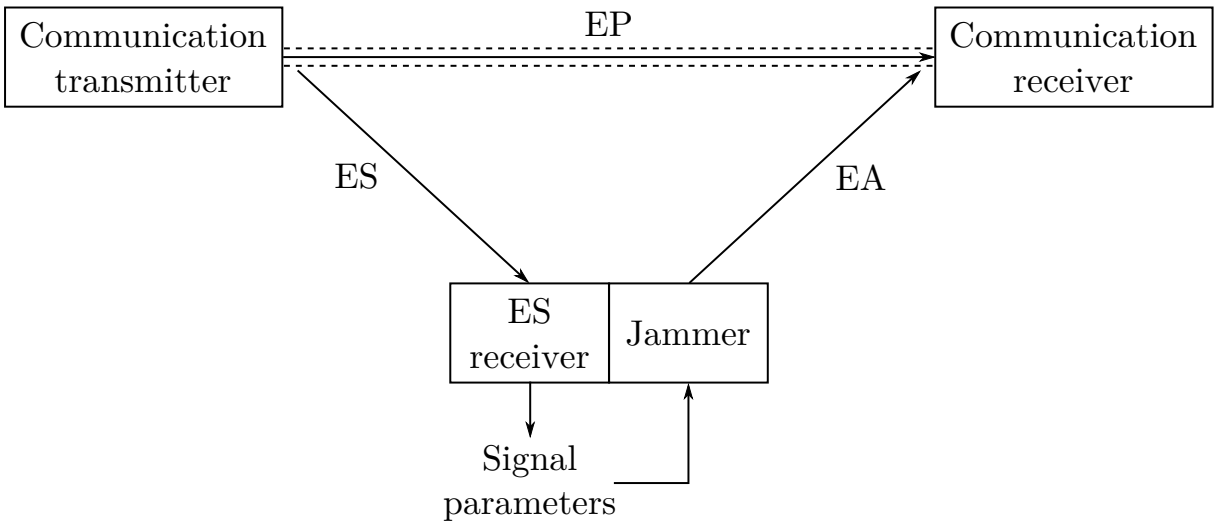


Figure 1.3: Electronic warfare applied in a wireless communication scenario.

In a strategic scenario, the electronic support system (or intercept receiver) will record the intercepted signals for further off-line analysis. The processed signal or estimated parameter values may then be used in future communications EW operations.

1.1.3 Military and non-military communication

Wireless communication has become an integral component of both military and non-military life [13], where non-military refers to the use of communication systems in support of civilian, criminal and terrorist activities. The proliferation and ubiquitous use of wireless communication systems, coupled with the decreasing availability of RF spectrum, have lead communication engineers to develop and adopt strategies to efficiently manage system resources in a dense user environment. These strategies include employing the transmission security techniques discussed in Section 1.1.1, which were originally developed for military applications in order to communicate in unfavourable channel conditions with LPD [2]. These techniques allow several users to communicate concurrently using the same frequency band, while minimising the transmission power. Resultant advantages to the users and operators of a communication system utilising such techniques include enhanced battery life of the mobile device, increased system capacity, improved spectrum utilisation, and lower detection probability for unintended receivers.

A growing trend in the military domain is the increasing reliance on commercial communication devices, and the more prevalent interaction and confrontation with non-military (or unconventional) forces using commercial communication systems [9,13]. This has resulted in a requirement for countermeasure technology to focus more on commercial wireless communication systems. Furthermore, terms like “enemy”, “warfare” and “military” have veered away from their original definitions as traditional conflicts are being replaced with

unconventional and asymmetrical war, and the distinction between military and non-military communication is becoming increasingly vague. Detecting potentially harmful communication transmissions within commercial communication networks and the monitoring and interception of signals from a wide variety of sources have therefore gained importance.

1.1.4 Non-cooperative signal interception

In cooperative[1] communication system design, optimal signal detection techniques such as correlation or matched filtering can be used, as the receiver has perfect knowledge of all the signal parameters (e.g. carrier frequency and modulation type) used by the transmitter [7]. Many RF signals originating from commercial communication systems or networks can also be detected optimally, as the majority of the specifications are openly available. The signal parameters may however vary due to channel effects such as noise, time dispersion and Doppler shift. Estimation techniques are therefore implemented to find and track the parameters around their known values [15]. Detection and estimation techniques are also used to identify the beginning of transmissions (as in asynchronous communication systems) or to identify inactive channels or "white spaces" (as in cognitive radio systems [14]).

However, in non-cooperative applications such as spectrum surveillance and electronic interception, the receiver has little or no knowledge available of the communication parameters used by the transmitter. If no knowledge is available, the intercept receiver must rely on blind detection and estimation techniques, which are typically more complex and processor-intensive compared with non-blind detection. If some parameter information is indeed available, semi-blind techniques may be used. Performing detection in a non-cooperative context may also involve parameter estimation techniques. In order to detect a specified signal, it may be required to estimate some of its parameters to confirm that the detected signal is indeed the signal of interest. Estimation techniques can also be adapted to function as detection algorithms by using the estimated parameters as detection test statistics [16, 17].

Blind detection is further complicated by the fact that target communication signals are typically weak, hidden in background noise, for the reasons given below.

- The intercept receiver is not necessarily located within the communication range or

[1] Cooperative in this sense means the transmitter and receiver are cooperating (they are part of the same communication system or network) in order to achieve successful communication. In cognitive radio [14], cooperative spectrum sensing has a different meaning and refers to several sensing nodes collaborating to determine collectively whether a signal of interest is present or not.

within the main lobe of the transmitting antenna as illustrated in Fig. 1.4. Modern communication systems employ feedback power control mechanisms to conserve energy (and to limit interference) such that the intended receiver receives just enough signal power to recover the transmitted message. The communication range is therefore continually adjusted and minimised, according to changes in the channel caused by movement of the transmitter or receiver, and various other environmental effects [6, 18].

- Certain covert communication systems employ digital signalling techniques (such as the LPD/LPI schemes discussed in Section 1.1.1) to further reduce transmitted power, in order to communicate at power levels below the noise floor [7]. Only the intended receiver, having knowledge of the signalling technique, can therefore recover the transmitted message from the noisy signal with relative ease.

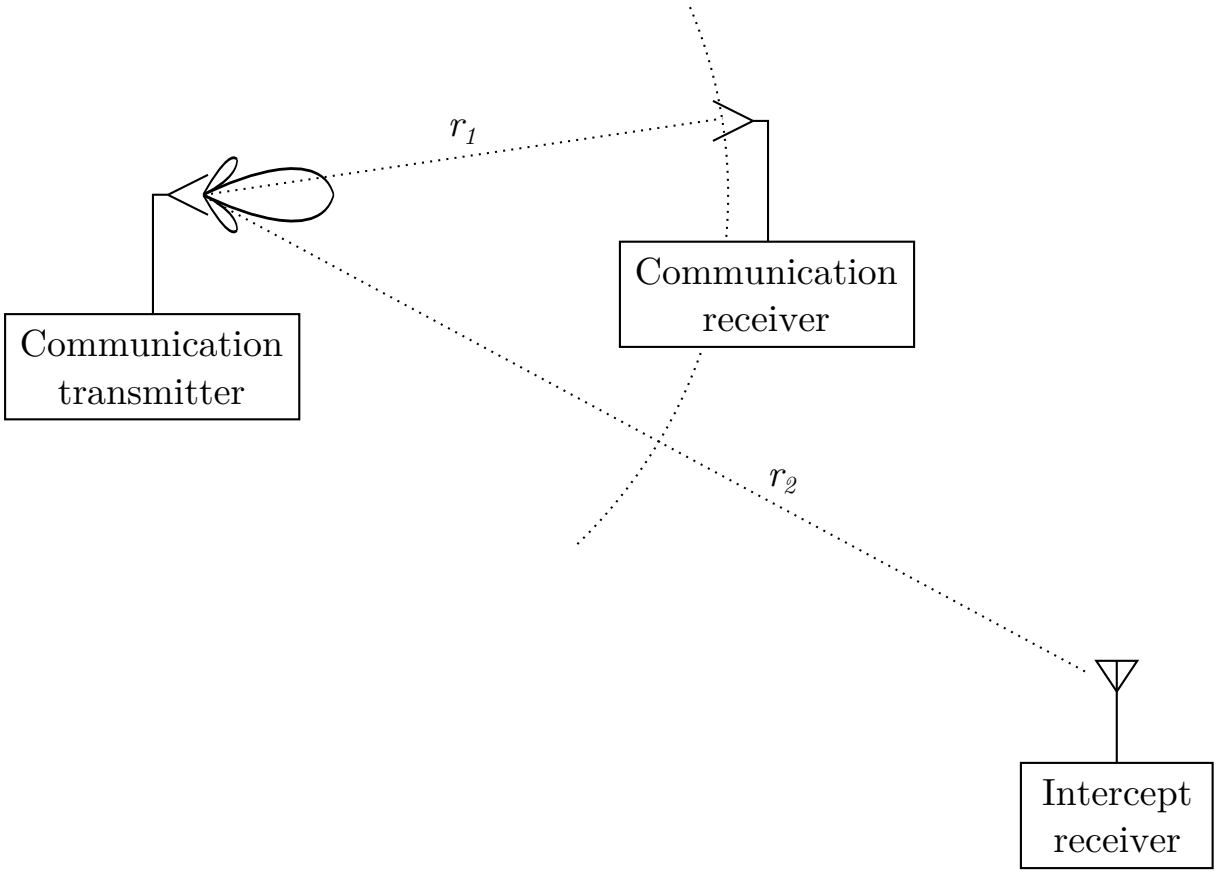


Figure 1.4: Signal interception scenario where r_2 is much larger than r_1.

Detection of a given communication transmission is also difficult in a dense and dynamic wireless signal environment. Due to the multitude of concurrent communication users and their mobility, a large number of simultaneous signals is typically present in the surveillance band with fluctuating signal powers. A specific signal therefore not only competes against background channel noise, but also against numerous other signals, effectively weakening the signal further.

The development and implementation of blind weak signal detection techniques is therefore of critical importance to non-cooperative signal interception, especially when LPD/LPI signals are concerned. When the presence of the signal of interest can successfully be detected, further actions can be taken such as locating or tracking the movement of the associated transmitter, extracting the message content, or denying communication transmission through jamming [2]. Intelligence gathering, direction finding, and communication denial efforts therefore depend on the implementation of efficient signal detection techniques.

1.2 Problem definition

The problem considered in this thesis is the detection of the presence of weak DSSS communication signals with unknown parameter values. Closely related is the estimation of the parameter values that are required to perform detection and to classify the detected signal. As blind detection and estimation techniques are processor-intensive, real-time detection is difficult to achieve [5, 15] and unknown DSSS signals therefore present a significant challenge to the COMINT or intercept receiver.

As mentioned in Section 1.1, DSSS signaling was originally developed as a covert or LPD modulation technique for use in wireless military networks [6]. The increasing demand for wireless communication and the associated diminishing availability of radio spectrum have however resulted in the widespread adoption of DSSS technology in commercial wireless communication standards [6, 19], including wireless local area networks [20], satellite telephones [21], and almost all 3G mobile cellular standards [18].

The popularity of DSSS signaling in wireless communications is due to its interference-rejection, multipath-resistance, co-existence[2], and transmission-security properties [18]. This popularity and the fact that DSSS transmissions are difficult to detect have made it a threat in terms of hiding potentially hostile transmissions below the noise floor and within dense user environments [2, 9].

Furthermore, as military systems increasingly rely on commercial communication technology [13] (and the difference between military and non-military communications becomes less significant), techniques to counter both traditional military and modern commercial DSSS transmissions are required. The development of sophisticated techniques to detect weak DSSS signals is the first step towards achieving this goal.

[2]DSSS signaling can be used to allow multiple concurrent access to the same frequency band, by either assigning orthogonal spreading codes (with strict time synchronisation) or non-orthogonal spreading codes (with multi-user detection) to each participating user or device. These schemes are known as direct-sequence code division multiple access (CDMA) methods as multiple access is achieved through the code domain [22].

1.3 Objectives of work

The objectives of the work presented in this thesis are given below.

- Evaluation of existing communication detection and estimation techniques.
- Development of new and improved DSSS detection techniques.
- Analysis and comparison of the detection performance of existing and new techniques.
- Analysis and comparison of the computational complexities of these techniques in order to determine the feasibility of hardware development and implementation.

1.4 Signal detection approach

The signal detection approach followed in this thesis is based on statistical decision theory where the detector must decide whether the signal of interest is present or not, given a set of noisy data [5]. The classical signal detection approach involves the following steps.

1. A test statistic value is calculated from the noisy data (or received signal).
2. The value is compared with a threshold, which is determined from the detection performance specifications.
3. If the threshold is exceeded, a detection is declared or an alarm signal is activated.

The threshold differentiates between the two hypotheses of "signal present" and "signal absent". The possible outcomes are given in Table 1.1, where the aim is to maximise the detection probability, while constraining the error probabilities.

Table 1.1: Possible outcomes of binary hypothesis testing.

	Detector decision	
Reality	"Signal is absent"	"Signal is present"
Signal is absent ($\mathcal{H}_0$)	Correct rejection	False alarm (Type I error)
Signal is present ($\mathcal{H}_1$)	Miss (Type II error)	Correct detection

The threshold level is an indication of how stringently the hypothesis testing is performed. If the threshold level is set high, only high signal-to-noise ratio (SNR) signals will be correctly detected. Using a high threshold level in low-SNR scenarios will therefore result in a low probability of correct detection P_D and a high probability of miss P_M. Likewise, the false alarm probability P_{FA} will be low and the probability of correct rejection high. Using this threshold-detection approach, it is not possible to minimise both error probabilities (P_{FA} and P_M) simultaneously. Furthermore, P_D and P_{FA} will increase and decrease together as the threshold level is adjusted. A popular approach is to fix one error probability to a constant value, while minimising the other one, as the two error probabilities can be

traded off. The Neyman-Pearson (NP) approach is to fix P_{FA} and then to minimise P_M, which is equivalent to maximising P_D, since $P_D + P_M = 1$ [5].

The detection approach followed in this thesis is based on the NP theorem. The challenge is to design detection algorithms to extract a feature from the received signal that can be used to differentiate between the two hypotheses at low SNRs. Further details regarding the signal detection approach are discussed in Chapter 2.

1.5 Techniques and methods used

The research design (types of studies undertaken) and the methodology followed in this thesis are subsequently discussed.

1.5.1 Research design

1.5.1.1 Theoretical and empirical analyses

The design of new detection and estimation algorithms require both mathematical analysis and empirical tests or numerical evaluation of components of these algorithms. The evaluation of the detection and estimation performances, and the computational complexities also require mathematical and numerical analyses.

1.5.1.2 Statistical modelling

Development and evaluation of the algorithms to be implemented in the intercept receiver require statistical models of the transmitted data and noise. For simplicity of modeling and algorithm development, it will be assumed that the data is uniformly and the noise normally distributed [7, 23]. Data and noise sequences can be generated experimentally using established software implementations of pseudo-random sequences, such as the Mersenne-Twister pseudo-noise (PN) generator implemented in the GNU scientific library (GSL) [24].

1.5.1.3 Computer simulation

Closed-form mathematical analysis of digital communication techniques and algorithms is often not possible due to the level of complexity involved [5, 25]. Monte-Carlo computer simulation provides a solution with a high measure of control, which can be used to predict the performance of communication and receiver systems in the real world under different scenarios. The approach followed in this thesis to determine the performance of each detection algorithm and the effects of several parameters such as the SNR relies on Monte-Carlo analysis of computer models (and mathematical analysis where applicable), which include the receiver architecture, detection algorithms, and channel effects.

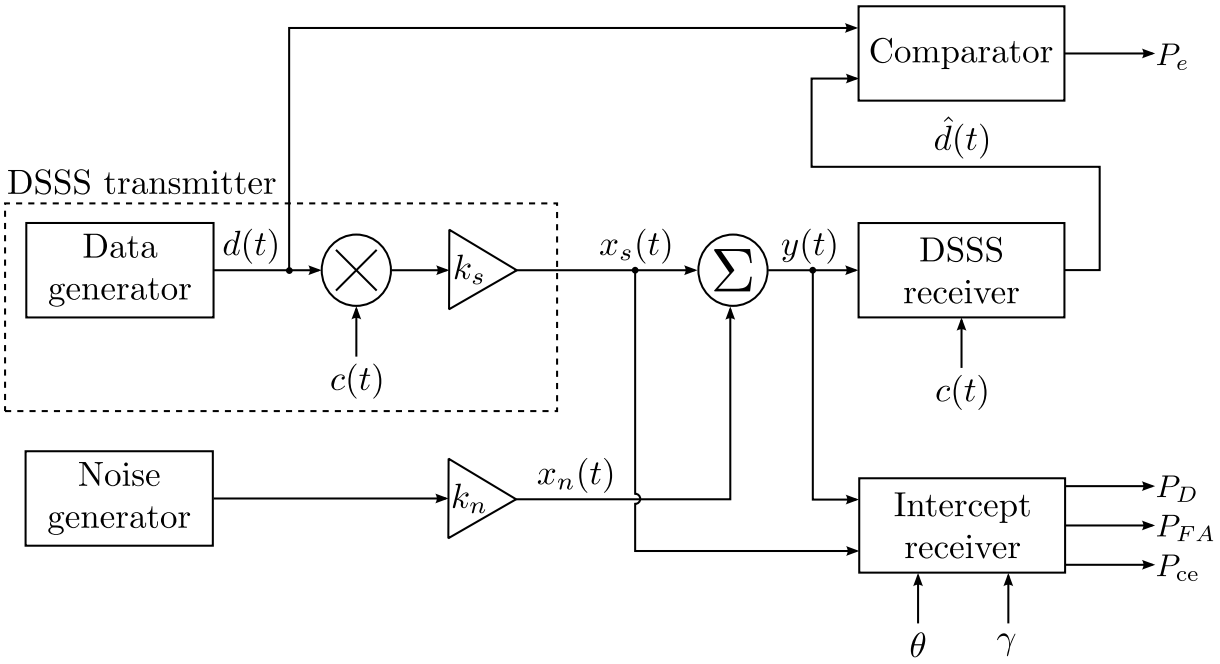


Figure 1.5: Software simulation model of the communication system and intercept receiver.

1.5.2 Research methodology

The research design was executed mainly through computer simulation. The DSSS communication platform and intercept receiver were implemented in C using the basic linear algebra subprograms (BLAS) and GSL libraries [24]. Using this lower-level programming approach provides an execution speed advantage, as performance evaluation of digital communication systems typically requires a very large number of data samples in order to obtain reliable statistics [25].

1.5.2.1 Software simulation model

The baseband simulation model of the DSSS communication system and single-channel intercept receiver shown in Fig. 1.5 were used to evaluate existing and newly developed detection and estimation algorithms. Several assumptions are made in the model to simplify the design, including the fact that no carrier is used and only additive white Gaussian noise (AWGN) channel effects are considered (see also Appendix A).

The data generator shown in Fig. 1.5 produces a binary phase shift keying (BPSK) data signal $d(t)$ with uniformly distributed samples with values ± 1. The data is spread by multiplying $d(t)$ with the spreading code $c(t)$, and the result is amplified by a factor k_s to produce the spread spectrum signal $x_s(t)$.

A noise generator produces zero-mean, unity-variance, normally-distributed samples which

are amplified by a factor k_n to produce the noise signal $x_n(t)$. AWGN is thus introduced by adding $x_n(t)$ to $x_s(t)$ to produce the received signal $y(t)$. The values of k_s and k_n are chosen according to the desired SNR (at the input of the receiver) defined as

$$\text{SNR} = \frac{k_s^2}{k_n^2}. \tag{1.1}$$

The intended DSSS receiver recovers an estimate $\hat{d}(t)$ of the original data from the received signal $y(t)$ using a synchronised copy of $c(t)$. To confirm that the DSSS transmitter and receiver structures are implemented correctly, the error probability P_e is calculated by comparing $d(t)$ with $\hat{d}(t)$ over a range of SNR values. The measured P_e vs. SNR is then compared with the theoretical performance curve of BPSK DSSS in AWGN, which can easily be derived mathematically [26].

The intercept receiver executes the detection or estimation algorithm under evaluation using the received signal $y(t)$ as main input. Other possible inputs include the transmitted DSSS signal $x_s(t)$, the correct value of a parameter to be estimated θ, and the detection threshold γ. The performance indicators provided as outputs are the probabilities of detection P_D, false alarm P_{FA}, and correct estimation P_{ce}.

1.5.2.2 Detection performance

Following the NP approach [5], P_D can be measured (experimentally or in simulation) over a range of SNR values using the threshold γ, calculated from a set P_{FA} value. P_{FA} can also be measured and compared with the set value. The resultant P_D vs. SNR curve can then be compared with reference performance curves, such as energy detection (ED).

To calculate the optimal value for γ, perfect knowledge of the data and noise statistics are assumed. Prior knowledge of whether $x_s(t)$ is transmitted or not, is also assumed to calculate P_D and P_{FA}.

1.5.2.3 Estimation performance

Estimation accuracy is defined in terms of the difference or error between the actual parameter value θ and the estimated parameter value $\hat{\theta}$ [15]. A similar measure, the probability of correct estimation value P_{ce}, was used in this thesis to evaluate the performance of sequence-length-estimation algorithms over a range of SNR values.

Table 1.2: Parameter summary to evaluate different scenarios.

Parameter	**Value/Type**
Channel effect	AWGN
Modulation	BPSK
Spreading code	Barker-11, m-sequence-63, Walsh-64
SNR range	As required such that $P_D, P_{ce} \in [0, 1]$
P_{FA}	0.1 to 10^{-6}
Receiver architecture	Single channel

1.5.2.4 Computational cost analysis

Computational complexity can be derived mathematically in terms of the total number of elementary arithmetic operations $(+, -, \times, \div)$, or the equivalent number of additions and multiplications, required to execute an algorithm [27]. The execution time (or equivalently the number of clock cycles) can also be determined in software.

1.5.2.5 Scenario selection

To evaluate different detection and estimation algorithms, a number of different scenarios must be considered. Each scenario depends on environmental effects, the type of DSSS transmission, and the configuration of the intercept receiver.

Table 1.2 contains a summary of the parameters and their values or types that are considered in this thesis. Representative spreading code types used in communication systems [26, 28] are considered, including a short code (Barker-11), and longer orthogonal (Walsh) and non-orthogonal (m-sequence) codes. The P_{FA} range is chosen to include large values that can easily be verified experimentally, and smaller values in line with practical detection system design [5].

1.6 Outline of thesis structure

The structure of this thesis follows the *PhD by publication* format, where the core of the document consists of three published papers. The outline of the thesis is considered below.

1.6.1 Literature review

An extensive literature review is provided in Chapter 2, where the existing body of relevant knowledge is summarised. The theory considered includes communication signal detection, DSSS technology, and existing detection and estimation approaches relevant to DSSS. Gaps identified in the literature are also discussed.

1.6.2 Thesis core: publications

Chapters 3 to 5 contain the three papers embedded into the thesis text, such that section, figure and table numbering are followed sequentially throughout the thesis. Each paper is however a self-contained publication following from an introduction to conclusion in each case. Some overlap between the three publication chapters and the rest of this thesis is therefore present. The reference list of each paper was however removed and embedded into the reference list of this thesis, provided after Chapter 6 and before the appendices.

The bibliographic details of the three published chapters are given below.

- Chapter 3: "Analytic approximation to the largest eigenvalue distribution of a white Wishart matrix," *IET Communications*, vol. 6, no. 12, pp. 1804–1811, Aug. 2012.
- Chapter 4: "Non-cooperative detection of weak spread-spectrum signals in additive white Gaussian noise," *IET Communications*, vol. 6, no. 16, pp. 2513–2524, Nov. 2012.
- Chapter 5: "Blind sequence-length estimation of low-SNR cyclostationary sequences," *IET Communications*, vol. 8, no. 9, pp. 1578–1588, Jun. 2014.

Following are short descriptions of the three publications.

1.6.2.1 Publication 1

The first publication (Chapter 3) presents an approximation to an eigenvalue distribution that is used in the second publication to express the performance of the newly developed DSSS detection algorithms. The distribution under consideration is that of the largest eigenvalue of a white Wishart matrix, which is formed when performing principal component analysis (PCA) on the detection or data matrix in the noise-only case ($\mathcal{H}_0$).

1.6.2.2 Publication 2

The second publication (Chapter 4) presents new DSSS detection algorithms, based on PCA and cyclic shifting of the data matrix. The algorithms are shown to perform better than classical ED in AWGN, although at increased computational complexity.

1.6.2.3 Publication 3

The third publication (Chapter 5) presents blind methods to determine the spreading code length of DSSS signals hidden within AWGN. Sequence-length estimation algorithms are important to detect DSSS signals, as several detection algorithms assume prior knowledge of the sequence length.

1.6.3 Discussion

The thesis is concluded in Chapter 6 where major findings of the research work are highlighted. The challenge and importance of DSSS detection research, and possible ethical issues surrounding the research topic are also discussed. Finally, possibilities of future research are briefly considered.

1.6.4 Reference list and appendices

Finally, the reference list and a number of appendices are given after Chapter 6.

CHAPTER 2

LITERATURE REVIEW

2.1 Detection of communication signals

Communication signal detection and parameter estimation refer to statistical signal processing performed on signals at the input of a communication receiver. The objectives of such processing are to determine the presence or absence of signals within noise, to classify the signals, and to extract information from the signals [29]. This section considers detection theory applied to communication signals, as the focus of this thesis is on communication signal detection. Parameter estimation theory is briefly considered in Appendix B.

Detecting the presence of a signal is based on statistical decision theory where binary hypothesis testing is applied to decide between the following two hypotheses when analysing the received signal [5].

$\mathcal{H}_0$: The communication signal of interest is absent (only noise is present).

$\mathcal{H}_1$: The communication signal of interest is present within noise.

The possible outcomes of binary decision making are shown in Table 1.1 in Section 1.4. Discriminating between $\mathcal{H}_0$ and $\mathcal{H}_1$ involves the calculation of a function value (the test statistic) from the received signal samples and comparing this value with a threshold. When the test statistic exceeds the threshold value, $\mathcal{H}_1$ is chosen (a detection is declared). The central problem of detection theory is finding the correct function to calculate the test statistic value and making a decision based on this value [5].

There are a number of signal-detection approaches that can be applied, which can be categorised according to the following criteria [5].

- Number of hypotheses:
  - Binary: A decision must be made between only two hypotheses, such as $\mathcal{H}_0$ and $\mathcal{H}_1$ given above.

  - Multiple: A decision must be made between more than two hypotheses. This approach is used for example in classification problems, where a signal or parameter must be identified from multiple possibilities.

- Knowledge of parameters available to the receiver:
  - Simple test: All the parameters are known, such that the probability density function (PDF) describing the received signal under each hypothesis is completely known. This approach is used e.g. in a communication receiver where the signal of interest and the noise statistics are known exactly.
  - Composite test: Unknown parameters are present, such that the PDF describing the received signal is not completely known. This approach is applicable in e.g. detecting RF signals where the carrier frequency and noise statistics are not known completely.

- Use of prior probabilities:
  - Classical detection: Prior probabilities are not used, e.g. in radar systems and blind-detection of communication signals where prior knowledge regarding the occurrence of the signal of interest is not available.
  - Bayesian detection: Prior probabilities are known or assigned to the hypotheses, e.g. in a communication system where the probability of occurrence of symbols is known a priori.

In this thesis, only binary hypothesis testing using a simple classical approach will be considered to derive and evaluate the detection performance of various detection techniques. Statistics of the signal will also be assumed known to the receiver under both hypotheses.

2.1.1 Simple classical approach

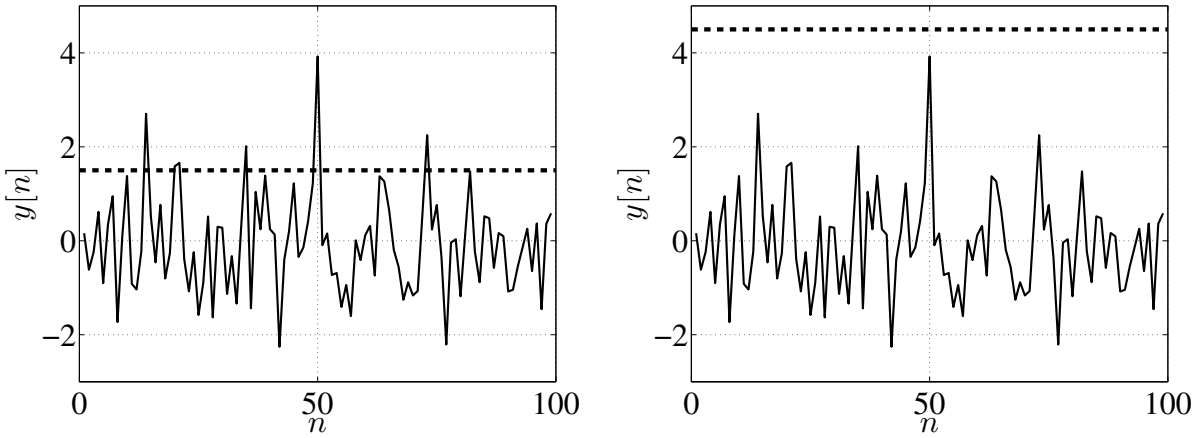
The simple classical approach forms the basis of non-cooperative signal detection from which the other approaches can be understood. Using this approach, the two types of decision errors given below can be made when deciding whether $\mathcal{H}_0$ or $\mathcal{H}_1$ is true (see also Table 1.1).

- Type I error (false alarm): Choosing $\mathcal{H}_1$ when the signal is in fact absent.
- Type II error (miss): Choosing $\mathcal{H}_0$ when the signal is in fact present.

Although one would like to minimise both errors simultaneously, this is not possible using threshold detection.[1] When the probability of false alarm P_{FA} is minimised, the probability of miss P_M is maximised and vice versa. The classical approach is to trade the two errors off against one another.

[1]Both errors may possibly be constrained concurrently using a serial detection approach [30].

An illustration of a sampled noisy signal $y[n]$ where the signal of interest is present at $n = 50$ is shown in Fig. 2.1. By choosing a low threshold as shown in Fig. 2.1(a), several noise samples will exceed the threshold resulting in a large P_{FA}. However, the low threshold will also ensure a small P_M since the signal of interest will usually exceed the threshold and will therefore not be missed. Likewise, by choosing a high threshold value as shown in Fig. 2.1(b), the P_{FA} will be small and the P_M large since the probability that any received sample will exceed the threshold is small (for the case shown in Fig. 2.1(b) no correct or false detections are made). Clearly, the error probabilities depend on the threshold, and the optimal threshold level will result in a trade-off between P_{FA} and P_M. The optimal decision rule can be defined in terms of the likelihood ratio test (LRT) and the NP theorem [5].



(a) Low threshold scenario. (b) High threshold scenario.

Figure 2.1: Different threshold scenarios to detect a signal (the spike at $n = 50$) in noise.

2.1.2 Likelihood ratio test

The likelihood ratio measures the likelihood that hypothesis $\mathcal{H}_1$ is true relative to the likelihood that hypothesis $\mathcal{H}_0$ is true, given an observed data set $\mathbf{y}$ [23]. The likelihood ratio is therefore given by

$$\Lambda(\mathbf{y}) = \frac{p(\mathbf{y}; \mathcal{H}_1)}{p(\mathbf{y}; \mathcal{H}_0)} \tag{2.1}$$

with $p(\mathbf{y}; \mathcal{H}_1)$ and $p(\mathbf{y}; \mathcal{H}_0)$ respectively the likelihood functions of $\mathcal{H}_1$ and $\mathcal{H}_0$ given the data set $\mathbf{y}$. (The difference between likelihood and probability is discussed in Appendix C.) The LRT is a statistical test, based on the likelihood ratio, that can be used to decide optimally between the two hypotheses, given that the PDFs of the data under both hypotheses are known. The LRT is given by

$$\frac{p(\mathbf{y}; \mathcal{H}_1)}{p(\mathbf{y}; \mathcal{H}_0)} > \gamma \tag{2.2}$$

with γ a threshold value, that forms the decision region. The data set that satisfies the inequality of (2.2) falls into the region where the signal is hypothesised to be present (where $\mathcal{H}_1$ is assumed to be true).

2.1.3 Neyman-Pearson theorem

The NP approach [5] is to maximise the detection probability P_D for a fixed false alarm rate P_{FA} using the LRT. If (2.2) is satisfied, the NP theorem states that $\mathcal{H}_1$ should be chosen, where the optimal value for the threshold γ is obtained from the P_{FA} expression. The LRT can also be used to obtain the test statistic function (i.e. how the data must be processed to perform detection) and the expressions for the detection performance.

It is important to note that the decision region defined by the LRT in (2.2) is not affected if the inequality expression is changed by any operation that does not affect the values that satisfy the inequality [23]. The LRT can therefore be simplified, for example, by taking the logarithm on both sides of (2.2).

2.1.4 Example to explain LRT and NP theorem

The example considered here is based on the detection of a direct current (DC) level in AWGN with zero mean [5]. Assuming a single sample is used to perform detection and the signal is the DC level $x = A$, the hypothesis problem can be stated as

$$\mathcal{H}_0 : y = w \tag{2.3}$$

$$\mathcal{H}_1 : y = A + w \tag{2.4}$$

with $w \sim \mathcal{N}(\mu = 0, \sigma^2)$. The likelihood functions are then given by

$$p(y; \mathcal{H}_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{y^2}{2\sigma^2}\right] \tag{2.5}$$

$$p(y; \mathcal{H}_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y-A)^2}{2\sigma^2}\right]. \tag{2.6}$$

The likelihood ratio can then be written using (2.1) as

$$\Lambda(y) = \frac{p(y; \mathcal{H}_1)}{p(y; \mathcal{H}_0)} \tag{2.7}$$

$$= \exp\left[\frac{2Ay - A^2}{2\sigma^2}\right]. \tag{2.8}$$

Using the LRT of (2.2), a detection should therefore be declared if

$$\exp\left[\frac{2Ay - A^2}{2\sigma^2}\right] > \gamma \tag{2.9}$$

according to the NP theorem. As the decision space is unaffected by modifying both sides of the inequality with logarithms, (2.9) can be simplified as

$$\ln \exp\left[\frac{2Ay - A^2}{2\sigma^2}\right] > \ln \gamma \tag{2.10}$$

$$\therefore \frac{2Ay - A^2}{2\sigma^2} > \ln \gamma \tag{2.11}$$

$$\therefore y > \frac{(2\sigma^2 \ln \gamma) + A^2}{2A} \tag{2.12}$$

by collecting all data-dependent terms on the left-hand side of the inequality. The inequality can further be simplified by replacing the threshold value with a single variable as

$$y > \gamma' \tag{2.13}$$

where the new threshold level will be determined directly from the false alarm rate. The inequality in (2.13) indicates that the test statistic is $T(y) = y$, such that the value of the single sample will be used as is to decide between $\mathcal{H}_0$ and $\mathcal{H}_1$.

The false alarm probability P_{FA} and detection probability P_D are related to the likelihood functions as illustrated in Fig. 2.2. P_{FA} can be expressed using the likelihood function for $\mathcal{H}_0$ as

$$P_{FA} = p(\mathcal{H}_1; \mathcal{H}_0) \tag{2.14}$$

$$= p(T(y) > \gamma'; \mathcal{H}_0) \tag{2.15}$$

and since $T(y) \sim \mathcal{N}(0, \sigma^2)$ under $\mathcal{H}_0$ from (2.5), the false alarm probability can be given as

$$P_{FA} = Q\left(\frac{\gamma'}{\sigma}\right) \tag{2.16}$$

as explained in Appendix D. The threshold level is therefore

$$\gamma' = \sigma \, Q^{-1}(P_{FA}). \tag{2.17}$$

The detection probability P_D can similarly be derived as

$$P_D = p(\mathcal{H}_1; \mathcal{H}_1) \tag{2.18}$$

$$= p(T(y) > \gamma'; \mathcal{H}_1) \tag{2.19}$$

$$= Q\left(\frac{\gamma' - A}{\sigma}\right). \tag{2.20}$$

Equations (2.16) and (2.20) define the theoretical performance for the detection problem given by (2.3) and (2.4), which can be attained by using the optimal threshold level defined in (2.17).

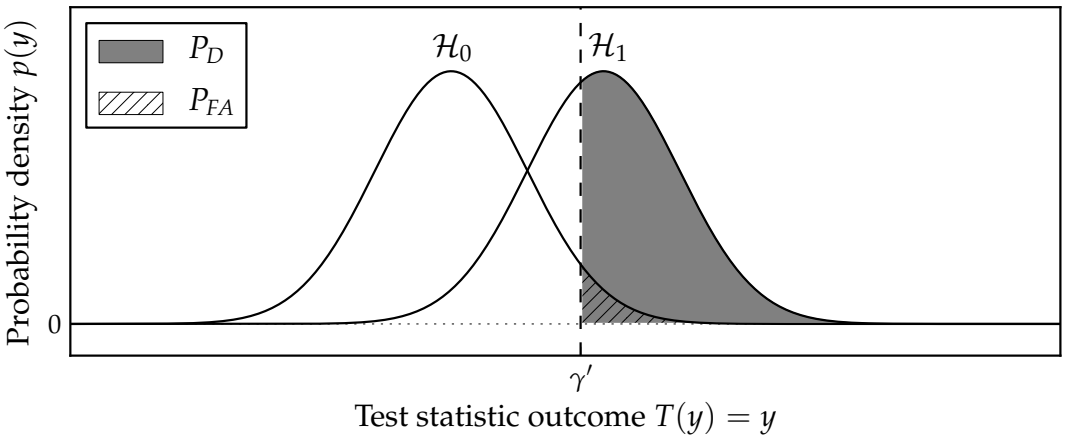


Figure 2.2: Detection performance related to the likelihood functions.

The performance of a detector also depends on the SNR, and two detectors can be compared by evaluating their performances over a range of SNR values. The detector that achieves a higher P_D at the same SNR for a given P_{FA} has better detection performance (though the computational cost of achieving a higher P_D should also be considered when comparing two detectors).

To determine the performance of the example detector as a function of SNR, it should be noted that the power of a DC signal of amplitude A is $P_s = A^2$, and the noise power is simply the variance of the AWGN variable w, given by $P_n = \sigma^2$. The SNR can therefore be expressed as

$$\text{SNR} = \frac{P_s}{P_n} = \frac{A^2}{\sigma^2}. \tag{2.21}$$

The detection performance can then be obtained by substituting (2.17) and (2.21) into (2.20) as

$$P_D = Q\left(\frac{\sigma\, Q^{-1}\left(P_{FA}\right) - A}{\sigma}\right) \tag{2.22}$$

$$= Q\left(Q^{-1}\left(P_{FA}\right) - \sqrt{\text{SNR}}\right). \tag{2.23}$$

Fig. 2.3 shows the theoretical performance obtained using (2.23) and the simulated performance for different P_{FA} values over a range of SNR values. The noise power was chosen as $P_n = 1$ W and the simulated performance was calculated over 1000 iterations for each SNR value. By increasing the number of iterations, the performance will converge to the theoretical curve in each case.

The detection performance can also be displayed as a receiver operating characteristic (ROC), which is a graph depicting P_D against P_{FA} for a fixed SNR value. A number of ROCs using (2.23) is displayed in Fig. 2.4. Clearly, the further a curve is away from the 45° reference line, the better the performance. By allowing a small P_{FA}, a good detector should achieve a high P_D value.

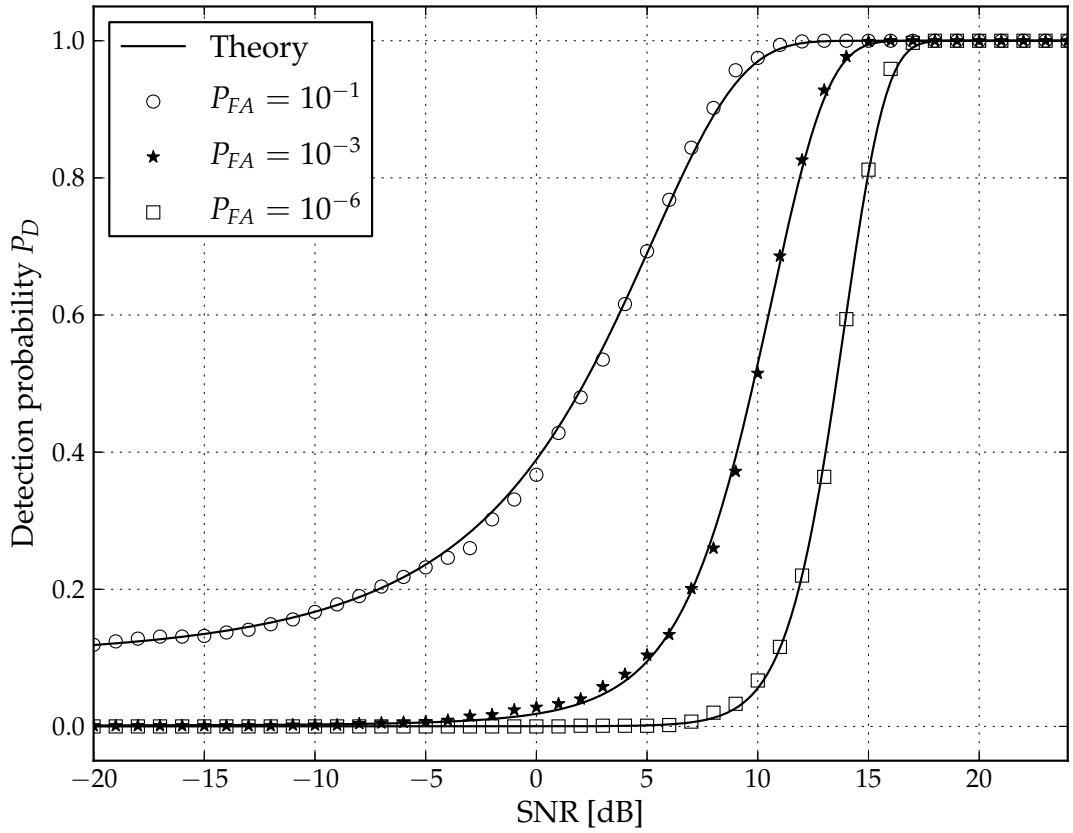


Figure 2.3: Theoretical and simulated performance for DC level in AWGN.

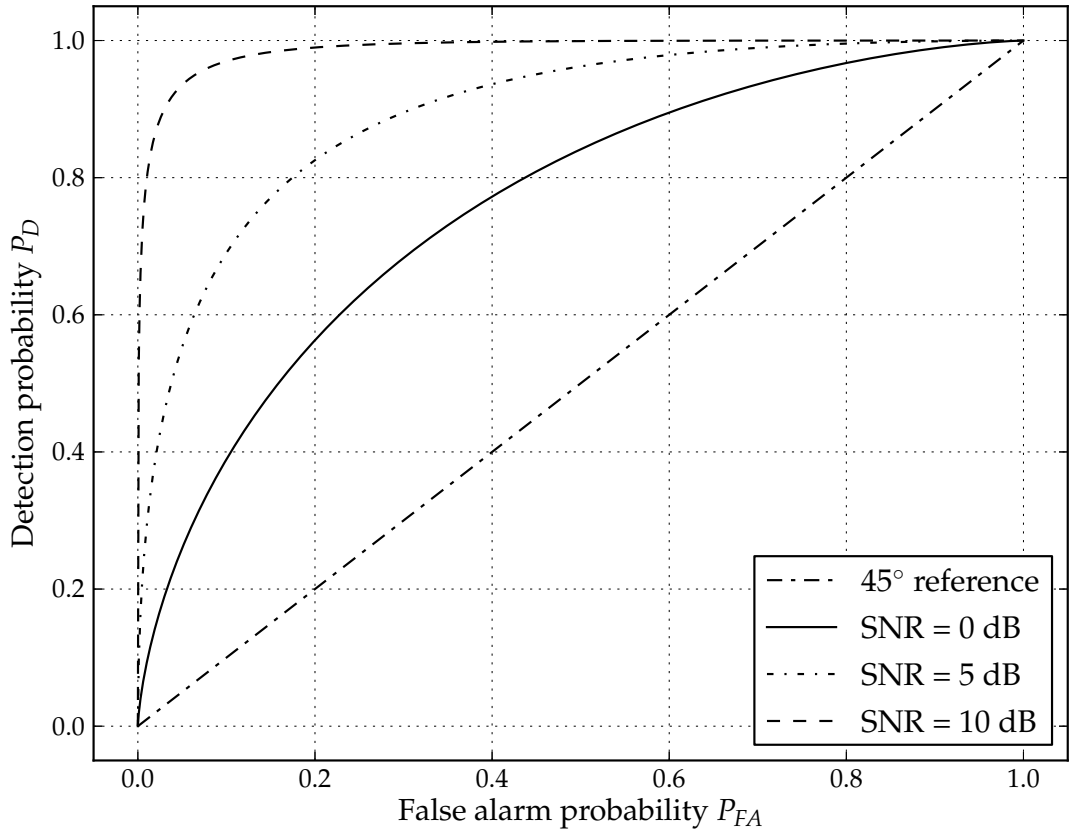


Figure 2.4: Theoretical ROC curves for different SNR values.

2.2 DSSS Technology

DSSS signaling is a class of spread spectrum modulation techniques[2] where the transmitted waveform has a much wider bandwidth than the required Nyquist bandwidth [28] to contain the information signal. The modulated RF bandwidth is determined by some function unrelated to the information signal.

2.2.1 Transmitter and receiver architectures

The essence of spread spectrum communication is expanding the bandwidth of the information signal, transmitting the expanded signal, and recovering the desired signal by collapsing the spread signal back into its original information bandwidth at the receiver [19]. DSSS is typically implemented in combination with phase shift keying (PSK) modulation. In its simplest form, BPSK is used such that the transmitted DSSS signal can be expressed as

$$s(t) = d(t)c(t)\cos(w_c t) \tag{2.24}$$

with $d(t)$ a real input data signal, $c(t)$ the real spreading code, and w_c the sinusoidal RF carrier. The data signal consists of a stream of antipodal bits (taking on values -1 and 1) each having duration T_b, and the spreading code consists of a sequence of antipodal samples or chips each having duration T_c. Higher order modulation formats such as quadrature PSK (QPSK) can also be used, especially when $c(t)$ is complex.

The bandwidth of $s(t)$ is determined by the chip period T_c, while the bandwidth expansion factor is determined by the spreading code length, which is equal to the number of chips in $c(t)$ used to represent a single data bit. The bandwidth expansion factor (or processing gain) can therefore be given as [19]

$$G_P = \frac{W_{RF}}{W_{BB}} = \frac{T_b}{T_c} \tag{2.25}$$

with W_{RF} the spread RF bandwidth and W_{BB} the narrow base-band information bandwidth. Each data bit is thus represented by a sequence of chips (the spreading code) having a bandwidth G_P times that of the data stream. The power spectral density (PSD) of the spread signal is also reduced G_P times (assuming the amplitudes of the original data and spread signals are equal). The bit energy is therefore spread over a wide frequency band, potentially hiding the information signal within noise.

To recover the original data sequence, the DSSS receiver performs coherent demodulation and performs despreading by matching segments of $s(t)$ with a local copy of $c(t)$ using

[2]Other spread spectrum techniques include FHSS, time hopping, pulsed techniques and hybrid forms [19]. Only DSSS will be considered in this thesis.

a correlation-type architecture. Apart from carrier synchronisation (including frequency and phase recovery), an important aspect of recovering the original data is code synchronisation which is required to synchronise the local spreading code (originating from a sequence generator or memory bank) with the spreading code within the received signal. Time synchronisation is obtained through an initial acquisition stage and maintained through a tracking stage to align the incoming signal and the local spreading code, usually via a delay-locked or code-locked loop [19].

2.2.2 Properties of spreading codes

Spreading codes are the signature sequences or keys, shared by the spread spectrum transmitter and receiver, which contain the transmitted message signal after spreading. The properties of spread spectrum transmissions are inherited from the type and quality of spreading codes used. Desired properties of spreading codes are considered below [6,7,28].

2.2.2.1 Low detectability or randomness

Since DSSS is an LPD scheme, the spreading code should ideally be a random sequence resembling white noise. The communication signal and background noise would then be indistinguishable for an unintended receiver. The randomness of a digital sequence can be evaluated in terms of its statistical distribution, predictability of patterns, and correlation between its elements [28]. A spreading code should have a uniform distribution in which the probabilities of occurrence of all symbols in the sequence alphabet are equal. Repeating patterns within the spreading code are undesirable and should be limited to make the sequence less predictable. Lastly, the occurrence of a given value in the sequence must be independent from the others, such that a given symbol value may not be inferred by observing the others.

2.2.2.2 Determinism or pseudo-randomness

Spreading codes are usually generated (or stored) at both the transmitter and receiver, which require them to be deterministic. Pseudo-random or PN sequences provide the possibility for DSSS codes to resemble noise while still being deterministic.

2.2.2.3 Correlation characteristics

An important aspect of spreading codes is their correlation characteristics. Autocorrelation peaks are produced during despreading when the receiver's copy of the spreading code aligns with a spreading code within the received signal. To achieve reliable communication in low SNR conditions, these autocorrelation peaks should clearly exceed the sidelobe correlation values (when the spreading codes are not aligned) and noise [28].

It is desirable that spreading codes exhibit low (ideally zero) cross-correlation sidelobes. Correlation values (both aligned and misaligned) between different spreading codes within a family should also be low. This property affords spread spectrum a co-existence capability, where multiple access to the same frequency band can be provided to several users as in CDMA [28].

The two desired correlation characteristics are therefore: high autocorrelation and low cross-correlation. Correlation can be subdivided into two types, including aperiodic and periodic correlation. The aperiodic autocorrelation (AAC) function of a length-N spreading code $\mathbf{C} = [c_1, c_2, \ldots, c_N]$ can be expressed as [31]

$$R_{cc}(k) = \sum_{n=1}^{N} c_n \; c_{n+k} \tag{2.26}$$

with the shift parameter $k \in [-N+1; N-1]$ since R_{cc} is zero outside these ranges. The shifted index value $n+k \in (0; N]$ since c_{n+k} is zero outside these ranges. Equation (2.26) means the complete spreading code $\mathbf{C}$ is shifted from left to right over itself and the dot product (or partial dot product since a complete dot product is calculated only when $k = 0$) is calculated for each shift. Note that the index value $n+k$ could also be expressed as $n-k$, as the autocorrelation can be performed in either direction (equivalently, (2.26) can be manipulated by subtracting k from the index of c_n and c_{n+k}).

The AAC function only considers the correlation given over one period of a spreading sequence. The periodic autocorrelation (PAC) function, on the other hand, provides the autocorrelation where the sequence repeats itself and is given similarly to (2.26) by

$$R_{cc}(k) = \sum_{n=1}^{N} c_n \; c_{\{(n+k) \bmod N\}} \quad \text{with} \quad c_0 = c_N \tag{2.27}$$

with k now unrestricted due to the modulus operator. When the shifted index value $\{(n+k) \bmod N\} = 0$, such that c_0 is required in (2.27), the chip c_0 must be replaced with c_N. This is to ensure that the index value remains in the range $[1, N]$ as required by the definition of the spreading code $\mathbf{C} = [c_1, c_2, \ldots, c_N]$. The PAC function therefore provides the correlation between a spreading code and a long sequence consisting of several repetitions of the same spreading code. If the sign of a given code within the long sequence differs from the original code, the correlation peak will simply be $-N$ (as opposed to $+N$).

The periodic and aperiodic correlation characteristics are closely related and both are important. However, in a practical burst-communication scenario, several spreading codes will be transmitted in sequence and the periodic correlation characteristics are therefore more important. Good aperiodic correlation characteristics will however translate to good periodic correlation characteristics.

2.2.2.4 Anti-jamming and interference resistance

When a DSSS communication system is attacked with a narrowband interfering or jamming signal, the DSSS receiver will act on the narrowband interferer much like the DSSS transmitter acts on the narrowband information signal. The wideband, low-PSD information-bearing received signal will be despread into a narrowband, high-PSD signal, while the narrowband, high-PSD interferer will be despread into a wideband, low-PSD signal. The resultant effect is that only a small portion of the interfering energy will affect the DSSS signal [18, 28].

2.2.2.5 Multipath resistance

DSSS systems require accurate synchronisation of spreading codes in order to function correctly. Delayed versions of the transmitted signal caused by multipath effects in the channel will therefore typically be rejected, since they are out of phase with the direct-path transmitted signal [18, 28].

2.2.3 Classification of spreading codes

Several different types of spreading codes exist, which are subsequently considered.

2.2.3.1 Analogue and digital codes

Analogue spreading codes are finite-length sequences of which the samples (or chips) are taken from a continuum of values rather than from a finite-size alphabet, as is the case with digital spreading codes [1]. Examples of analogue sequences include Huffman sequences [32] and analogue chaotic sequences [33].

2.2.3.2 Orthogonal and non-orthogonal codes

Orthogonal codes are sets of sequences (belonging to a code family) in which all pairwise cross correlations are zero. Using orthogonal codes allows concurrent transmission of multiple data streams on the same frequency band if the codes can be synchronised or aligned in time. Different data streams may originate from different devices sharing the same communication medium (e.g. in a wireless communication system), or it may involve a single device where more than one channel is used to transmit or receive data (or control information). The cross-correlation properties between codes within a family determine how many concurrent data streams can be accommodated, and how the performance will degrade as the number of data channels is increased [18, 28].

Non-orthogonal spreading codes are sequences that are not designed as part of an orthogonal code family, and their main function is to spread the bandwidth. Non-orthogonal codes can however also be used in multi-user systems, although other strategies (e.g. multi-user detection and scrambling codes as in WCDMA) must then be used to separate different channels or transmitters [22].

2.2.3.3 Real and complex codes

Real codes only contain real components and are typically used in conjunction with BPSK modulation as explained in Section 2.2.1. A complex spreading code can be expressed as

$$c(t) = c_1(t) + jc_2(t) \tag{2.28}$$

with c_1 the real and c_2 the imaginary component sequences. c_1 and c_2 can either be derived from the primitive root of unity [34], or be chosen as two unique sequences of equal length.

Complex codes are typically transmitted on quadrature carriers to isolate the real and imaginary components. Either one or two real (equivalent to one complex) data streams can then be spread using $c(t)$. Depending on the correlation characteristics between c_1 and c_2, a multi-dimensional quadrature modulation architecture [35] can alternatively be used. Complex codes potentially have superior correlation characteristics compared with real codes, depending on the implementation.

2.2.3.4 Linear and nonlinear codes

Spreading codes can be generated either using linear or nonlinear techniques. Nonlinear codes [36] provide enhanced information security properties, since they are difficult to decipher.

2.2.4 Example spreading codes

Following are several examples of popular spreading codes that fall within the classifications discussed in Section 2.2.3 above.

2.2.4.1 Barker codes

Barker codes are short binary antipodal codes ($c_n = \pm 1$) with the property that the sidelobe autocorrelation values are bounded by unity:

$$|R_{cc}(k)| \leq 1;\ k \neq 0. \tag{2.29}$$

When $k = 0$ the autocorrelation reaches its peak $R_{cc}(0) = N$. There are only a few known Barker codes of lengths $N = 2, 3, 4, 5, 7, 11$ and 13 (see for example [28]), which are given in Table 2.1 in binary format. Variations on the codes given in Table 2.1 are also possible; Barker codes remain Barker codes after some transformations such as negation [28]. Barker codes are used in DSSS schemes (e.g. the $N = 11$ code is used in the Institute of Electrical and Electronic Engineers (IEEE) 802.11 (Wi-Fi) standard) and pulse-compression radar.

Table 2.1: All known Barker codes in binary format where 0 represents -1.

N	**Binary code**	N	**Binary code**
2	11 and 10	7	1110010
3	110	11	10110111000
4	1110 and 1101	13	1111100110101
5	11101		

The AAC and PAC functions of three Barker codes ($N = 7, 11$ and 13) are shown in Fig. 2.5. The characteristic of (2.29) is also illustrated. Each AAC function shows an autocorrelation peak of amplitude N at $k = 0$, and sidelobe correlation values varying between 0 and either -1 or $+1$. Each PAC function shows autocorrelation peak values (also of amplitude N) at $k = 0$ and at integer multiples of N. The sidelobe values are however constant at either -1 or $+1$. The peak-to-sidelobe difference is also an important feature that can be used to predict the performance of the spreading code in noise. It is desirable that longer codes have higher peak-to-sidelobe values, although from the PAC graphs shown in Fig. 2.5, the peak-to-sidelobe differences are $[8, 12, 12]$ respectively for $N = [7, 11, 13]$. The Barker-11 code therefore has the same peak-to-sidelobe difference as the Barker-13 code.

2.2.4.2 Maximum-length sequences

A maximum-length (or maximal-length) shift register sequence or simply m-sequence is a PN or pseudo-random binary sequence which can be generated using a linear feedback shift register (LFSR) from a given seed value. A PN sequence is a noise-like sequence, as its autocorrelation properties are similar to those of white noise [26]. A LFSR containing m registers can generate a periodic sequence of maximum length $2^m - 1$ (the m-sequence) consisting of 2^{m-1} ones and $2^{m-1} - 1$ zeros, if the feedback taps are chosen correctly.

The feedback taps (or register positions that should be connected to be fed back to the input through the modulo-2 adder) of the LFSR can be determined using finite or Galois field mathematics. The structure of a LFSR can be expressed using a polynomial of variable X. This polynomial is known as the generator polynomial as it defines the sequence that will be generated with the LFSR.

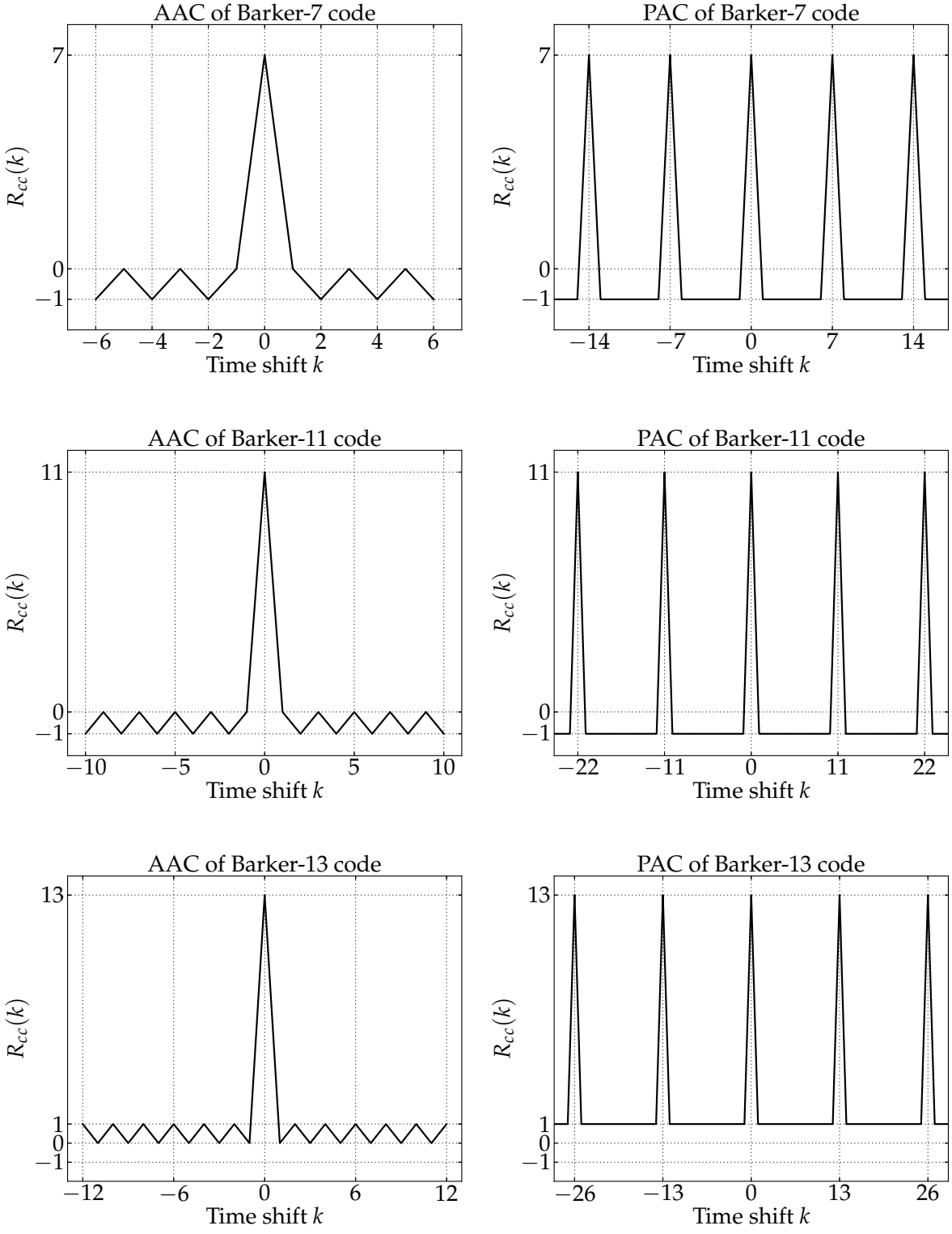


Figure 2.5: Autocorrelation characteristics of several Barker codes.

Fig. 2.6 shows an illustration of a LFSR with generator polynomial terms associated with each register position. The generator polynomial for the LFSR shown in Fig. 2.6 is given in terms of the connected taps as

$$g(X) = X^m + X^{m-1} + X^{m-2} + X^2 + 1. \tag{2.30}$$

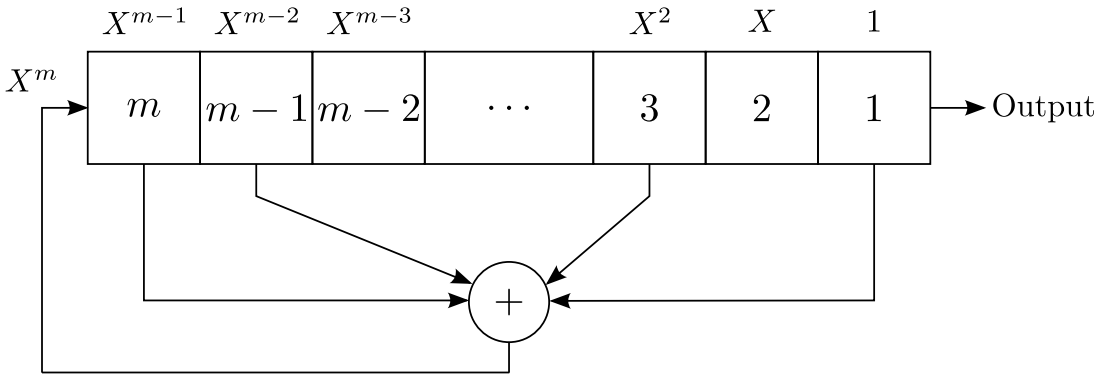


Figure 2.6: Linear feedback shift register with taps connected according to $g(X)$.

If the generator polynomial is a primitive (minimal or irreducible) polynomial of degree m, the LFSR will have maximum cycle length $M = 2^m - 1$ [1]. Such generator polynomials can be obtained by factoring $X^M + 1$ and choosing the irreducible factor polynomials of degree m. Alternatively, the tap positions and generator polynomials can be obtained from tables in literature (e.g. [26, 28]). Table 2.2 shows some examples of the connected taps and generator polynomials for $m = 4, 5, 6$. Each generator polynomial will produce a unique m-sequence. For example, the two $m = 4$ generator polynomials will produce the following two sequences

$$g(X) = X^4 + X^3 + 1 \quad \rightarrow \quad \mathbf{c} = [0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1] \tag{2.31}$$

$$g(X) = X^4 + X + 1 \quad \rightarrow \quad \mathbf{c} = [0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1] \tag{2.32}$$

These sequences can be determined using the LFSR structure or directly from the generator polynomial as the coefficients of the terms of $1/g(X)$ [28].

Table 2.2: Register tap connections and generator polynomial examples for m-sequences.

m	N	**Connected taps**	**Generator polynomial** $g(X)$
4	15	1,4	$X^4 + X^3 + 1$
		1,2	$X^4 + X + 1$
5	31	1,4	$X^5 + X^3 + 1$
		1,3	$X^5 + X^2 + 1$
6	63	1,6	$X^6 + X^5 + 1$
		1,2	$X^6 + X + 1$

Fig. 2.7 shows the AAC and PAC functions of m-sequences of lengths $N = 15, 31, 63$ where the first generator polynomial listed in Table 2.2 was used in each case. The PAC functions show constant sidelobe values of -1 in each case.

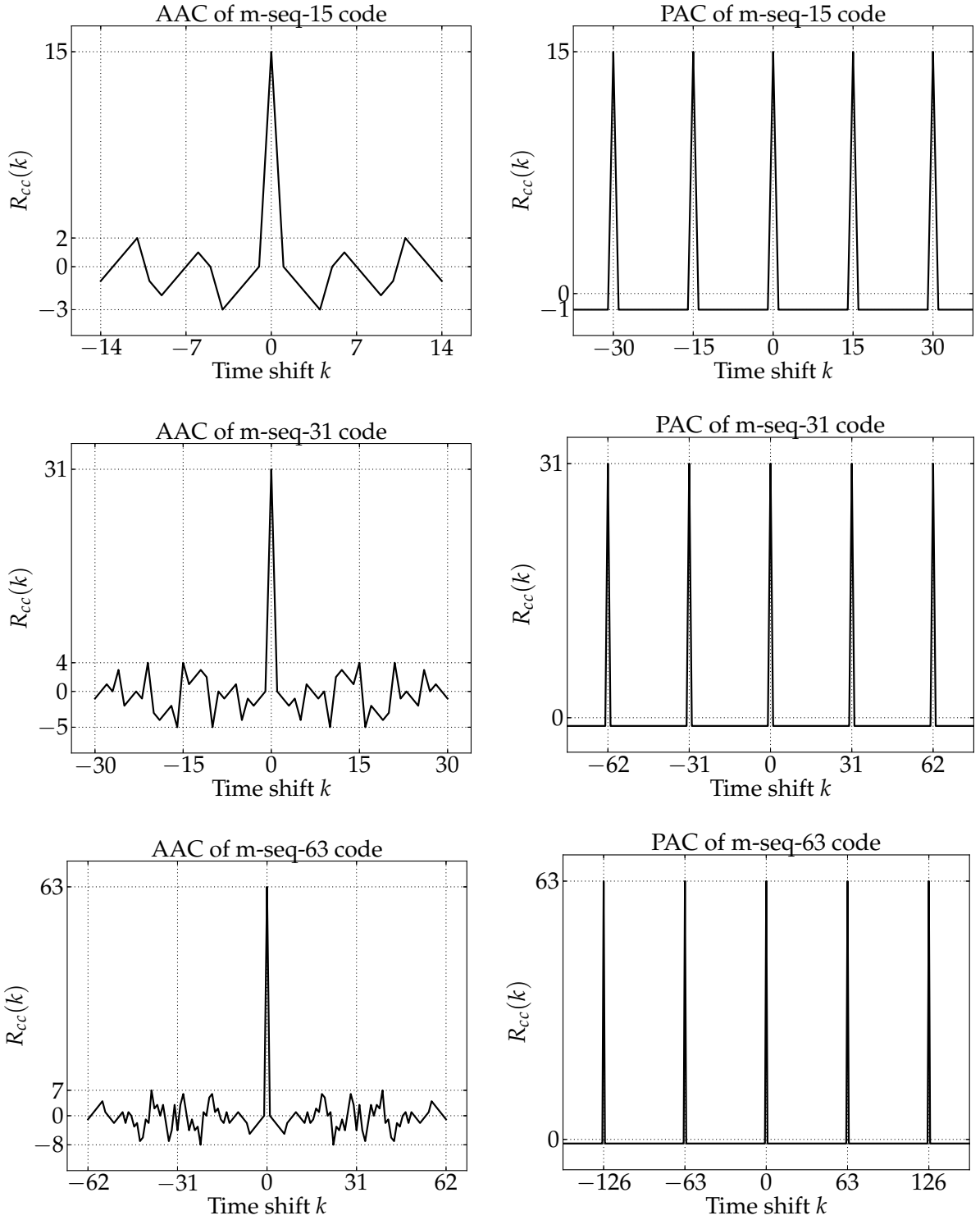


Figure 2.7: Autocorrelation characteristics of several m-sequences.

2.2.4.3 Sequences derived from m-sequences

M-sequences can be used in DSSS or FHSS systems, but generally not in multi-user systems where every user requires a unique orthogonal spreading sequence within a code family. The cross correlation between all unique codes used in such a system should be low (ideally zero) in order to limit interference between concurrent users. M-sequences typically do not exhibit such desirable cross-correlation properties. It is however possible to construct code families with improved cross-correlation properties using m-sequences. Two such code types are Gold and Kasami sequences [28].

Gold sequences

A set of Gold sequences is generated from two specific ("preferred pairs") m-sequences. The first two Gold codes in the set are these two specific m-sequences. For a length-N Gold code (with $N = 2^m - 1$), up to N additional codes can be generated by calculating the modulo-2 sum of the first m-sequence and a cyclic shift of the second m-sequence. For every cyclic shift, a new Gold code is formed. A given Gold code family therefore has a total of $N + 2$ sequences.

Kasami sequences

Kasami sequences are generated similarly to Gold sequences. A Kasami code is produced by the modulo-2 sum of an m-sequence and cyclic shifts of a sequence derived from the original m-sequence. The derived sequence is formed through decimation and replication of the original m-sequence. The cross correlation between unique codes in the Kasami set is smaller than for Gold codes.

2.2.4.4 Walsh codes

Walsh codes are orthogonal codes taken from a $N \times N$ matrix (with N a power of 2) where each row contains a spreading code which is orthogonal to all other rows in the matrix [28]. The matrix is known as the Walsh matrix, which is recursively defined as

$$W_{2n} = \begin{bmatrix} W_n & W_n \\ W_n & \overline{W}_n \end{bmatrix} \tag{2.33}$$

starting from $W_1 = [0]$. $\overline{W}_n$ is the logical complement matrix of W_n. To construct antipodal codes, all 0s are simply replaced with -1s.

Walsh codes are orthogonal because the cross correlation between any pair of codes (or any pair of rows in the Walsh matrix) is zero. The dot product between code words $\mathbf{c}_i$ and $\mathbf{c}_j$ is therefore zero:

$$\sum_{n=1}^{N} c_{i,n}\ c_{j,n} = 0;\ i \neq j. \tag{2.34}$$

Note that the codes need to be aligned perfectly in order for the cross correlation to be zero. Furthermore, Walsh codes also exhibit large autocorrelation sidelobes, which are not desirable. Fig. 2.8 shows the AAC and PAC functions of Walsh codes with lengths $N = 16$, 32 and 64. For each code, the middle row of the Walsh matrix is chosen as the spreading code (e.g. for the $N = 32$ code, row 16 is the chosen code). Due to the large sidelobe values (in both the AAC and PAC functions), the multi-user system that uses Walsh codes must ensure strict synchronisation between all concurrent users accessing the shared transmission medium.

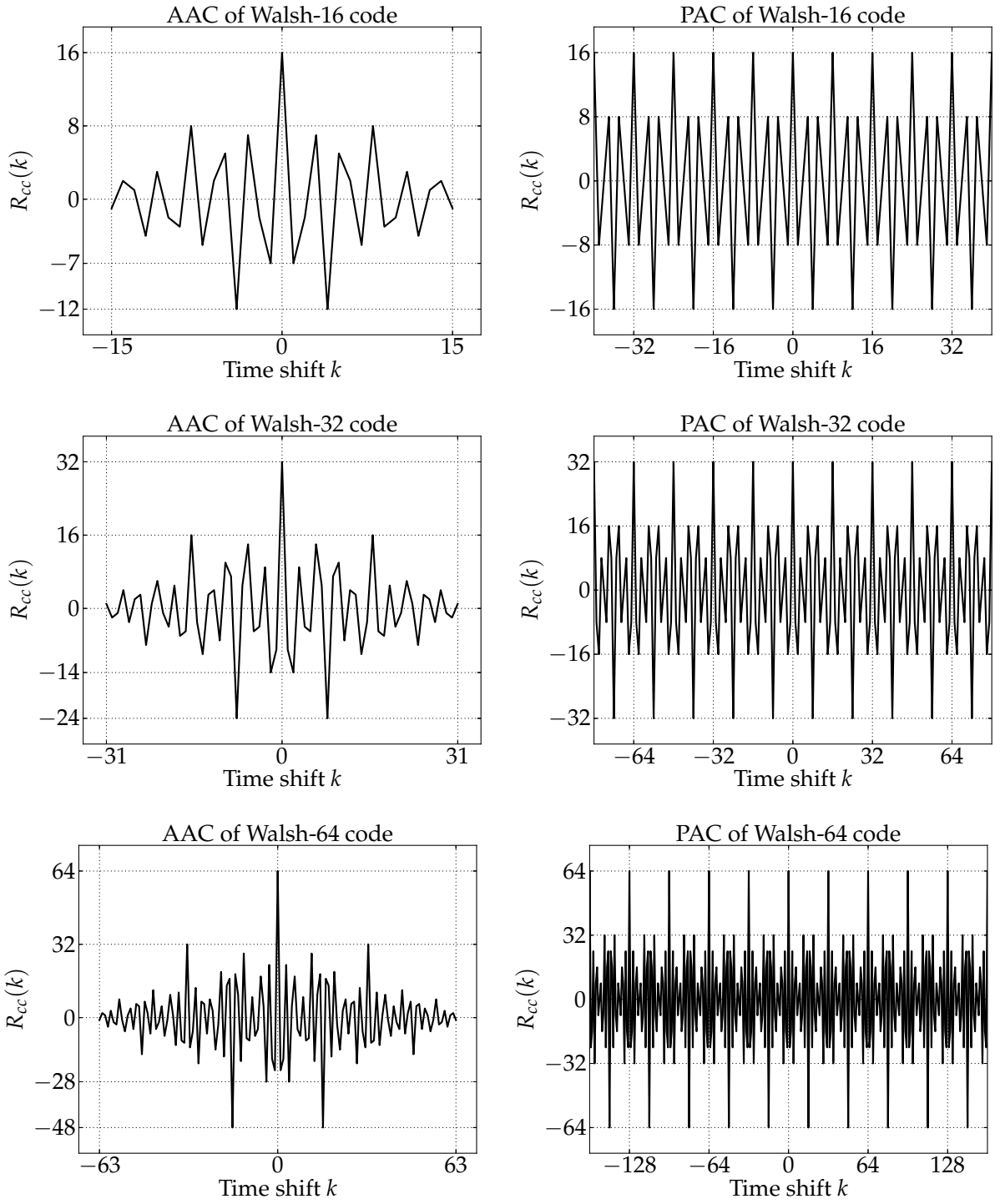


Figure 2.8: Autocorrelation characteristics of several Walsh codes.

2.2.4.5 Complex spreading codes

The example spreading codes considered above are all binary and real. Real codes can be implemented as complex codes by using them on quadrature carriers as discussed in Section 2.2.3.3. Alternatively, complex codes can be designed by choosing their sample (or chip) values from the primitive root of unity. Example complex codes include generalised chirp-like polyphase [34], Frank-Zadoff-Chu [37], zero cross correlation [38], and complementary code keying [20] sequences. Complex spreading codes exhibit improved correlation characteristics in comparison with real codes (although the data throughput rate may be lower) and can also be used to ensure constant envelope transmission [31].

2.2.5 Commercial applications of DSSS

Spread spectrum signaling was originally developed to counter detection, interception and jamming in the military context. The inherent interference-resistance and co-existence capabilities of spread spectrum makes it a lucrative technique with robust performance in mobile wireless channels, where effects such as multipath are prevalent. DSSS communication has therefore, since its military origin, found multiple applications in civilian or commercial systems and also forms part of a number of wireless and mobile communication standards [6]. A number of these applications are discussed in this section.

2.2.5.1 Mobile cellular networks

DSSS is the basis of CDMA technology which was first implemented in the Interim Standard 95 (IS-95) 2G standard [39]. CDMA has also become the primary multiple access technique in almost all 3G mobile cellular systems, such as CDMA2000 and Universal Mobile Telecommunications Service (UMTS), which includes wideband CDMA (W-CDMA) and time division synchronous CDMA (TD-SCDMA). Table 2.3 contains the spreading codes used in these systems [18]. Orthogonal variable spreading factor (OVSF) codes are sets of orthogonal codes with different lengths (supporting different data rates) used to separate physical channels.

Table 2.3: Spreading codes used in mobile cellular networks.

Mobile standard	**Spreading code**
IS-95 (cdmaOne)	Walsh ($N = 64$) Long PN ($N = 2^{15}$)
CDMA2000	Variable length Walsh (maximum $N = 256$) Long PN
W-CDMA	OVSF, PN codes Gold codes ($N = 2^{18}$)
TD-SCDMA	OVSF codes

2.2.5.2 Wireless local area networks

The IEEE 802.11b standard known as Wi-Fi [20] employs the DSSS Barker-11 and complementary-code-keying codes. The standard defines a contention-based wireless protocol employing the carrier-sense-multiple-access-with-collision-avoidance (CSMA/CA) media access protocol operating in the 2.4 GHz unlicensed industrial, scientific and medical (ISM) band. As the spreading codes used in Wi-Fi do not form part of an orthogonal family, multiple access is provided by combining frequency division with the CSMA/CA protocol.

2.2.5.3 Global navigation satellite systems

Satellite navigation systems that use DSSS technology include the USA global positioning system (GPS) [40] and the European Galileo system [41]. The GPS system consists of a constellation of 24 to 32 medium-Earth-orbit satellites and a number of terrestrial monitoring and control stations. The system uses CDMA technology by assigning a unique length-1023 PN Gold code to each satellite for civilian use. The set of codes were carefully selected to optimise CDMA performance; from the family of 1025 codes, only 32 are used. Increased accuracy is available for USA military purposes, where encryption and much longer spreading codes are used.

2.2.5.4 Wireless telephone systems

Wireless telephone systems other than cellular systems that employ DSSS (or CDMA) technology include the following.

- Cordless home telephone systems, which typically operate in unlicensed ISM bands. Spread spectrum technology is therefore used to mitigate interference from similar devices [42].
- Satellite telephones such as Globalstar, which consists of a constellation of 48 satellites offering telephone and low-rate (9600 bps) data communication services using CDMA [21].

2.2.5.5 Radio telemetry systems

Telemetry refers to technology that allows remote measurement and reporting of information. A few examples of such systems, which use a wireless physical layer employing DSSS signaling, are considered below.

- Automatic meter reading systems are used to monitor (and to control) energy and water consumption for the purposes of billing and analyses. Different technologies, including DSSS, are used to transmit data from the measuring units to a central processing station [43].

- Radio-tracking techniques are widely used in research activities to monitor wildlife movements. Such a tracking system utilising a light weight animal tracking device employing CDMA (Gold codes) is considered in [44]. This specific system enables simultaneous tracking of many animals, including small birds.

- Supervisory control and data acquisition (SCADA) systems are typically used in industrial control systems to monitor and control processes e.g. on manufacturing and mining plants. These systems are often connected via RF which uses spread spectrum techniques [45].

2.3 Approaches to detect DSSS

Parameters that can possibly be used to perform detection of DSSS signals hidden in noise include the spreading code length, modulation type, data and chip rates, signal bandwidth, carrier frequency, and the statistical parameters of the signal and noise. These parameters are typically unknown in the non-cooperative context, although some may be estimated. Noise statistics can for example be obtained by monitoring the surveillance band over long periods of time, assuming the signal of interest is not always present.

There are two features of the RF spectrum that change when a wireless communication system transmits information. If the surveillance band is aligned with the portion of the spectrum in which the communication system operates, these two features may be used to perform detection.

The first feature is the energy or power present in the surveillance band which will increase from a reference power level as soon as an information signal is transmitted. Determining the reference power level is a challenge in itself, since it is dynamic and depends on background noise and the behaviour of probably many other communication transmitters that might not be of interest. There are however several techniques which may be used to estimate the signal-free noise level [46].

The second feature is related to the first one and encompasses the statistics of the spectrum. Since information is typically conveyed using signals that exhibit periodicity, the statistics of the received signal will change when a communication signal is present. Knowing what the statistics are when the signal of interest is not present is again a challenge. It can usually be assumed that noise is a stationary process, while communication signals are cyclostationary. Stationarity means the statistics do not change over time, while cyclostationarity means the statistics change with a certain periodic pattern [47]. The cyclostationary nature of communication signals can therefore be exploited to perform detection.

A number of possible approaches that may be used to perform DSSS detection is subsequently discussed. Exact details regarding which detection algorithms are being used on which platforms are either sensitive military or proprietary information and are therefore not widely available in the open literature (see Section 1.1.2.4). However, academic publications, classical textbooks [1, 19, 48] and information in company brochures [10, 11] provide a good indication of what is possible, and which types of techniques are probably implemented.

2.3.1 Energy detection

It can be shown mathematically (using the NP approach) that ED is the optimal detection technique of deterministic signals when no parameter knowledge is available to be used in the detection process [5]. As illustration, the derivation of the detection performance of an AWGN signal hidden within AWGN is given in Appendix E.

2.3.1.1 Ideal energy detection

The continuous-time energy content of a continuous-time voltage signal $y(t)$, over the interval T can be expressed as [48]

$$E_s(t) = \int_{t-T}^{t} y^2(\tau)\ d\tau. \tag{2.35}$$

If $y(t)$ is periodic with symbol period T_{sym} and the integration period $T = T_{\text{sym}}$, (2.35) will provide the signal energy of a single symbol of $y(t)$ at some time instant t. If $T < T_{\text{sym}}$, only a part of the symbol energy will be calculated. The energy of a received signal $y(t)$ can be calculated by implementing (2.35) using a sliding-window integrator. Alternatively, an integrate-and-dump approach can be followed [48], which calculates the energy in non-overlapping segments of $y(t)$ every T seconds, which may or may not coincide with T_{sym}.

2.3.1.2 Energy detection receivers

Devices that employ ideal ED to measure the energy of RF signals are known as radiometers [48, 49]. Fig. 2.9 shows a single-channel ED receiver structure consisting of a radiometer and an alarm circuit. The received signal $r(t)$ is first passed through a band-pass filter tuned to a certain spectral region in order to reduce the effect of noise, and to focus on a specified portion of the RF spectrum. To cover a wider surveillance bandwidth, a filter-bank system (consisting of a number of parallel single-channel receivers focussing on separate narrow bands) or a frequency-scanning receiver (where the centre frequency of the filter is swept over a frequency range) can be used [48]. The alarm circuit must decide between the two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ as discussed in Section 2.1. The alarm circuit here is a simple thresholding device that analyses the energy measurement $E_s(t)$ and declares a detection if a predefined energy threshold is exceeded.

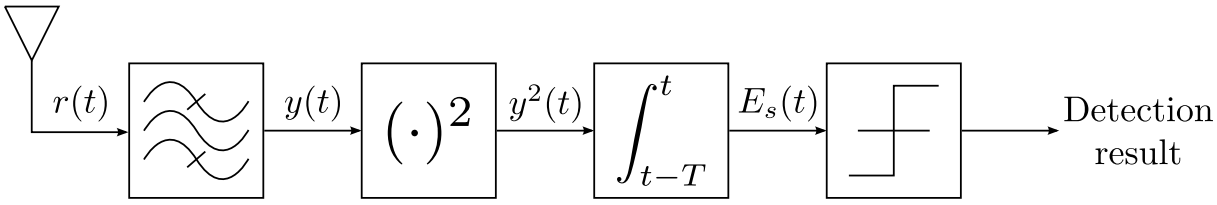


Figure 2.9: Single-channel radiometer-based receiver structure.

2.3.1.3 Performance of energy detection

The detection result provided by the intercept receiver of Fig. 2.9 can be quantified in terms of P_{FA} and P_D as explained in Section 2.1. The test statistic function of ED for a single window of the received signal can be given by (2.35) in its discrete form as [5]

$$T(y) = \sum_{n=1}^{N} (y[n])^2 \tag{2.36}$$

with $y[n]$ the $n^{\text{th}}; n = 1, 2, \ldots, N$ sample of the discrete-time equivalent of $y(t)$ where the window size is N samples. If a deterministic signal $x(t)$ is to be detected in noise $w(t)$, the binary hypothesis problem can be stated in terms of the sampled signals as [5]

$$\begin{aligned} \mathcal{H}_0 &: y[n] = w[n] \\ \mathcal{H}_1 &: y[n] = x[n] + w[n]. \end{aligned}$$

If the signal of interest $x(t)$ is a base-band BPSK DSSS signal, with symbol amplitude A, the ED test statistic can be expressed as

$$T(\mathbf{y}) = \begin{cases} \sum_{n=1}^{N} (w[n])^2 & \text{under } \mathcal{H}_0 \\ \\ \sum_{n=1}^{N} (\pm A + w[n])^2 & \text{under } \mathcal{H}_1 \end{cases} \tag{2.37}$$

In order to calculate the detection performance, it is necessary to determine the statistical distributions of the test statistic under each hypothesis. The sum of squares of N independent standard normal random variables (RVs) $\mathcal{N}(0, 1)$ is distributed according to the central Chi-squared distribution with N degrees of freedom $\mathcal{X}_N^2$ [5]. By normalising the received samples under $\mathcal{H}_0$, it can therefore be shown that (see Appendix E.4)

$$\mathcal{H}_0 : \frac{T(\mathbf{y})}{\sigma_w^2} \sim \mathcal{X}_N^2 \tag{2.38}$$

under the assumption that $w(t)$ is AWGN with variance σ_w^2.

Likewise, the normalised sum of squares of N independent normal RVs (where the n^{th} RV has mean μ_n and variance σ_n^2) is distributed according to $\mathcal{X}_N'^2(\lambda)$, which is the non-central Chi-squared distribution with N degrees of freedom and non-centrality parameter given by [5]

$$\lambda = \sum_{n=1}^{N} \left(\frac{\mu_n}{\sigma_n} \right)^2 . \tag{2.39}$$

It can therefore be shown that

$$\mathcal{H}_1 : \ \frac{T(\mathbf{y})}{\sigma_w^2} \ \sim \mathcal{X}_N'^2(\lambda) \tag{2.40}$$

with

$$\lambda = \sum_{n=1}^{N} \left(\frac{\pm A}{\sigma_w} \right)^2 = \frac{A^2}{\sigma_w^2} N = \frac{\sigma_x^2}{\sigma_w^2} N = \text{SNR} \times N. \tag{2.41}$$

The false alarm rate can then be expressed as (see Section 2.1.4)

$$\begin{aligned} P_{FA} &= p(T(\mathbf{y}) > \gamma'; \mathcal{H}_0) && (2.42)\\ &= p\left(\frac{T(\mathbf{y})}{\sigma_w^2} > \frac{\gamma'}{\sigma_w^2}; \mathcal{H}_0 \right) && (2.43)\\ &= Q_{\mathcal{X}_N^2}\left(\frac{\gamma'}{\sigma_w^2} \right) && (2.44) \end{aligned}$$

from which the optimal threshold level can be calculated as

$$\gamma' = \sigma_w^2 \, Q_{\mathcal{X}_N^2}^{-1} \left(P_{FA} \right). \tag{2.45}$$

Similarly, the detection probability can be obtained as

$$\begin{aligned} P_D &= p(T(\mathbf{y}) > \gamma'; \mathcal{H}_1) && (2.46)\\ &= p\left(\frac{T(\mathbf{y})}{\sigma_w^2} > \frac{\gamma'}{\sigma_w^2}; \mathcal{H}_1 \right) && (2.47)\\ &= Q_{\mathcal{X}_N'^2}\left(\frac{\gamma'}{\sigma_w^2} \right). && (2.48) \end{aligned}$$

To detect the signal of interest in a practical or simulation setup, the threshold level given in (2.45) will be required. However, to draw theoretical detection performance curves, (2.45) can be substituted into (2.48) to obtain

$$P_D = Q_{\mathcal{X}_N'^2}\left(Q_{\mathcal{X}_N^2}^{-1} \left(P_{FA} \right) \right) \tag{2.49}$$

which is plotted in Fig. 2.10, for different P_{FA} values and $N = 100$.

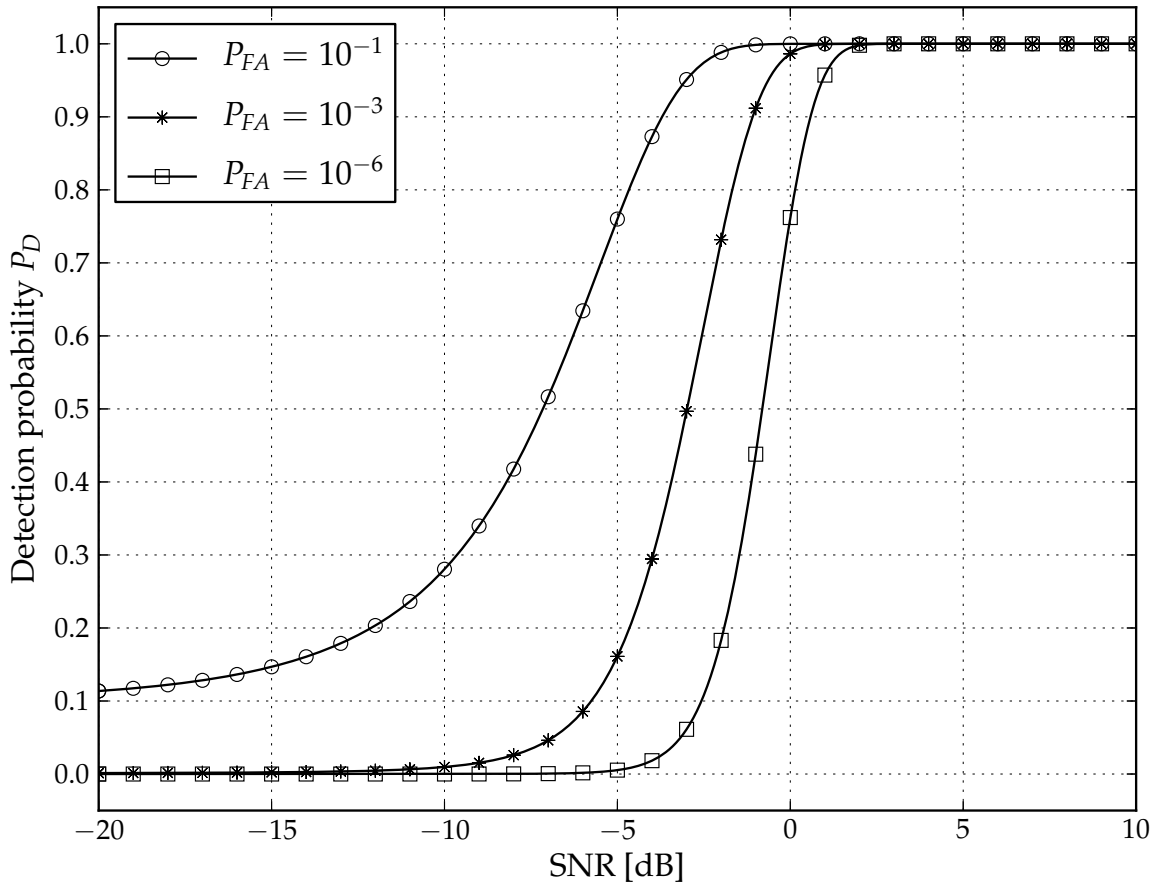


Figure 2.10: Theoretical energy detection performance bounds for $N = 100$.

The curves shown in Fig. 2.10 represent the theoretical ED performance bounds, as perfect knowledge of the parameters and PDFs under both hypotheses were assumed to derive (2.49). These performance curves are valid in general when ED is applied to signals in zero-mean AWGN, such that each received sample is normally distributed under both hypotheses.

It can however be expected that actual detection performance will be worse than depicted in Fig. 2.10, as perfect noise estimates will typically not be available to the detector. Uncertainty regarding the noise level will result in an inaccurate threshold level (see (2.45)), which will negatively affect detection performance.

Assuming energy is calculated over N consecutive samples (as in Fig. 2.10) using a sliding window, the value N does not have to equal the symbol period (or spreading sequence length) of the signal of interest. The results in Fig. 2.10 are valid as long as the signal of interest (or a section of the signal) is present over the entire energy window (irrespective of whether N matches the symbol period or not), as the detector compares the calculated energy (not necessarily equal to the symbol energy) over the window with a threshold level determined from the noise-only energy level. In these conditions, the underlying (and unknown) symbol period will not influence the energy calculated over the length-N window.

The sliding-window energy detector therefore does not require accurate knowledge of the symbol period of the signal to be detected.

However, if the detector calculates energy over non-overlapping and consecutive segments of the received signal, the position and period of symbols within the received signal become important. In such conditions, the results shown in Fig. 2.10 will only be valid if each segment contains at least one complete signal symbol and no noise-only samples. The sliding-window approach can therefore be expected to perform better, especially for blind detection applications.

2.3.1.4 DSSS detection

For the intercept receiver, DSSS falls under the class of unknown deterministic signals [5, 49]. The signal parameters are typically unknown or difficult to obtain due to the low SNR of DSSS signals. ED is therefore an important detection technique when non-cooperative detection of weak DSSS signals is considered [48–51].

2.3.1.5 Critical evaluation of energy detection

ED provides a very simple (low computational complexity) method to perform detection, since only the sum of squares of intercepted samples is computed. Due to the simplicity of the mathematics involved, theoretical ED performance curves can easily be derived by assuming perfect knowledge of the signal and noise statistics (see Section 2.3.1.3).

However, the theoretical ED performance curves also indicate that relatively high SNRs are required to successfully perform detection [52]. ED might therefore not be the optimal detection technique in low-SNR scenarios, as are typical for DSSS transmissions. ED does not capitalise on the cyclic features of communication signals [3] and might therefore easily be outperformed by other detection techniques.

The theoretical ED performance curves may also be useful as reference curves for Monte-Carlo simulation (or experimental results) of more advanced detection techniques and scenarios that are mathematically intractable.

2.3.2 Autocorrelation detection

The cyclostationarity of DSSS signals and similar communication transmissions can be exploited by correlating the received signal, or a portion thereof, with itself. Detection is then performed by analysing the output of the correlation process for the presence of correlation peaks. Correlation detection may also be used in combination with ED [48].

2.3.2.1 Single-channel system

The simplest form of correlation detection is to isolate a segment $\tilde{y}(t)$ from the intercepted signal $y(t)$ and to correlate the two as illustrated in Fig. 2.11. Assuming the SNR is sufficiently high, the output of the correlator R_{yy} will contain correlation peaks if $\tilde{y}(t)$ contains a spreading code (or a large section thereof) that repeats within $y(t)$. These peaks can either be positive or negative which can all be made positive using a squaring (or absolute value) device. The detection test statistic R_{yy}^2 is then used as input to a thresholding device to decide whether the DSSS signal is present or not. The threshold value can be calculated using the required P_{FA} and noise statistics, in a similar way as for ED explained in Section 2.3.1.3.

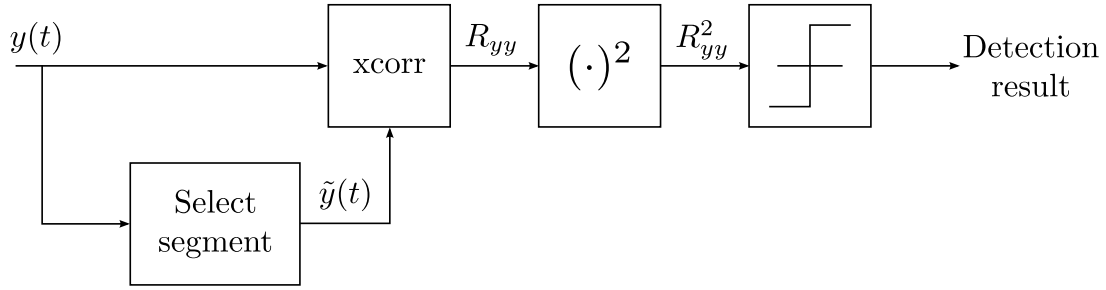


Figure 2.11: Single-channel autocorrelation detection system.

2.3.2.2 Dual-channel system

Correlation detection can also be implemented using two spatially separated receiver antennas [48, 53]. The two received signals are then correlated and the output analysed for the presence of correlation peaks. Provided the antennas are sufficiently separated and isolated, the two receivers will produce uncorrelated noise waveforms, which will have a smaller effect during the correlation process compared with the correlated noise samples produced using the single-channel system. The dual-channel system will therefore produce stronger correlation peaks resulting in a performance advantage over the single-channel system.

2.3.2.3 Compounded autocorrelation

A compounded autocorrelation approach uses the "fluctuations of autocorrelation estimators" [54, 55] or mean-square correlation sequence to perform detection. The intercepted signal $y(t)$ is split into M segments of duration T each, and the correlation is then calculated for every segment using [54]

$$R_{yy}^{(m)}(\tau) = \frac{1}{T}\int_0^T y(t)y^*(t-\tau)\,dt \tag{2.50}$$

where $m = 1, 2, \ldots, M$ is the segment number.

The mean-square correlation sequence is then calculated using [54]

$$\rho(\tau) = \frac{1}{M} \sum_{m=1}^{M} \left[R_{yy}^{(m)}(\tau) \right]^2 \tag{2.51}$$

which will produce large correlation spikes that will exceed a threshold value (calculated from the noise-only statistics) if the DSSS signal is present with sufficiently high SNR.

2.3.2.4 Critical evaluation of autocorrelation detection

Although autocorrelation detection can outperform ED, a relatively high SNR or many samples that contain several repetitions of the DSSS code are still required. The performance of this technique is however limited because a noisy signal segment is correlated with another noisy signal segment from the same intercepted signal. The noise samples are therefore correlated (if the single-channel correlator is used), which will hamper the performance, as the correlation peaks will be less distinguishable from the noise. Detection performance can however be improved by ensuring that the noise samples are uncorrelated by using two (or more) independent receiver channels feeding into the same correlator [53].

2.3.3 Higher-order statistical analysis

The autocorrelation function and power spectrum are important second order statistical functions which also form part of the higher order class [56]. However, higher order statistical (HOS)[3] analysis usually refers to correlation and spectral analysis beyond the second order, and is defined in terms of cumulants [57] and cumulant spectra [56]. HOS analysis can be used in detection and estimation problems to suppress AWGN, to reconstruct signals, to identify nonlinearities in stochastic processes, and to perform classification [58].

2.3.3.1 Cumulant sequences and spectra

The N^{th} order spectrum of $x(t)$ can be expressed in terms of the Fourier transform of its N^{th} order cumulant sequence c_N as [56]

$$C(w_1, w_2, ..., w_{N-1}) = \mathcal{F}\left\{c_N(\tau_1, \tau_2, \ldots, \tau_{N-1})\right\}$$

$$= \sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{N-1}=-\infty}^{\infty} c_N(\tau_1, \tau_2, \ldots, \tau_{N-1}) \exp\left\{-j(w_1\tau_1 + \cdots + w_{N-1}\tau_{N-1})\right\} \tag{2.52}$$

with w the frequency and τ the time-shift operators. The second-order cumulant sequence (or autocorrelation function) of the zero-mean, stationary random process $x(t)$ can be written as [56]

$$c_2(\tau_1) = R_x(\tau_1) = E\left[x(t)\ x(t+\tau_1)\right]. \tag{2.53}$$

[3]HOS can also refer to higher order spectrum/spectra/spectral.

Similarly, the third-order cumulant sequence (or triple correlation function) can be written as

$$c_3(\tau_1, \tau_2) = E\left[x(t)\ x(t+\tau_1)\ x(t+\tau_2)\right]. \tag{2.54}$$

By substituting (2.53) into (2.52), the power spectrum $C(w_1)$ is obtained by taking the Fourier transform of the autocorrelation function (also known as the Wiener-Khinchin theorem [26]). Similarly, when (2.54) is substituted into (2.52), the bispectrum $C(w_1, w_2)$ is obtained.

An advantage of the bispectrum (and trispectrum and even higher orders) over the power spectrum is the fact that the bispectrum of a Gaussian signal has a mean of zero for all combinations of frequency pairs (w_1, w_2) [59]. The bispectrum of a non-Gaussian signal however exhibits a specific pattern with non-zero means. HOS can therefore be used to identify deviations from Gaussianness [56] which may be applied to detect DSSS signals within AWGN.

2.3.3.2 Detection based on higher order statistics

Higher order cumulant sequences of spreading codes show unique patterns in terms of the location and magnitude of peak values, which may be used to identify and discriminate such sequences [60]. Similarly, the bispectrum also exhibits well-defined features with a peak at $(w_1, w_2) = (0, 0)$. Statistical tests may therefore be conducted to detect DSSS transmissions using averaging and thresholding techniques [59].

2.3.3.3 Spectral correlation and cyclic feature detection

The cyclic spectrum [3] is similar to the bispectrum where correlations between power spectral elements are represented on a two-dimensional or bi-frequency plane. The cyclic spectrum can exploit the cyclostationarity of DSSS signals by extracting their cyclic features in order to perform detection [4].

2.3.3.4 Critical evaluation of higher-order statistical analysis

HOS analysis has higher computational requirements than second order techniques, such as the autocorrelation detection technique of Section 2.3.2. Claims that HOS may provide better detection performance have been made [52], although only limited results to support these claims have been published. Potential performance gains achievable by using these HOS detection algorithms should therefore be investigated further.

2.3.4 Time-frequency analysis

Signal detection may also be performed using time-frequency analysis, which refers to methods where the received signal is concurrently analysed in the time and frequency domains. Frequency analysis techniques, such as the Fourier transform, provide a spectral view of the signal without temporal localisation [61]. The amplitude (or energy) of each spectral line is related to the strength and total duration of a given frequency component, but the Fourier transform does not indicate the instance of occurrence and duration of each spectral component.

2.3.4.1 Short-time Fourier transform

Temporal localisation can be obtained by performing the Fourier transform over shorter temporal windows of the signal. The characteristics of the short-time Fourier transform depend on the window function and its duration. Short windows provide good temporal resolution (the time instance when a frequency change occurs is clear), but poor frequency resolution (the exact frequency values present are unclear). Longer windows have the opposite characteristics (the frequency values present are clear, but the time instances of change are unclear) [62].

2.3.4.2 Wavelet transform

The wavelet transform addresses the resolution challenges of the short-time Fourier transform using multi-resolution analysis [63] where different resolutions are used for different spectral bands by changing the window duration according to frequency. Good temporal (poor frequency) resolution is used for high frequencies, and good frequency (poor temporal) resolution for low frequencies. This approach makes sense for most signals, since high and low frequency components require short and long durations for analysis respectively. The wavelet transform is performed by calculating a transform using a wavelet basis function instead of a sine wave as used in the Fourier transform.

Continuous wavelet transform

The continuous wavelet transform of the signal $x(t)$ can be expressed as [62]

$$X_w(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-\tau}{s} \right) \, dt \tag{2.55}$$

with $\psi(t)$ the mother wavelet, translated in time using τ, and scaled using s. Large values of s (low frequencies) correspond to a global view (large segments of the signal) and small values of s (high frequencies) correspond to a detailed view (small segments of the signal). The continuous wavelet transform provides a measure of similarity between $x(t)$ and wavelets derived from the mother wavelet. Several mother wavelets with unique properties are available, including the Haar, Mexican hat, Morlet and Daubechies wavelets [62]. The choice of mother wavelet depends on the signal type and application [64].

Discrete wavelet transform

An alternative method to obtain the wavelet representation of a signal is the discrete wavelet transform. The signal $x(t)$ is passed through a series of high and low pass filters to analyse the signal at high and low frequencies respectively. The scale and frequency resolution are changed by up or down sampling of the filtered signal, resulting in a sequence of discrete wavelet coefficients [62].

2.3.4.3 Wavelet denoising

Denoising can be performed by computing the discrete wavelet transform of a signal and setting all coefficient values less than a certain threshold to zero. The original denoised signal can then be obtained by calculating the inverse discrete wavelet transform [65]. Instead of using a threshold to decide which values should be set to zero, the positions of the coefficients can also be used.

2.3.4.4 DSSS detection

Examples of time-frequency techniques, suggested in the literature, to detect DSSS signals are considered below.

- A PN sequence can be detected within AWGN by analysing the short-time Fourier transform of the intercepted signal for periodic components [66].
- Similar to performing the compounded autocorrelation technique in the time domain as described in Section 2.3.2.3, the mean-square correlation sequence can also be calculated in the wavelet domain. By transforming a spreading code to the wavelet domain and then calculating the correlation between wavelet domain segments, DSSS transmissions can be detected [52].
- The performance of a radiometric DSSS intercept receiver in the presence of narrow band interference can be improved by removing the interference using wavelet denoising [67].

2.3.4.5 Critical evaluation of time-frequency analysis

The time-frequency techniques presented here are not very processor intensive, since they can be implemented using fairly simple filtering and correlation processes. It is however doubtful that time-frequency techniques will outperform the other techniques in low-SNR conditions. Simply combining temporal and frequency representations (which inherently carry the same signal information) will probably not provide significant gains in terms of detection performance, especially when weak DSSS signals are considered.

2.3.5 Principal component analysis

Blind detection techniques based on PCA can be used to detect communication signal activity with improved performance compared with other techniques, such as cyclic feature analysis and ED [68, 69].

PCA is a technique where the essential or principal components of a correlated data set are isolated to form a reduced and uncorrelated data set, with the aim of retaining most of the information [70]. PCA has found application in several fields including pattern recognition, image compression [71], and more recently in spectrum sensing used especially in cognitive radio [72].

2.3.5.1 Isolating the principal components

The first step in finding the principal components of a real data set, consisting of N variables (with M observations each) is to remove the mean from each variable [70]. The M observations of each zero-mean variable are organised as column vectors to form the $M \times N$ data matrix $\mathbf{X}$. The $N \times N$ sample covariance matrix (SCM) of $\mathbf{X}$ is then [73]

$$\mathbf{C} = \frac{1}{M}\mathbf{X}^T\mathbf{X} \tag{2.56}$$

with $\mathbf{X}^T$ the transpose of $\mathbf{X}$. The N eigenvalues and associated eigenvectors of $\mathbf{C}$ are then calculated and sorted in decreasing order such that $\lambda_1 > \lambda_2 > \ldots > \lambda_N$. The eigenvectors associated with the largest eigenvalues of $\mathbf{C}$ are the principal components of $\mathbf{X}$. The number of eigenvectors used depends on the application and level of reduction required.

2.3.5.2 Spectrum sensing application

The eigenvalues of $\mathbf{C}$ in (2.56) can be used to perform detection if $\mathbf{X}$ contains the samples of the received signal. A number of cognitive radio spectrum sensing algorithms, that use these eigenvalues in their detection test statistics are listed below [68, 72].

- The largest-eigenvalue (LE) method uses the largest eigenvalue directly with [74]

$$T_{\text{LE}} = \lambda_1(\mathbf{C}). \tag{2.57}$$

- The maximum-minimum-eigenvalue (MME) method uses the ratio between the largest and smallest eigenvalue with [69]

$$T_{\text{MME}} = \frac{\lambda_1(\mathbf{C})}{\lambda_N(\mathbf{C})}. \tag{2.58}$$

- The energy-with-minimum-eigenvalue (EME) method uses the ratio between the received signal power P_s and smallest eigenvalue with [69]

$$T_{\text{EME}} = \frac{P_s}{\lambda_N(\mathbf{C})}. \tag{2.59}$$

- The scaled-largest-eigenvalue (SLE) method normalises the largest eigenvalue with the mean of all the eigenvalues with [68]

$$T_{\text{SLE}} = \frac{\lambda_1(\mathbf{C})}{\frac{1}{N}\sum_{n=1}^{N}\lambda_n(\mathbf{C})} = \frac{N\lambda_1(\mathbf{C})}{\text{tr}(\mathbf{C})} \tag{2.60}$$

  since the trace $\text{tr}(\cdot)$ of a matrix (sum of its diagonal elements) equals the sum of its eigenvalues [75].

In addition to outperforming classical detection algorithms, the PCA algorithms associated with (2.57) to (2.60) do not need prior information of the signal to be detected. If accurate noise estimates are available, (2.57) will outperform ED and the other PCA algorithms. If accurate noise estimates are however unavailable, (2.58)-(2.60) can be used instead, as they are insensitive to noise estimation error [68].

2.3.5.3 DSSS detection

Similar to the detection of primary users in cognitive radio applications [74], detection of DSSS signals can be performed using the largest eigenvalue of the SCM of the data matrix $\mathbf{X}$ containing the intercepted signal [16]. The technique is semi-blind since it depends on knowledge of the spreading code length. Knowledge of the noise statistics is also required to determine the threshold [76]. More details of this technique are given in Chapter 3.

2.3.5.4 Critical evaluation of principal component analysis

The PCA techniques considered here are popular in especially cognitive radio spectrum sensing applications [68]. Although these techniques are processor intensive (they require the calculation of eigenvalues), they show promising performance for DSSS detection purposes. A number of publications with mathematical proof and simulated performance results indicate the superiority of PCA techniques to perform both DSSS detection and sequence estimation (compared with ED and autocorrelation techniques) [14, 16, 68, 77].

2.3.6 Chaos theory

Chaos theory involves nonlinear dynamical systems that exhibit apparent disordered behaviour [78, 79]. An important characteristic of chaotic systems is their sensitive dependence on initial conditions or system parameters [80]. This sensitivity is also referred to as the "butterfly effect" which expresses the idea that a seemingly insignificant event (the flap of a butterfly's wings) may possibly have dramatic consequences (setting off a distant tornado) [81].

Chaotic systems (e.g. Duffing oscillators [79, 82]) can potentially be made very sensitive to the presence of sinusoidal signals of a given frequency [83, 84]. The presence of a very weak signal (the butterfly's flap) will then have a dramatic effect (the tornado) on the behaviour of the chaotic oscillator, such that signal detection at low SNR is made possible.

2.3.6.1 DSSS detection

The Duffing oscillator may be used to detect the presence of a DSSS transmission, after performing a nonlinear operation (such as squaring) on the intercepted signal [85]. The nonlinear operation causes distinct spectral lines to reappear, that are not present in the intercepted digital transmission. These spectral lines indicate the presence of sinusoidal signals (which will be weak in the case of DSSS) which can be used to perform detection.

2.3.6.2 Critical evaluation of chaos theory

It was claimed that chaotic techniques can perform detection at much lower SNR levels compared with classical approaches [83], although no simulation (or measured) results to support this claim have yet been published. Also, several important details (e.g. the number of samples used and the false alarm rate) were not considered in [83].

Furthermore, two important issues still need to be resolved to realise detection using chaotic systems. Firstly, the sensitive-dependence principle not only implies the detector will be sensitive to the signal of interest, but also to many other factors (including noise). Methods to limit the false alarm rate should therefore be developed. Secondly, solving nonlinear dynamical systems involves highly complex and computationally expensive numerical methods, which may limit the practical application of these methods.

2.4 Approaches to estimate DSSS sequences

An intercepted communication transmission affected by noise and other unwanted channel effects can only be identified if its signal parameters can be estimated accurately. The parameters are also required to reconstruct the original signal, and to extract the information from the intercepted transmission.

To classify a communication transmission as DSSS with the aim of extracting the information (or to at least identify the transmitter system), the spreading code should be estimated. Important parameters mentioned in Section 2.3 that are needed to estimate the code sequence include the chip and symbol periods and the code length.[4] Many detection techniques also require these parameters as inputs.

Spreading code parameters can be estimated by exploiting the cyclostationary nature of DSSS transmissions. The chip and symbol periods (or rates) can be obtained using cyclic feature detection [3], HOS analysis [59], or correlation estimators [86]. The code length can be determined from the chip and symbol periods or directly from the spacing between correlation estimator peaks [55], or using a technique based on PCA [87].

Subsequently, techniques to estimate the spreading code itself are considered. Most of these techniques assume prior knowledge of the spreading code length.

2.4.1 Higher order statistical analysis

2.4.1.1 Triple correlation

The generator polynomials or LFSR structures of a number of overlapping m-sequences can be determined from the peaks on the two dimensional triple correlation function of the intercepted signal [59, 60].

2.4.1.2 Bispectral averaging

Bispectral techniques can be used to extract cyclostationary waveforms hidden in noise [88, 89], and can therefore possibly be applied to estimate spreading codes. These techniques calculate the average bispectrum from different segments of the noisy signal, and then extract the Fourier coefficients from the bispectrum. The original waveform can then be obtained using the inverse Fourier transform.

[4] The symbol period and code length are closely related, as the symbol period is equal to the period of a single spreading code, and the code length is the number of chips in a single spreading code.

2.4.2 Principal component concatenation

The spreading code of an intercepted DSSS transmission can be obtained from the data matrix, constructed using non-overlapping signal segments of duration equal to the DSSS symbol period. The first two principal components (i.e. the eigenvectors associated with the two largest eigenvalues; see Section 2.3.5) of the data matrix contain separate fragments of the spreading code [90]. The position and offset of the fragments within each eigenvector can be determined from the eigenvalues, and the spreading code can then be estimated by concatenating these eigenvectors [91].

2.4.3 Data matrix correlation

By constructing a data matrix as in the PCA technique discussed above, the spreading codes will typically not be aligned in the matrix, such that each row will contain fragments of two codes with each code having its own polarity. The matrix misalignment can be determined by cyclically shifting the matrix and calculating a measurement function for each cyclic shift [77]. The measurement function is a correlation between different matrix elements, which is used to determine the starting position of the spreading code within each row.

The data matrix can then be aligned such that each row will contain a complete spreading code. The original spreading code can be estimated from the aligned data matrix by calculating the correlation between the first and all other columns [77]. Assuming the first chip in the code is +1, the sign of each subsequent correlation value will indicate the value (± 1) of each remaining chip. The spreading code (or its reversed polarity) is thus obtained.

2.4.4 Artificial neural network approaches

Artificial neural networks (NNs) can be used to estimate spreading codes blindly if the spreading code length is known. Two such examples are considered below.

2.4.4.1 Single-layer networks

The weights of a single-layer unsupervised NN controlled by Oja's learning rule (or modified Hebbian rule) will converge to the principal component of the input data [92]. The single-layer NN can be implemented as a simple tapped delay line filter structure as shown in Fig. 2.12. The filter weight vector $\mathbf{w} = [w_1, w_2, \ldots, w_N]$ is updated according to [93]

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \Delta\ y\ [\mathbf{x}_n - y\ \mathbf{w}_n] \tag{2.61}$$

with Δ the step size, y the filter output, and $\mathbf{x}$ the input data vector.

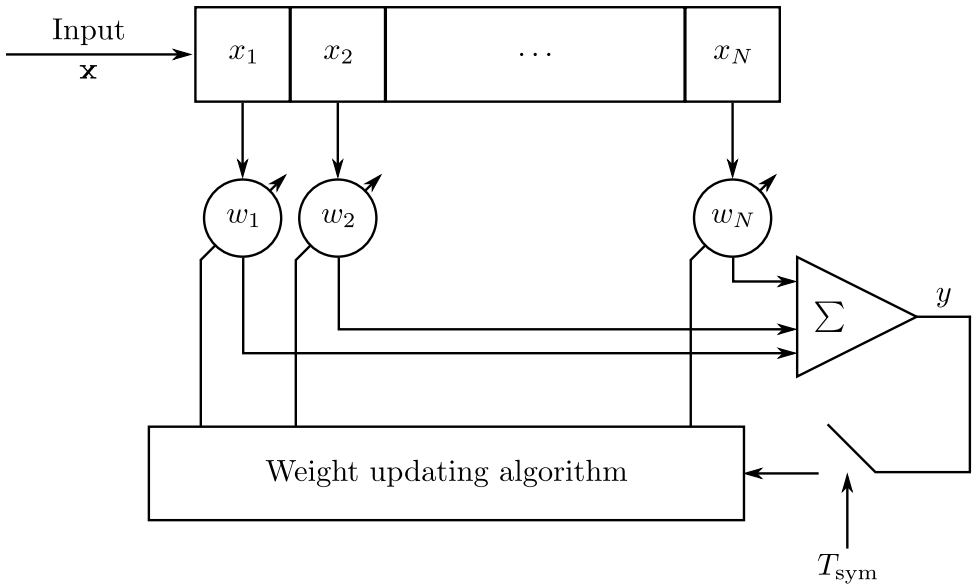


Figure 2.12: Tapped delay line filter structure forming a single-layer neural network.

Since the principal component of the aligned data matrix (see Sections 2.4.2 and 2.4.3) *is* the spreading code, sequence estimation can be performed using this tapped delay line filter structure. It is required that the filter (or NN) output be sampled at the DSSS symbol rate T_{sym} in order for the weights to be calculated according to synchronised spreading codes. Equivalently, the aligned data matrix can be fed into the NN row-by-row such that no weight updates are performed between complete DSSS symbols (or spreading codes).

2.4.4.2 Multi-layer networks

Using a three-layer NN with two neurons in the hidden layer, and feeding the network with the data matrix (not necessarily aligned but with the number of columns equal to the code length), the spreading code can be estimated as the converged weights of the hidden layer [94].

2.5 Gaps identified in the literature

Several areas of possible future investigation were identified in the literature considered in this chapter. For example, the HOS analysis and chaos theory techniques could be developed further as suggested in Sections 2.3.3.4 and 2.3.6.2. The detection and estimation performance of the techniques reviewed here could also be analysed and compared further. These aspects are however not pursued in this thesis.

The gaps in the literature that are focussed on in this thesis, are considered subsequently in this section. Although several aspects of the research problem were considered, which resulted in three journal publications, research on this topic could be continued as discussed in Section 6.5.

2.5.1 Simplifying mathematical performance expressions

The first publication (Chapter 3) considers the simplification of mathematical expressions which require computationally complex numerical evaluation. These expressions can be used in the analysis and application of detection receivers where decision thresholds in noisy environments are concerned. The simplified expressions do not need numerical evaluation and therefore provide a faster method of calculating threshold levels and performance predictions, at the cost of reduced numerical accuracy.

2.5.2 Development of improved DSSS detection techniques

The second publication (Chapter 4) presents improved DSSS detection techniques in terms of detection performance. A number of other aspects to improve current detection techniques are also considered, including the following.

2.5.2.1 Blind detection

The parameters of the DSSS signal to be detected are typically unknown to the intercept receiver. Blind (or semi-blind) detection techniques are therefore required. The methods presented in Chapter 4 are semi-blind, although they can be made fully-blind using the techniques presented in Chapter 5.

2.5.2.2 Computational complexity

Detection techniques that show promising detection performances in low-SNR conditions typically have high computational complexities. The number and type of arithmetic operations to perform detection, and possible techniques to reduce the computational complexity, are also considered in Chapter 4.

2.5.3 Sequence length estimation algorithms

Many of the existing DSSS detection and sequence-estimation algorithms assume knowledge of the sequence length. Very little is however published on sequence-length estimation, and the development of such techniques is therefore critically important. The third publication (Chapter 5) presents two possible sequence-length estimation techniques to address this gap.

2.5.4 Variety of spreading codes

Different spreading codes have different applications and unique properties [18, 28]. It can therefore be expected that a given detection or estimation technique will perform differently on different spreading codes. Therefore a need exists to develop different algorithms to cater for the variety of spreading codes. This fact is also highlighted in Chapter 5 by evaluating two types of codes.

CHAPTER 3

PUBLICATION 1

 The bibliographic details of the paper are given below and in the Reference list as [76].

Title	Analytic approximation to the largest eigenvalue distribution of a white Wishart matrix
Authors	J.D. Vlok and J.C. Olivier
Journal	IET Communications
Publication date	14 August 2012
Volume	6
Issue	12
Pages	1804–1811
doi	10.1049/iet-com.2011.0843
Print ISSN	1751-8628
Online ISSN	1751-8636

The submission timeline is given below.

Original manuscript submitted	11 November 2011
Decisioned: Revision requested	28 March 2012
Revised manuscript submitted	19 April 2012
Decisioned: Accepted	28 May 2012
Publication	14 August 2012

Analytic approximation to the largest eigenvalue distribution of a white Wishart matrix

J.D. Vlok[1] and *J.C. Olivier*[2]

[1]Defence, Peace, Safety & Security (DPSS), Council for Scientific and Industrial Research (CSIR), Pretoria 0001, South Africa

[2]School of Engineering, University of Tasmania, Hobart 7005, Australia

E-mail: jvlok@csir.co.za

Abstract: Eigenvalue distributions of Wishart matrices are given in the literature as functions or distributions defined in terms of matrix arguments requiring numerical evaluation. As a result the relationship between parameter values and statistics is not available analytically and the complexity of the numerical evaluation involved may limit the implementation, evaluation and use of eigenvalue techniques using Wishart matrices. This paper presents analytic expressions that approximate the distribution of the largest eigenvalue of white Wishart matrices and the corresponding sample covariance matrices. It is shown that the desired expression follows from an approximation to the Tracy-Widom distribution in terms of the Gamma distribution. The approximation offers largely simplified computation and provides statistics such as the mean value and region of support of the largest eigenvalue distribution. Numeric results from the literature are compared with the approximation and Monte Carlo simulation results are presented to illustrate the accuracy of the proposed analytic approximation.

3.1 Introduction

The eigenvalue spectrum of noise covariance matrices plays an important role in such fields as PCA [73], singular value decomposition (SVD), multiple-input multiple-output (MIMO) communication systems [95] and signal detection [96, 97]. The behaviour of the largest eigenvalue can be used to predict the performance of MIMO systems in a fading channel and the performance of eigenvalue-based signal detection techniques. The exact distributions of individual eigenvalues can be obtained from the joint distribution, which is defined in terms of hypergeometric functions if the covariance matrix has a Wishart distribution [95, 98]. The individual distributions are then expressed in terms of Laguerre polynomials [99] which can be simplified as matrix arguments [95, 100]. These however require numerical evaluation which can be performed using extensive tables or special purpose software [101]. However, it was shown in [102] that the asymptotic distribution of the scaled largest eigenvalue of a white Wishart matrix can be described by the Tracy-Widom

(TW) law [103, 104] which can be evaluated numerically [105–107] or approximated using a logit transform [101]. The TW distribution was also shown to be reasonably accurate for non-asymptotic cases [101, 102]. This paper presents a closed-form analytical expression to approximate the TW distribution in order to derive simple expressions for the largest eigenvalue distribution of the Wishart distributed covariance matrix and the associated sample covariance matrix, similar to an approximation given in [68]. Simple expressions describing the statistics and region of support of the largest eigenvalue distribution are also given. The rest of the paper is organised as follows. In Section 3.2 a mathematical background is given. Section 3.3 presents the approximation and Section 3.4 the expression for the largest eigenvalue distribution. Section 3.5 provides a simulation study where numeric results from the literature are compared with Monte Carlo simulation results and finally Section 3.6 summarises the main results and concludes the paper. The focus of this paper is on the TW law of order 1 and 2, denoted respectively by TW_1 and TW_2. TW_4 is briefly considered in the appendix.

3.2 Mathematical background

3.2.1 Noise matrix

Let $\mathbf{X}$ be an $M \times N$ matrix where each row of $\mathbf{X}$ is real and independently drawn from $\mathcal{N}_N(0, \sigma_x^2\mathbf{I})$, the N-variate normal distribution with zero mean and covariance matrix $\sigma_x^2\mathbf{I}$. The $N \times N$ matrix

$$\mathbf{Y} = \mathbf{X}^H\mathbf{X} \tag{3.1}$$

will then have a white Wishart distribution $\mathcal{W}_N(M, \sigma_x^2\mathbf{I})$ [102], where $\mathbf{X}^H$ is the Hermitian transpose of $\mathbf{X}$. If $\mathbf{X}$ is complex and the complex components of each row are independently drawn from $\mathcal{N}_N(0, (\sigma_x^2/2)\mathbf{I})$, $\mathbf{Y}$ will have a complex white Wishart distribution. The largest eigenvalue λ_1 of $\mathbf{Y}$ in the edge scaling limit, that is when $M \to \infty$, $N \to \infty$ and $\frac{M}{N} \to \gamma \geq 1$, will obey [102]

$$\frac{(\lambda_1/\sigma_x^2) - \mu_{MN,\beta}}{\sigma_{MN,\beta}} \xrightarrow{\mathcal{D}} F_\beta \tag{3.2}$$

where F_β is the TW cumulative distribution function (CDF) CDF with $\beta = 1$ if $\mathbf{X}$ is real and $\beta = 2$ if $\mathbf{X}$ is complex. The centre and scaling parameters for $\beta = 1$ are given by [102]

$$\mu_{MN,1} = \left(\sqrt{M-1} + \sqrt{N}\right)^2 \tag{3.3}$$

$$\sigma_{MN,1} = \sqrt{\mu_{MN,1}}\left(\frac{1}{\sqrt{M-1}} + \frac{1}{\sqrt{N}}\right)^{\frac{1}{3}} \tag{3.4}$$

and similarly for $\beta = 2$ are given by [102]

$$\mu_{MN,2} = \left(\sqrt{M} + \sqrt{N}\right)^2 \tag{3.5}$$

$$\sigma_{MN,2} = \sqrt{\mu_{MN,2}} \left(\frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}}\right)^{\frac{1}{3}}. \tag{3.6}$$

According to the limit $\frac{M}{N} \to \gamma \geq 1$, (3.2) to (3.6) hold only for $M \geq N$. It is however stated in [102] that (3.2) applies equally well if $M < N$ when $M \to \infty$, $N \to \infty$, and the roles of M and N are reversed in (3.3) and (3.4). Following the same argument, (3.5) and (3.6) can also be used for $M < N$ since reversing the roles of M and N has no effect in this case. Although (3.2) is true in the limit, [102] showed that it can provide a satisfactory approximation for matrix dimensions M and N as small as 10.

Note that (3.2) is usually stated for the unit variance case $\sigma_x^2 = 1$ (as in [73, 102]). The normalisation of λ_1 is required to develop expressions for the largest eigenvalue distribution and the associated statistics for the general case of σ_x^2. To explain the normalisation and show how a given eigenvalue λ of $\mathbf{Y}$ scale in comparison with the unit variance case, suppose $\mathbf{X}_{(u)}$ represents $\mathbf{X}$ when $\sigma_x^2 = 1$ and $\mathbf{Y}_{(u)} = \mathbf{X}_{(u)}^H \mathbf{X}_{(u)}$ from (3.1). The corresponding eigenvalue of $\mathbf{Y}_{(u)}$ is $\lambda_{(u)}$. By substituting $\mathbf{X} = \sigma_x \mathbf{X}_{(u)}$ for the general case into (3.1), it follows that $\mathbf{Y} = \sigma_x^2 \mathbf{Y}_{(u)}$. From the definition of eigenvalues and eigenvectors ($\mathbf{Yv} = \lambda \mathbf{v}$ with $\mathbf{v}$ an eigenvector of $\mathbf{Y}$) it can then be shown that

$$\lambda = \sigma_x^2 \lambda_{(u)}. \tag{3.7}$$

The eigenvalues of $\mathbf{Y}$ therefore scale with σ_x^2 compared with the unit variance case $\mathbf{Y}_{(u)}$. The eigenvalue λ_1 can therefore be normalised by dividing it with σ_x^2 as is done in (3.2).

3.2.2 Sample covariance matrix and relation to noise matrix

The sample covariance matrix of $\mathbf{X}$ is given as [73]

$$\mathbf{R} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}_m^H \mathbf{x}_m \tag{3.8}$$

with $\mathbf{x}_m$ the m^{th} $1 \times N$ row of $\mathbf{X}$. The ij^{th} element of the $N \times N$ matrix formed by $\mathbf{x}_m^H \mathbf{x}_m$ in (3.8) can by definition be expressed in terms of the elements of $\mathbf{X}$ as

$$\left[\mathbf{x}_m^H \mathbf{x}_m\right]_{ij} = \mathbf{X}_{mi}^* \mathbf{X}_{mj}. \tag{3.9}$$

From (3.8) and (3.9), each element of $\mathbf{R}$ can be expressed as

$$\mathbf{R}_{ij} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{X}^*_{mi}\mathbf{X}_{mj}. \tag{3.10}$$

Likewise, each element of $\mathbf{Y}$ can be expressed from (3.1) as

$$\mathbf{Y}_{ij} = \sum_{m=1}^{M} \mathbf{X}^*_{mi}\mathbf{X}_{mj} \tag{3.11}$$

which is the scalar product of the i^{th} row of $\mathbf{X}^H$ and the j^{th} column of $\mathbf{X}$. From (3.10) and (3.11) it is clear that

$$\mathbf{R} = \frac{1}{M}\mathbf{Y}. \tag{3.12}$$

The largest eigenvalue of $\mathbf{R}$ denoted by l_1 is therefore related to λ_1 as

$$l_1 = \frac{1}{M}\lambda_1. \tag{3.13}$$

Note that both λ_1 and l_1 are always real and non-negative since $\mathbf{Y}$ and $\mathbf{R}$ are always Hermitian (or symmetric if $\beta = 1$) and positive semidefinite.

3.2.3 Tracy-Widom law

The Tracy-Widom law [104] or distribution TW_β refers to a family of CDFs F_β and related PDFs f_β describing the limiting distributions of the largest eigenvalues of symmetric ($\beta = 1$), Hermitian ($\beta = 2$) or self-dual ($\beta = 4$) random matrices in the Gaussian ensembles[1].

The three TW CDFs are defined as [104, 105]

$$F_1(x) = \exp\left(-\frac{1}{2}\int_x^\infty q(w)\ dw\right)\sqrt{F_2(x)} \tag{3.14}$$

$$F_2(x) = \exp\left(-\int_x^\infty (w-x)\ q^2(w)\ dw\right) \tag{3.15}$$

$$F_4\left(\frac{x}{\sqrt{2}}\right) = \cosh\left(-\frac{1}{2}\int_x^\infty q(w)\ dw\right)\sqrt{F_2(x)} \tag{3.16}$$

with $q(w)$ the solution to the Painlevé II differential equation

$$q''(w) = wq(w) + 2q^3(w) \tag{3.17}$$

[1]The Gaussian ensembles include the Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and Gaussian symplectic ensemble (GSE) corresponding respectively to real ($\beta = 1$), complex ($\beta = 2$) and quaternion ($\beta = 4$) random matrices [98].

with the boundary condition $q(w) \sim \mathrm{Ai}(w)$ as $w \to \infty$ where $\mathrm{Ai}(w)$ is the Airy function. Calculation of F_β therefore requires evaluation of the Painlevé II differential equation which can be performed numerically and tabulated (see [107] for a review on the numerical evaluation of distributions defined in terms of Painlevé transcendents). A number of authors [105–107] developed and made available software modules to calculate double precision solutions of TW_β. Tables containing solutions of Painlevé II, TW_1 and TW_2 over $x \in [-40, 200]$ with step size $\Delta x = 0.0625$ as described in [106] are available at [108]. The numeric solutions of f_β obtained from [108] for $\beta = 1$ and 2 are shown in Fig. 3.1 and are used in this paper to develop the approximation.

3.3 Tracy-Widom approximation

In this section an approximation to TW_β using the Gamma distribution is proposed and the goodness-of-fit of the approximation is evaluated against the double precision numeric values of [108], which are exact to sixteen significant decimal digits. The numeric values of the PDF and CDF of TW_β obtained from [108] are denoted respectively by f_β and F_β. Likewise, the PDF and CDF of the Gamma approximation are denoted by g_β and G_β. Whereas only $\beta = 1$ and 2 are considered in this section, $\beta = 4$ is considered in the appendix.

3.3.1 Proposed Gamma approximation

By observing the numeric solutions of f_β in Fig. 3.1, the functions appear to resemble slightly asymmetric Gaussian density functions shifted on the x-axis. To incorporate the asymmetry, f_β could therefore be approximated using the Gamma PDF given by

$$g_\beta(x) = \frac{(x - x_0)^{k-1}}{\theta^k \Gamma(k)} \exp\left[\frac{-(x - x_0)}{\theta}\right] \tag{3.18}$$

with x_0 the location or shift parameter, k the shape, θ the scale and $\Gamma(k)$ the Gamma function. Values for these parameters (which are given in Table 3.1) were obtained by fitting g_β to the numeric values of f_β and minimising the sum of squared difference (SSD)

$$\epsilon_\beta^2 = \sum_{i=1}^{L} \left[f_\beta(x_i) - g_\beta(x_i)\right]^2 \tag{3.19}$$

over the full range of x in [108] such that $x_1 = -40$ and $x_L = 200$ with sample step size $\Delta x = 0.0625$. The statistics of the resultant Gamma approximation are given in Table 3.1, which resemble the TW statistics given in Table 1 of [104, 107]. In addition to the numeric solutions, Fig. 3.1 also shows the Gamma approximations using (3.18) with the parameter values from Table 3.1. The SSD values obtained using (3.19) are also given in Table 3.1.

Table 3.1: Parameter values and related results for the Gamma approximation to TW_β.

Parameter	**Symbol**	$\beta = 1$	$\beta = 2$
Shape	k	46.5651	79.3694
Scale	θ	0.1850	0.1010
Location	x_0	-9.8209	-9.7874
Mean	$k\theta + x_0$	-1.2064	-1.7711
Variance	$k\theta^2$	1.5937	0.8096
Skewness	$2/\sqrt{k}$	0.2931	0.2245
SSD	ϵ_β^2	2.8270×10^{-5}	9.3883×10^{-6}
SCvM statistic	W_β^2	1.0547×10^{-7}	4.7651×10^{-8}
Kolmogorov statistic	K_β	8.0577×10^{-4}	4.0428×10^{-4}

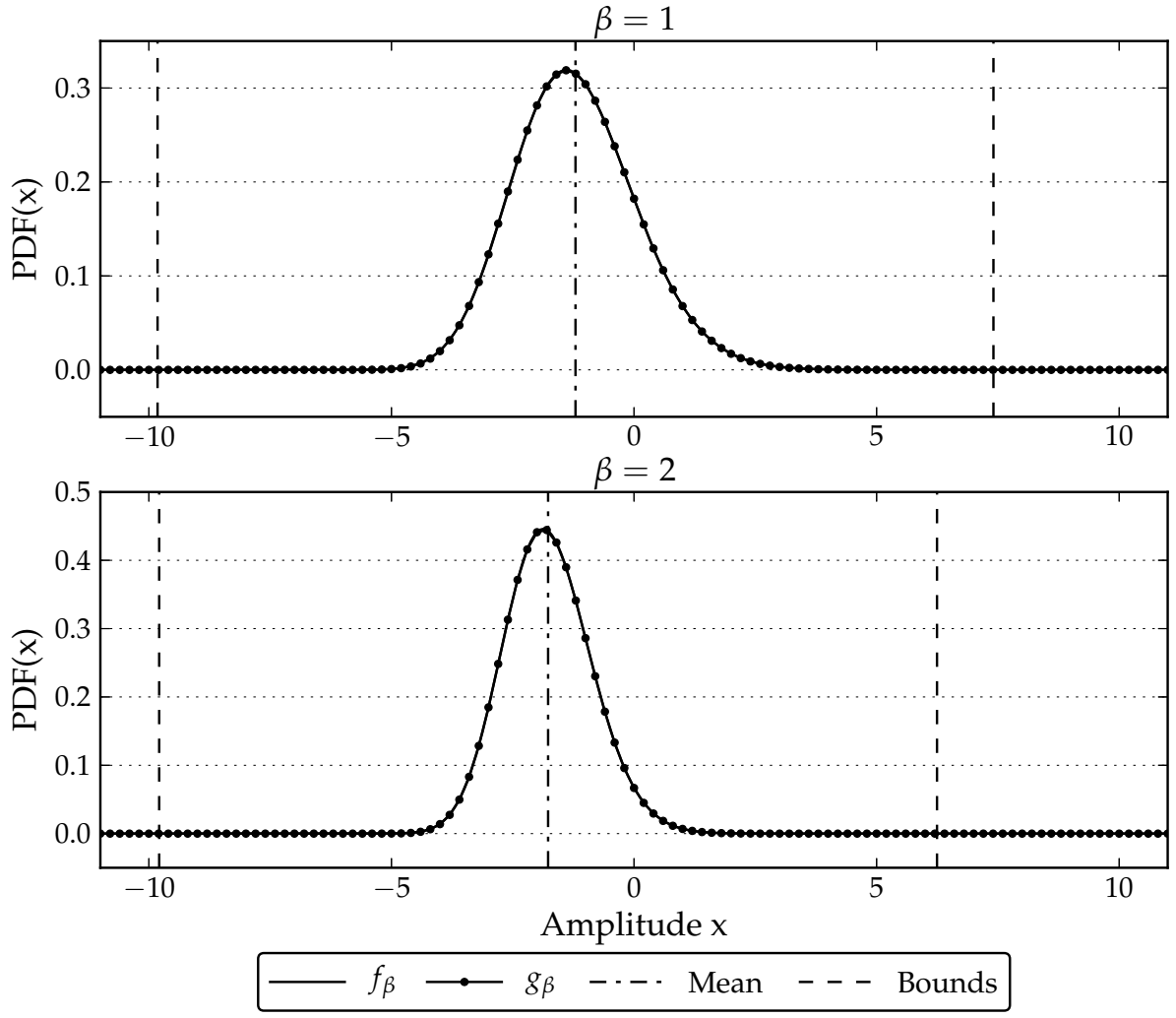


Figure 3.1: Numeric and approximated PDFs for $\beta = 1$ and $\beta = 2$. f_β is the numeric solution of the TW PDF obtained from [108] and g_β is the Gamma PDF given in (3.18).

3.3.2 Support region

Although the support of TW_β is not bounded, both left and right tails of f_β exhibit exponential decay [102]. It is therefore possible to truncate the support region to certain bounds $[b_-, b_+]$ without losing much probability mass. This section proposes a truncated support region for TW_β based on the Gamma approximation presented in Section 3.3.1. The probability mass lost when using the truncated support region is also considered.

The Gamma PDF given in (3.18) has support $[x_0, \infty)$ and the location parameter x_0 can therefore be used as the lower bound b_-. The upper bound b_+ is chosen such that the mean value of the Gamma distribution is also the mean of the lower and upper bounds[2]. The support of the truncated Gamma approximation is then

$$[b_-, b_+] = [x_0, 2k\theta + x_0] \tag{3.20}$$

which is also displayed in Fig. 3.1.

To illustrate the effect of the truncation, the PDFs and CDFs of TW_β and the associated Gamma approximations are shown with logarithmic ordinate axes in Figs. 3.2 and 3.3. To quantify the loss in probability mass due to the truncation, values from Fig. 3.3 for the mass of each tail in terms of the cumulative distribution outside the bounded region of (3.20) are given in Table 3.2. Interpolated values of [108] are used as reference solutions for F_β. The total probability mass lost in the truncation is the mass outside the support region.

Table 3.2: Parameter values related to truncated support and loss in probability mass.

Parameter	**Expression**	$\beta = 1$	$\beta = 2$
Lower bound	b_-	-9.8209	-9.7874
Upper bound	b_+	7.4082	6.2452
Left tail mass (reference)	$F_\beta(b_-)$	3.4799×10^{-21}	7.6093×10^{-35}
Right tail mass (reference)	$1 - F_\beta(b_+)$	4.3875×10^{-8}	1.0734×10^{-12}
Total mass lost (reference)	$F_\beta(b_-) + 1 - F_\beta(b_+)$	4.3875×10^{-8}	1.0734×10^{-12}
Left tail mass (approximation)	$G_\beta(b_-)$	0	0
Right tail mass (approximation)	$1 - G_\beta(b_+)$	3.4942×10^{-8}	1.1563×10^{-12}
Total mass lost (approximation)	$G_\beta(b_-) + 1 - G_\beta(b_+)$	3.4942×10^{-8}	1.1563×10^{-12}

[2]For the purpose of choosing bounds the TW PDFs can be assumed to be approximately symmetric - which can be seen clearly in Fig. 3.1.

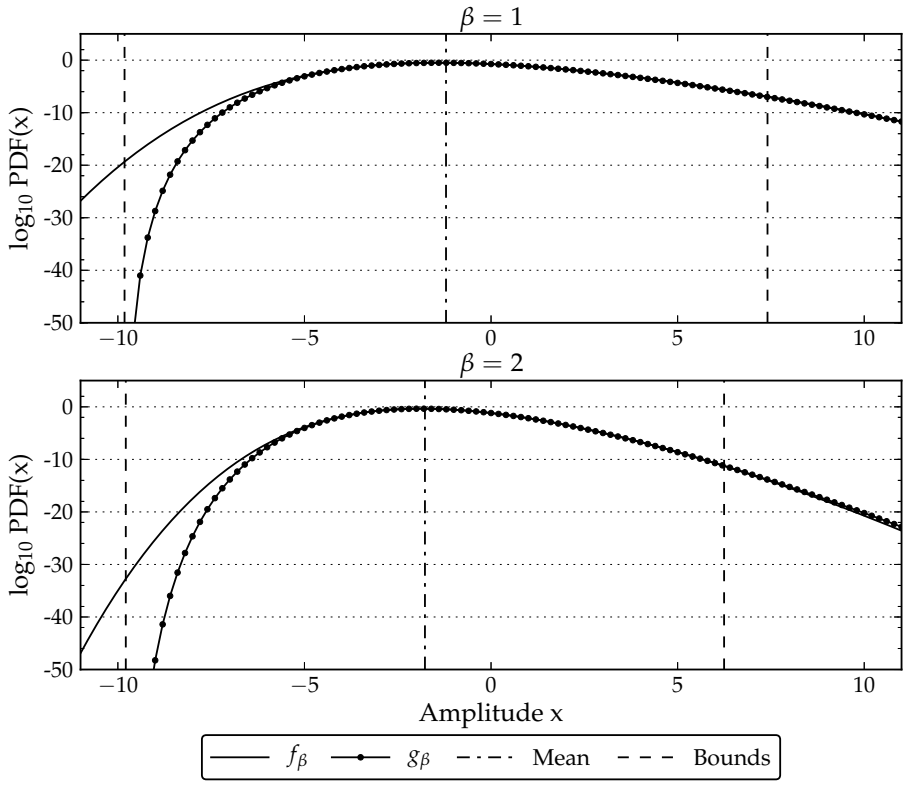


Figure 3.2: Numeric and approximated PDFs for $\beta = 1$ and $\beta = 2$ with logarithmic ordinate axes. f_β is the numeric solution of the TW PDF obtained from [108] and g_β is the Gamma PDF given in (3.18).

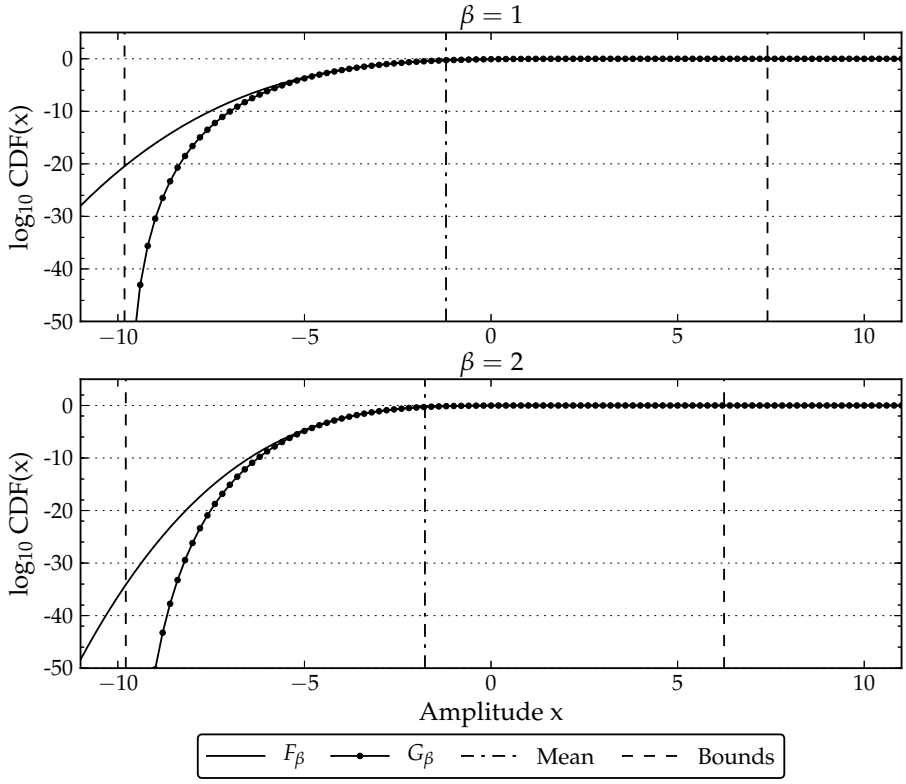


Figure 3.3: Numeric and approximated CDFs for $\beta = 1$ and $\beta = 2$. F_β is the numeric solution of the TW CDF obtained from [108] and G_β is the Gamma CDF derived from (3.18).

3.3.3 Goodness-of-fit

The approximation accuracy can also be measured using goodness-of-fit tests which indicate how close an empirical CDF is to a theoretical CDF. The difference or distance between the two CDFs is given for the purpose of this study as

$$D_\beta(x) = F_\beta(x) - G_\beta(x). \tag{3.21}$$

Two tests from [109] based on (3.21) are used in this paper to evaluate the approximation accuracy. The first test is the Smirnov-Cramér-Von-Mises (SCvM) test with test statistic

$$W_\beta^2 = \int_{b_-}^{b_+} D_\beta^2(x) g_\beta(x)\, dx. \tag{3.22}$$

The second test is the Kolmogorov test with test statistic

$$K_\beta = \max |D_\beta(x)|\,; \quad x \in [b_-, b_+]. \tag{3.23}$$

Both the SCvM and Kolmogorov test statistics are indications of how well the numeric values from [108] fit the analytic expression in (3.18). These statistics will approach zero as the goodness-of-fit improves. The values of (3.22) obtained through numerical integration with step size $\Delta x = 0.0625$ and (3.23) are given in Table 3.1. The values of the test statistics remain unchanged whether they are evaluated over $[-40, 200]$ or $[b_-, b_+]$ given in (3.20), confirming that the truncation has a negligible effect on the accuracy of the approximation. Graphs depicting the absolute value of (3.21) over x and the associated Kolmogorov statistics are shown in Fig. 3.4.

3.4 Expression for largest eigenvalue distribution

This section provides expressions for the largest eigenvalue distributions of the noise matrix $\mathbf{Y}$ and the sample covariance matrix $\mathbf{R}$ based on the TW approximation presented in Section 3.3. Other approximation methods are also briefly considered.

3.4.1 Noise matrix Y

Using (3.2) and (3.18) and linear random variable transformations [110] the PDF of λ_1 can be expressed as

$$p_{\lambda_1}(x) = \frac{1}{\sigma_x^2\ \sigma_{MN,\beta}} g_\beta \left\{ \frac{(x/\sigma_x^2) - \mu_{MN,\beta}}{\sigma_{MN,\beta}} \right\} \tag{3.24}$$

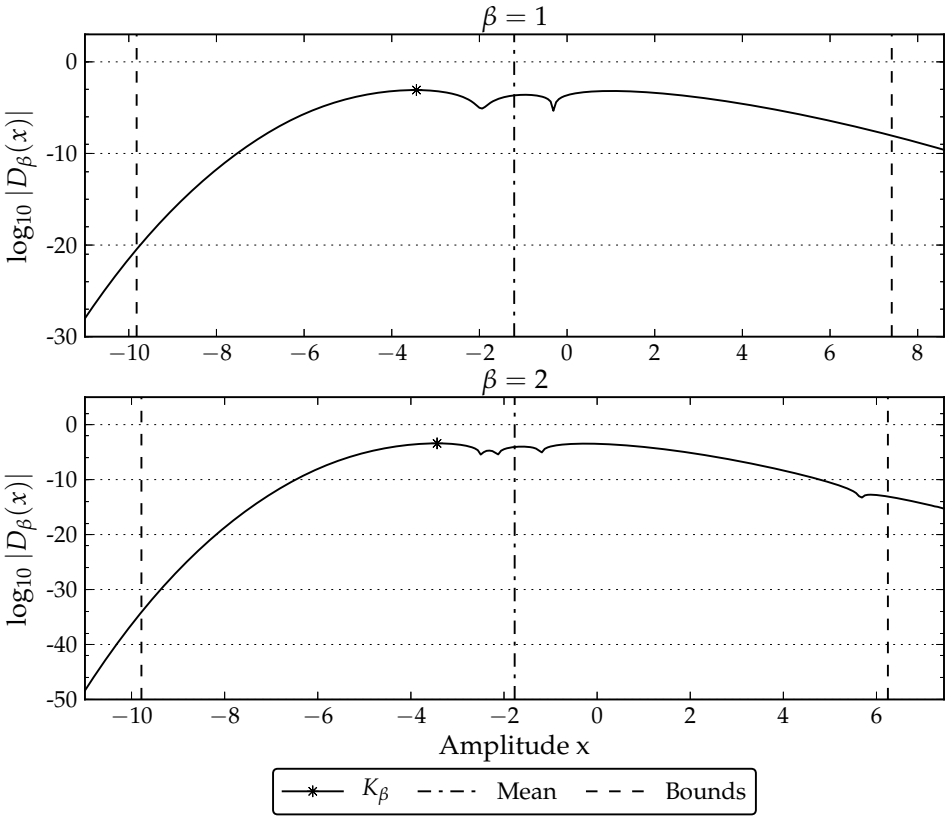


Figure 3.4: Absolute difference between the CDFs F_β and G_β as defined in (3.21) for $\beta = 1$ and $\beta = 2$. The Kolmogorov statistic defined in (3.23) is also shown on each graph.

which can be written in the form of (3.18) as

$$p_{\lambda_1}(x) = \frac{(x - x_0')^{k-1}}{\theta'^k \Gamma(k)} \exp\left[\frac{-(x - x_0')}{\theta'}\right] \tag{3.25}$$

with updated parameters

$$\theta' = \sigma_x^2 \, \sigma_{MN,\beta} \, \theta \tag{3.26}$$

$$x_0' = \sigma_x^2 \, (\mu_{MN,\beta} + x_0 \, \sigma_{MN,\beta}) \, . \tag{3.27}$$

The support of $p_{\lambda_1}(x)$ can then be written from (3.20) with the updated parameters given in (3.26) and (3.27) as

$$[b_-, b_+] = [x_0', 2k\theta' + x_0']. \tag{3.28}$$

3.4.2 Sample covariance matrix R

Using (3.13) the PDF of l_1 can be written from (3.24) as [110]

$$p_{l_1}(x) = M p_{\lambda_1}(Mx) \tag{3.29}$$

which can also be written in the form of (3.18) or (3.25) as

$$p_{l_1}(x) = \frac{(x - x_0'')^{k-1}}{\theta''^k \Gamma(k)} \exp\left[\frac{-(x - x_0'')}{\theta''}\right] \tag{3.30}$$

with parameters updated again, giving

$$\theta'' = \frac{\theta'}{M} \tag{3.31}$$

$$x_0'' = \frac{x_0'}{M}. \tag{3.32}$$

The support of $p_{l_1}(x)$ can then be written as (3.28) by replacing the updated parameters with the twice-updated parameters given in (3.31) and (3.32).

3.4.3 Other approximations

Other related approximations include the logit transform approximation to the TW law presented in [101] and a Gamma approximation describing the largest eigenvalue distribution in [68]. The logit transform approach considers only $\beta = 1$ and is computationally more complex than the approximation proposed in this paper. The approximation of [68] calculates the shape k and scale θ of the Gamma distribution by matching the first two moments of the largest eigenvalue and Gamma distributions using an equivalent of (3.2) and the TW law. The TW distribution is however not approximated directly and the shift parameter x_0 is not used. The focus of [68] is on spectrum sensing applicable to cognitive radio and only $\beta = 2$ is considered for matrix $\mathbf{Y}$. The approximation of [68] is however evaluated in the simulation study in Section 3.5 against the approximation presented in this paper for both $\beta = 1$ and 2 using the same scaling parameters given in Section 3.2.1 and the values of the first two TW moments given in [107].

3.5 Simulation study and results

A Monte Carlo computer simulation study was conducted with the aim of evaluating how accurate the proposed Gamma approximations can predict actual largest eigenvalue distributions. Empirical distributions of the largest eigenvalues of matrices $\mathbf{Y}$ and $\mathbf{R}$ for both $\beta = 1$ and 2 were obtained through simulation using 10^6 replications of these matrices for a given set of matrix dimensions (M, N) with $\sigma_x^2 = 1$. Every simulation set was started using identical random seed values. The empirical PDF for a given set was obtained from the simulated data by calculating the histogram over the support region given in (3.20) with the number of bins fixed to 100. To measure the approximation accuracy, the SCvM criterion given by (3.21) and (3.22) was used with $F_\beta(x)$ corresponding to the empirical CDF obtained through the Monte-Carlo simulations. Likewise, $G_\beta(x)$ and

$g_\beta(x)$ correspond to the Gamma approximations with densities defined by (3.25) or (3.30) depending on whether λ_1 or l_1 is concerned. Subsequently the SCvM results are presented. Section 3.5.1 considers an example set $(M, N) = (20, 40)$ and Section 3.5.2 a range of matrix dimensions. For the purpose of comparison, the SCvM statistics calculated for the approximation method given in [68] (see Section 3.4.3) are also given in Section 3.5.2. The results are discussed in Section 3.5.3.

3.5.1 Example set

Figs. 3.5 and 3.6 show the predicted and simulated distributions of the largest eigenvalues for $\beta = 1$ and $(M, N) = (20, 40)$. The predicted curves correspond to the Gamma approximations based on (3.18) and the simulated curves to the empirical data. Fig. 3.5 shows the results for λ_1 using (3.25) as prediction and Fig. 3.6 shows the results for l_1 using (3.30) as prediction. Table 3.3 shows parameter values for $\beta = 1$ (corresponding to Figs. 3.5 and 3.6) and $\beta = 2$. The goodness-of-fit statistics are identical for λ_1 and l_1 for each case of β since the random seed values used are identical and the number of histogram bins used in determining the empirical CDF is constant.

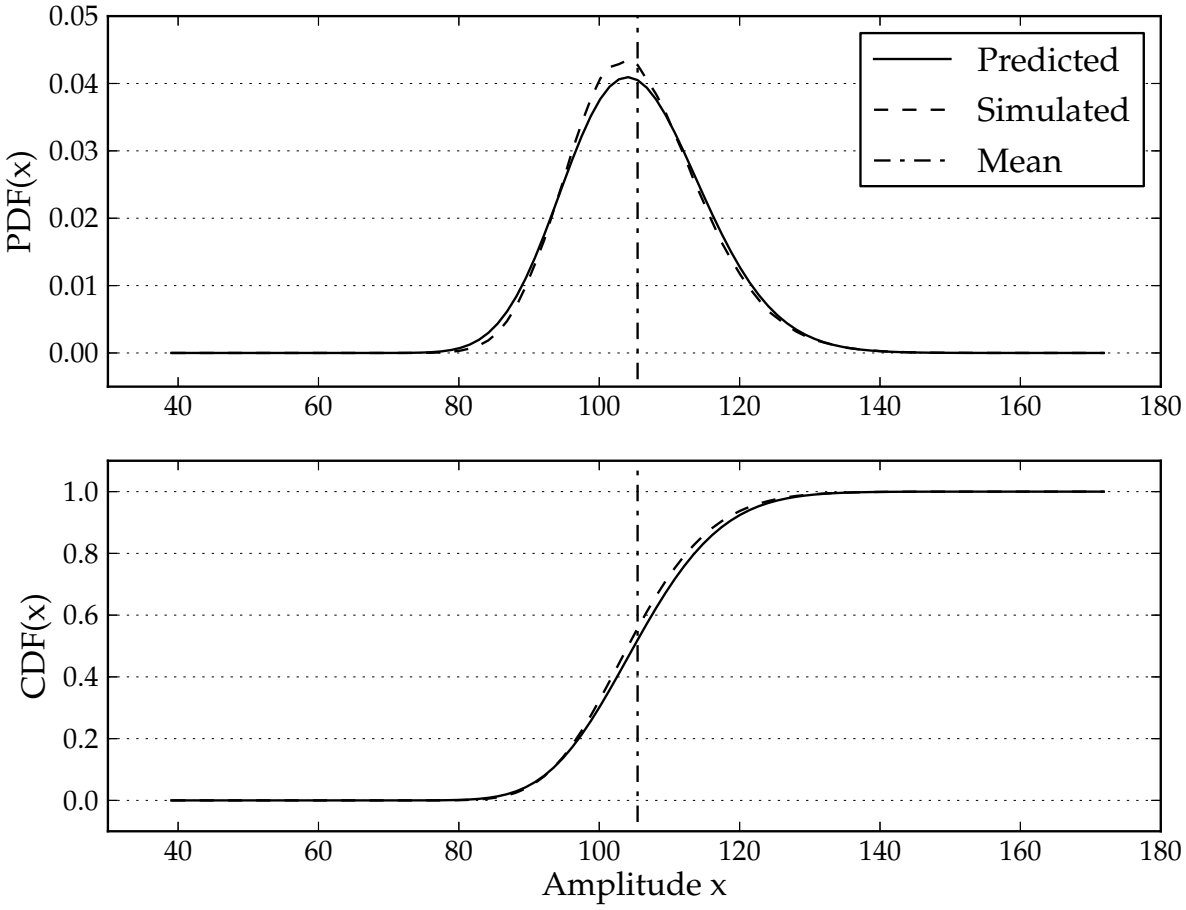


Figure 3.5: Predicted (from (3.25)) and simulated PDFs and CDFs of λ_1 for $\beta = 1, (M, N) = (20, 40)$ and $\sigma_x^2 = 1$.

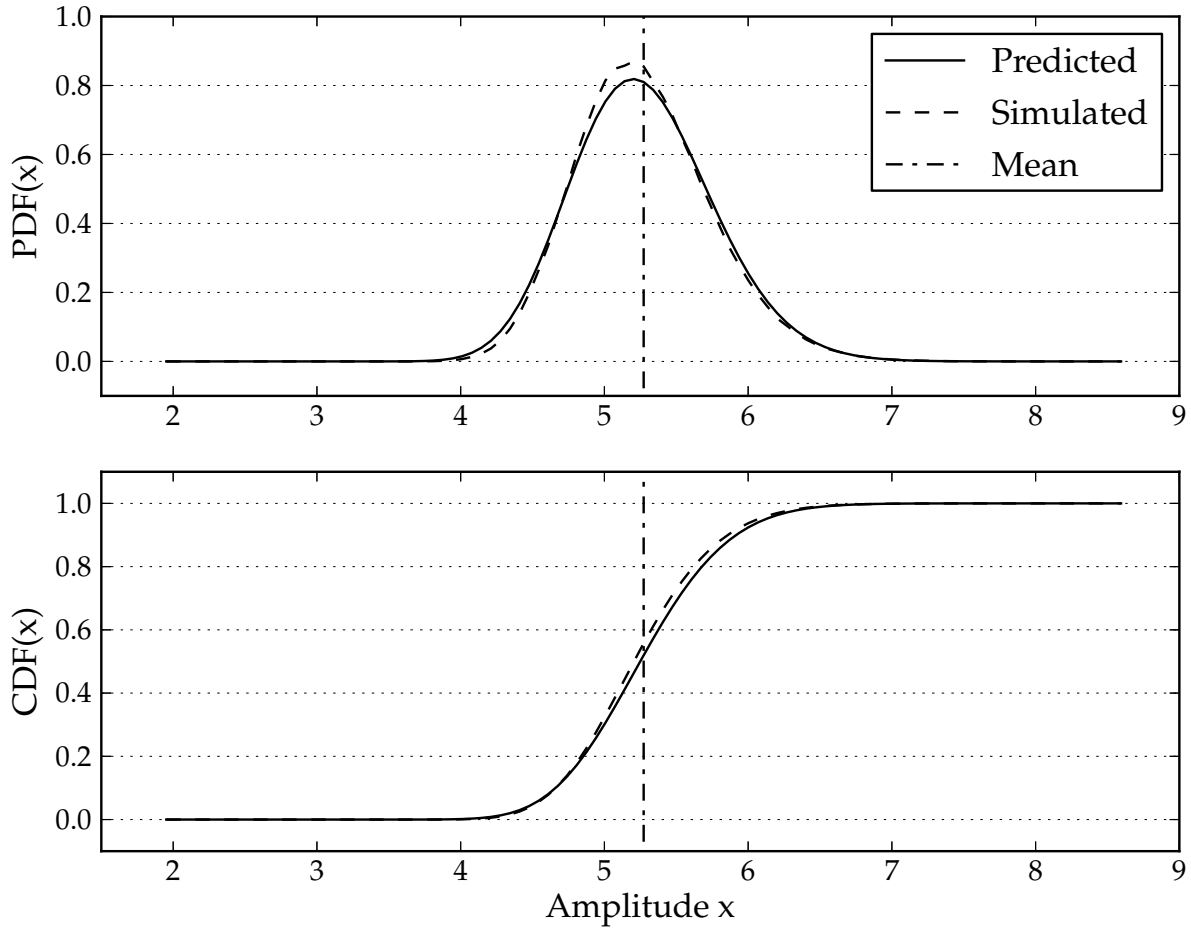


Figure 3.6: Predicted (from (3.30)) and simulated PDFs and CDFs of l_1 for $\beta = 1, (M, N) = (20, 40)$ and $\sigma_x^2 = 1$.

Table 3.3: Parameter values for the largest eigenvalue distributions for $(M, N) = (20, 40)$.

	$\beta = 1$		$\beta = 2$	
Parameter	λ_1	l_1	λ_1	l_1
Mean (Theory) $k\theta + x_0$	105.4619	5.2731	102.6974	5.1349
Mean (Measured)	105.4344	5.2717	103.1117	5.1556
Lower bound b_-	38.3724	1.9186	39.9136	1.9957
Upper bound b_+	172.5515	8.6276	165.4811	8.2741
Bin size Δx	1.3418	6.7090×10^{-2}	1.2557	6.2784×10^{-2}
SCvM statistic W_β^2	9.9414×10^{-5}	9.9414×10^{-5}	4.2237×10^{-4}	4.2237×10^{-4}
Kolmogorov statistic K_β	1.3901×10^{-2}	1.3901×10^{-2}	3.2812×10^{-2}	3.2812×10^{-2}

3.5.2 Range of matrix dimensions

The range of matrix dimensions from $(M, N) = (20, 20)$ to $(200, 200)$ for both cases of $M \geq N$ and $M < N$ is considered. Results for the square matrix case $M = N$ are given in Table 3.4 and plotted in Fig. 3.7. As in Table 3.3, the SCvM statistics for λ_1 and l_1 are identical in Table 3.4. Fig. 3.7 shows that as the matrix dimensions increase, the SCvM statistics decrease indicating an improvement in the approximation accuracy. For $\beta = 1$, the approximation given in (3.25) outperforms [68] up to a maximum SCvM difference of 9.2809×10^{-5} at $M = 200$. For $\beta = 2$ the two approximation methods show similar accuracies though for smaller values of M, [68] performs slightly better and for larger values of M, (3.25) performs slightly better.

Table 3.4: SCvM statistics for the largest eigenvalues when $M = N$.

	W_1^2 $(\beta = 1)$		W_2^2 $(\beta = 2)$	
M	λ_1 (3.25) and l_1 (3.30)	λ_1 [68]	λ_1 (3.25) and l_1 (3.30)	λ_1 [68]
20	2.8308×10^{-4}	2.8493×10^{-4}	7.4245×10^{-4}	6.9550×10^{-4}
40	9.5246×10^{-5}	1.0662×10^{-4}	2.7204×10^{-4}	2.0513×10^{-4}
60	5.2480×10^{-5}	8.7108×10^{-5}	1.6838×10^{-4}	1.1199×10^{-4}
80	3.1733×10^{-5}	7.8397×10^{-5}	1.0092×10^{-4}	6.7867×10^{-5}
100	2.5417×10^{-5}	8.2993×10^{-5}	7.9621×10^{-5}	5.4370×10^{-5}
120	1.8896×10^{-5}	7.9911×10^{-5}	5.9821×10^{-5}	4.4188×10^{-5}
140	1.1512×10^{-5}	9.0566×10^{-5}	3.7689×10^{-5}	4.2799×10^{-5}
160	1.2007×10^{-5}	9.3116×10^{-5}	3.6145×10^{-5}	4.3566×10^{-5}
180	8.3928×10^{-6}	9.2033×10^{-5}	2.7315×10^{-5}	4.1183×10^{-5}
200	6.8972×10^{-6}	9.9706×10^{-5}	2.7319×10^{-5}	4.0576×10^{-5}

SCvM results for fixed values of $M = 20$ and 200 (the extreme cases) over the range of $N \in [20, 200]$ when $\beta = 2$ are shown in Fig. 3.8. Again it is evident that larger matrix dimensions result in improved approximation accuracy. Fig. 3.8 also shows that the two methods (3.25) and [68] exhibit similar approximation accuracies, though [68] is slightly better for $M = 20$ and (3.25) is slightly better for $M = 200$.

3.5.3 Discussion of results

The presented results indicate that the Gamma approximation can provide an accurate prediction of the empiric distribution of the largest eigenvalue. It was also shown that the approximation accuracy improves as the matrix dimensions increase. This can be explained from (3.2) which is stated in terms of the edge scaling limits of the matrix dimensions. As the matrix dimensions increase, the TW law will provide a better prediction of the largest eigenvalue distribution. The approximation to the TW law will therefore

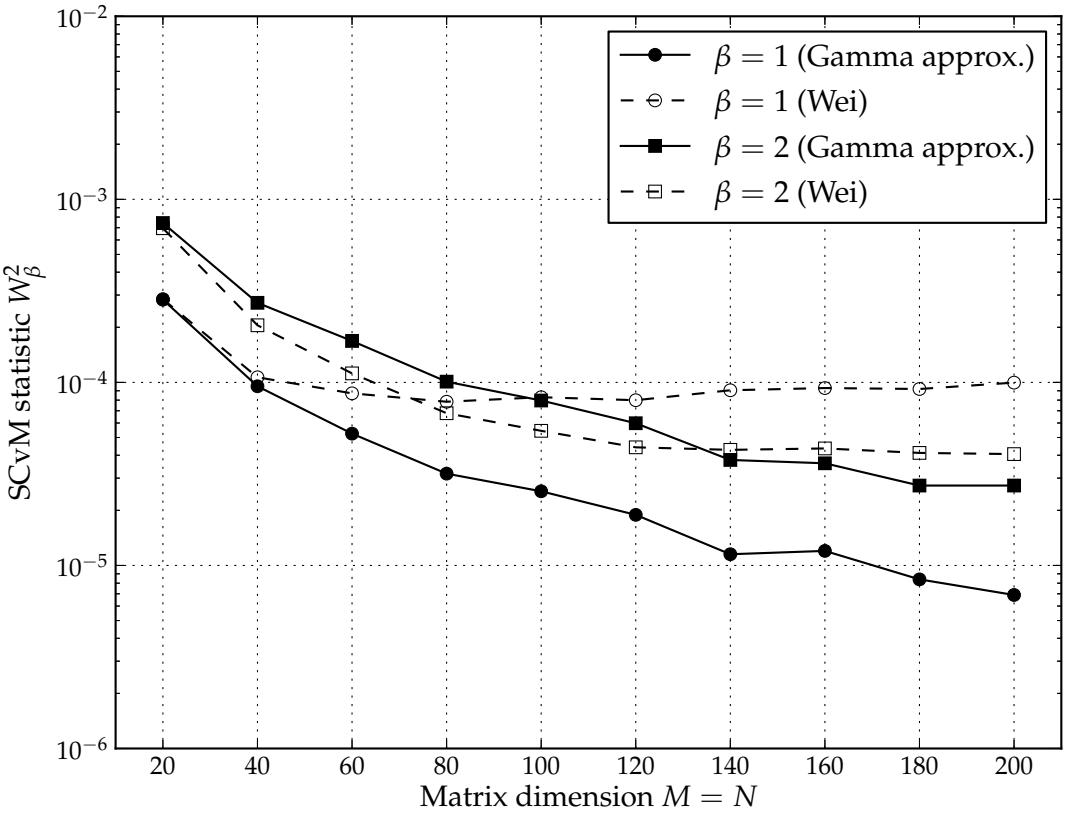


Figure 3.7: SCvM statistics for λ_1 as given in Table 3.4. The Gamma approximation and Wei's method correspond respectively to (3.25) and [68].

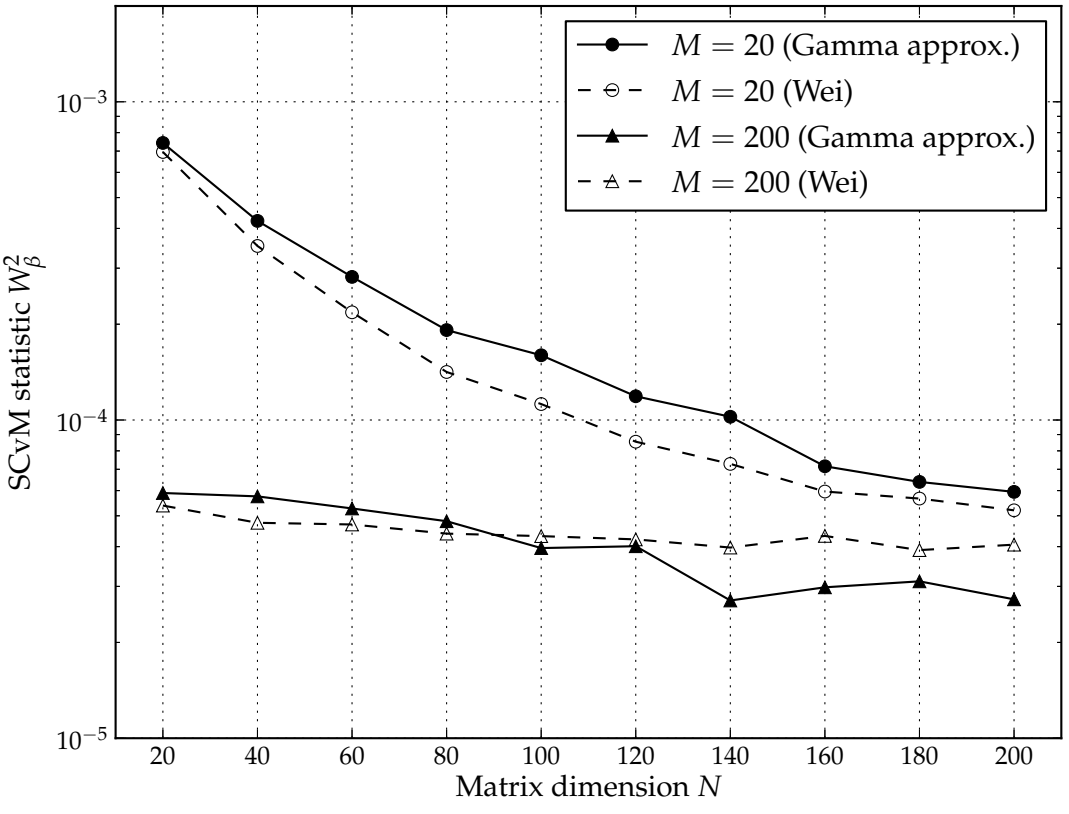


Figure 3.8: SCvM statistics for λ_1 ($\beta = 2$) with M fixed over the range of N. The Gamma approximation and Wei's method correspond respectively to (3.25) and [68].

also provide a more accurate prediction for larger matrix dimensions. Lastly, the approximation given by (3.25) is generally more accurate than [68] (especially for larger matrix dimensions). This can be ascribed to the different approximation methods. The method of [68] relies on matching moments (see Section 3.4.3) to find k and θ of the Gamma distribution. The method presented in this paper fits the Gamma distribution (k, θ and x_0) to the TW law directly and then uses (3.2) to derive the largest eigenvalue distribution. Using the shift parameter in the approximation provides a method to more accurately describe the TW law in terms of the Gamma distribution, which results in improved approximation accuracies.

3.6 Conclusion

This paper presented an approximation to the TW law based on the Gamma distribution which was shown, through Monte Carlo computer simulation and an analysis of the distributions, to be able to accurately predict the largest eigenvalue distribution of white Wishart matrices and their corresponding sample covariance matrices. The approximation provides a tractable and closed-form solution and does not require numerical evaluation. Furthermore, simple equations were derived to accurately predict the statistics and support region of the principal component of a noise matrix directly from the matrix dimensions. The results of this paper can be used to develop analytic expressions where the TW law forms part of the argument. Such expressions will be useful in the analysis and application of detection receivers where decision thresholds in noisy environments are concerned, e.g. in MIMO, cognitive radio and signal detection systems.

3.7 Acknowledgements

This work was supported by the Armaments Corporation of South Africa (Armscor) under contract KT521896. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

3.8 Appendix

This appendix considers two approximations to TW_4. The first approximation (referred to as the indirect Gamma approximation) is based on the Gamma approximations to TW_1 and TW_2 developed in this paper. The second approximation (direct Gamma approximation) is obtained using the method presented in Section 3.3.1.

3.8.1 Indirect Gamma approximation

The CDF F_4 given in (3.16) can be written in terms of F_1 and F_2 as

$$F_4\left(\frac{x}{\sqrt{2}}\right) = \cosh(\alpha(x))\sqrt{F_2(x)} \tag{3.33}$$

with

$$\alpha(x) = -\frac{1}{2}\int_x^\infty q(w)\ dw = \ln\left(\frac{F_1(x)}{\sqrt{F_2(x)}}\right) \tag{3.34}$$

from (3.14). The PDF f_4 can then be obtained by differentiation from (3.33) as

$$f_4\left(\frac{x}{\sqrt{2}}\right) = \frac{\sinh(\alpha(x))\sqrt{2F_2(x)}f_1(x)}{F_1(x)} + \frac{\exp(-\alpha(x))f_2(x)}{\sqrt{2F_2(x)}}. \tag{3.35}$$

It is required in (3.34) and (3.35) that $F_1(x) > 0$ and $F_2(x) > 0$. $F_4(x) = 0$ and $f_4(x) = 0$ wherever $F_1(x) = 0$ or $F_2(x) = 0$. By substituting the Gamma approximations g_β and G_β ($\beta = 1$ and 2) developed in Section 3.3 into f_β and F_β in (3.33) to (3.35), the indirect Gamma approximation $\tilde{\Gamma}_4$ is obtained. Using the double precision values obtained from [107] as reference (over $x \in [-10, 10]$ and $\Delta x = 0.0625$), the goodness-of-fit statistics (see Section 3.3.3) are calculated as $W_4^2 = 1.1455 \times 10^{-5}$ and $K_4 = 5.4584 \times 10^{-3}$.

3.8.2 Direct Gamma approximation

The direct Gamma approximation Γ_4 was obtained using the method described in Section 3.3.1 and the numeric values from [107]. The resultant parameter values are given in Table 3.5.

The PDFs and CDFs of TW_4 from [107], Γ_4 and $\tilde{\Gamma}_4$ are displayed in Fig. 3.9. It is evident from Fig. 3.9 and the goodness-of-fit statistics given in the previous section and Table 3.5 that $\tilde{\Gamma}_4$ is a less accurate approximation than Γ_4. This can be expected since the approximation $\tilde{\Gamma}_4$ is based on approximations to TW_1 and TW_2.

Table 3.5: Parameter values and related results for the direct Gamma approximation to TW_4.

Parameter	**Symbol**	$\beta = 4$
Shape	k	105.7442
Scale	θ	0.0700
Location	x_0	-9.7038
Mean	$k\theta + x_0$	-2.3017
Variance	$k\theta^2$	0.5181
Skewness	$2/\sqrt{k}$	0.1945
SSD	ϵ_β^2	1.0623×10^{-4}
SCvM statistic	W_β^2	1.9356×10^{-6}
Kolmogorov statistic	K_β	1.8025×10^{-3}

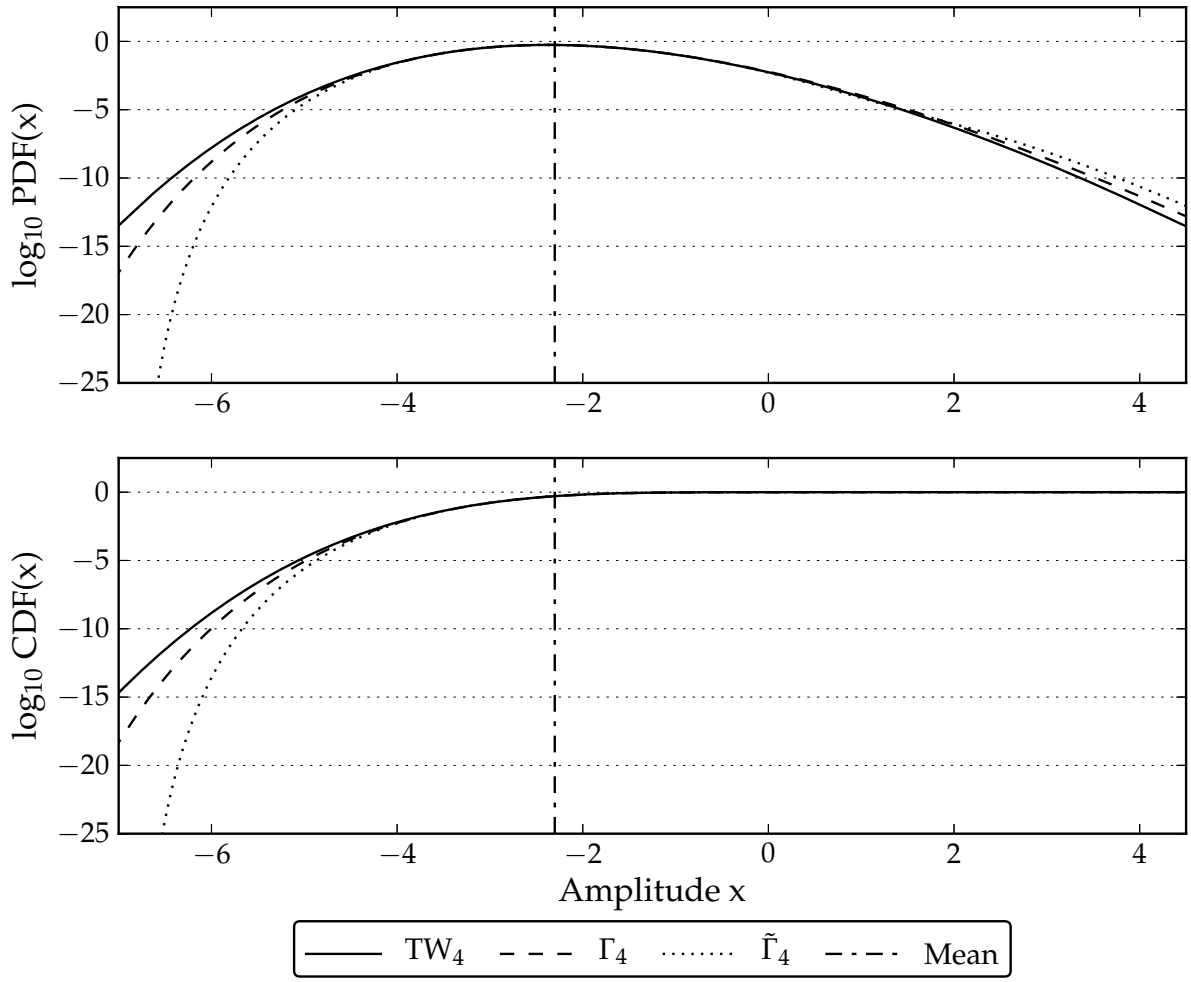


Figure 3.9: Numeric and approximated PDFs and CDFs for $\beta = 4$. TW_4 refers to the numeric solution obtained from [107]. Γ_4 and $\tilde{\Gamma}_4$ refer respectively to the direct and indirect Gamma approximations to TW_4.

CHAPTER 4

PUBLICATION 2

 The bibliographic details of the paper are given below and in the Reference list as [16].

Title	Non-cooperative detection of weak spread-spectrum signals in additive white Gaussian noise
Authors	J.D. Vlok and J.C. Olivier
Journal	IET Communications
Publication date	6 November 2012
Volume	6
Issue	16
Pages	2513–2524
doi	10.1049/iet-com.2011.0614
Print ISSN	1751-8628
Online ISSN	1751-8636

The submission timeline is given below.

Original manuscript submitted	26 August 2011
Decisioned: Revision requested	12 April 2012
Revised manuscript submitted	13 July 2012
Decisioned: Accepted	10 August 2012
Publication	6 November 2012

Non-cooperative detection of weak spread-spectrum signals in AWGN

J.D. Vlok[1] and *J.C. Olivier*[2]

[1]Defence, Peace, Safety & Security (DPSS), Council for Scientific and Industrial Research (CSIR), Pretoria 0001, South Africa

[2]School of Engineering, University of Tasmania, Hobart 7005, Australia

E-mail: jvlok@csir.co.za

Abstract: A semi-blind method based on PCA and sequence synchronisation is proposed to detect DSSS signals in a non-cooperative setting under low SNR conditions. The intercepted signal is segmented to form a detection matrix from which a feature is extracted through cyclic shifting. Signal detection is then performed using a test statistic based on this feature. The carrier frequency and sequence duration of the signal to be detected are assumed known. Theoretical analysis and a computer simulation study show that the performance of the new detection method is superior to classic ED in an AWGN channel.

4.1 Introduction

Cooperative detection techniques are used in communication systems where the receiver has perfect knowledge of all the parameters used by the transmitter. The receiver then uses detection and estimation of the parameters as necessary to identify the beginning of the transmission (e.g. in asynchronous communication) or to identify inactive channels (e.g. in cognitive radio). However, in non-cooperative applications such as spectrum surveillance and electronic interception, the receiver has no knowledge of the parameters used by the transmitter. Under these conditions blind detection and estimation techniques are used. If the receiver has some information on the parameters available or is able to estimate some of the parameters, semi-blind detection techniques may be used. This paper is concerned with the performance of semi-blind detection techniques, specifically for the detection of weak spread-spectrum signals.

In this paper new results and a new method for the detection of weak unknown deterministic signals in an AWGN channel are presented. The class of signals investigated is DSSS where a large transmission bandwidth is employed to hide the signal below the noise level. Several techniques addressing DSSS detection are available in the literature. ED of spread spectrum signals is presented in [48]. Techniques based on HOS analysis are presented in [111], and cyclostationary analysis where the autocorrelation of segments

of the intercepted signal is used as basis of detection is presented in [55]. Approaches to estimate the spreading sequence from the intercepted signal include signal correlation of synchronised sequences [77], PCA [91] and neural network techniques [94].

The new method presented in this paper is based on spreading sequence synchronisation [77] and PCA [91] which are used to extract the largest eigenvalue sequence as detection feature from the SCM of the intercepted signal. Two detection techniques are presented; the first technique uses the eigenvalue sequence directly and the second technique uses the frequency content of the eigenvalue sequence to perform detection. The techniques are semi-blind since certain aspects of the signal to be detected and the noise are assumed known. Signal knowledge assumed known include the carrier frequency and the sequence duration. These two parameters can be estimated using correlation techniques [55, 77] or the detection algorithm can sweep through a certain predefined range of these two parameters in order to search for signal activity. The noise statistics (including the PDF and power level) are assumed known in order to calculate the detection threshold. Noise statistics can be estimated by observing the surveillance band over long periods of time assuming that the signal of interest is not always present.

In this paper the performance of the detection techniques is evaluated and compared using the binary hypothesis testing approach (see [5]) over a range of SNRs. Although the detection algorithms do not assume knowledge of the SNR, the detection performance is expressed in terms thereof. This paper compares the performance of the two new techniques with classic ED, assuming the signal of interest is a baseband BPSK DSSS communication signal. It is shown that the new detection techniques have superior performance to classic ED under AWGN channel conditions.

The paper is organised as follows. Section 4.2 considers the target communication system and intercept receiver. Section 4.3 describes the new feature extraction technique and Section 4.4 evaluates the performance thereof in an AWGN channel. Section 4.5 reviews ED and presents the two new detection techniques. Section 4.6 investigates the computational complexity and Section 4.7 presents simulated detection performance results for each detection technique. The paper is concluded in Section 4.8.

4.2 Communication and detection systems

This section considers the target communication system and the intercept receiver used to detect communication activity originating from the target system in the surveillance band.

4.2.1 Target communication system

The target communication system considered in this paper is a BPSK DSSS system employing a Walsh spreading code of length $N = 64$. In order to establish the required SNR level at which non-cooperative detection must be performed, the error probability achievable by the intended receiver (the receiver of the target communication system for which the communication is intended) should be investigated. The error probability of BPSK in AWGN is well known as [26]

$$P_e = Q\left(\sqrt{\frac{2\epsilon_b}{N_0}}\right) \tag{4.1}$$

with $Q(\cdot)$ the tail probability of the standard normal distribution, ϵ_b the energy per bit and N_0 the single-sided noise power spectral density. The error probability can be expressed from (4.1) as (see Appendix A.3)

$$P_e = Q\left(\sqrt{N_s \text{ SNR}}\right) \tag{4.2}$$

with N_s the number of samples used to represent a single transmitted bit. Although (4.1) and (4.2) are equivalent, (4.2) is preferred in the non-cooperative context since the signal and surveillance bandwidths are not necessarily equal and (4.2) is required to illustrate the relation between the DSSS processing gain and the SNR advantage obtained by spreading.

If a single sample is used in the receiver to represent one chip of the spreading sequence then $N_s = N$. The bit error probability curves for the non-spread BPSK case ($N = 1$) and the spread case ($N = 64$) are shown in Fig. 4.1. The processing gain $P_G = 10 \log_{10} N \approx 18$ dB (for $N = 64$) is the SNR advantage obtained by spreading and corresponds to the SNR difference between the two curves for any P_e value in Fig. 4.1. The intended receiver would therefore be able to despread a DSSS signal with SNR $= -10$ dB (before despreading) to SNR $= 8$ dB (after despreading) to achieve communications at $P_e \approx 6 \times 10^{-3}$. Assuming the distances between the transmitter and intended receiver and the same transmitter and the intercept receiver are equivalent, the intercept receiver (not knowing the spreading code) is now faced with the challenge of detecting DSSS signal activity using a received signal with SNR $= -10$ dB. If the intercept receiver is further off in distance, an even lower SNR will result. Powerful detection algorithms are therefore required to enable the intercept receiver to compete with the target communication system and detect DSSS signals at very low SNR levels.

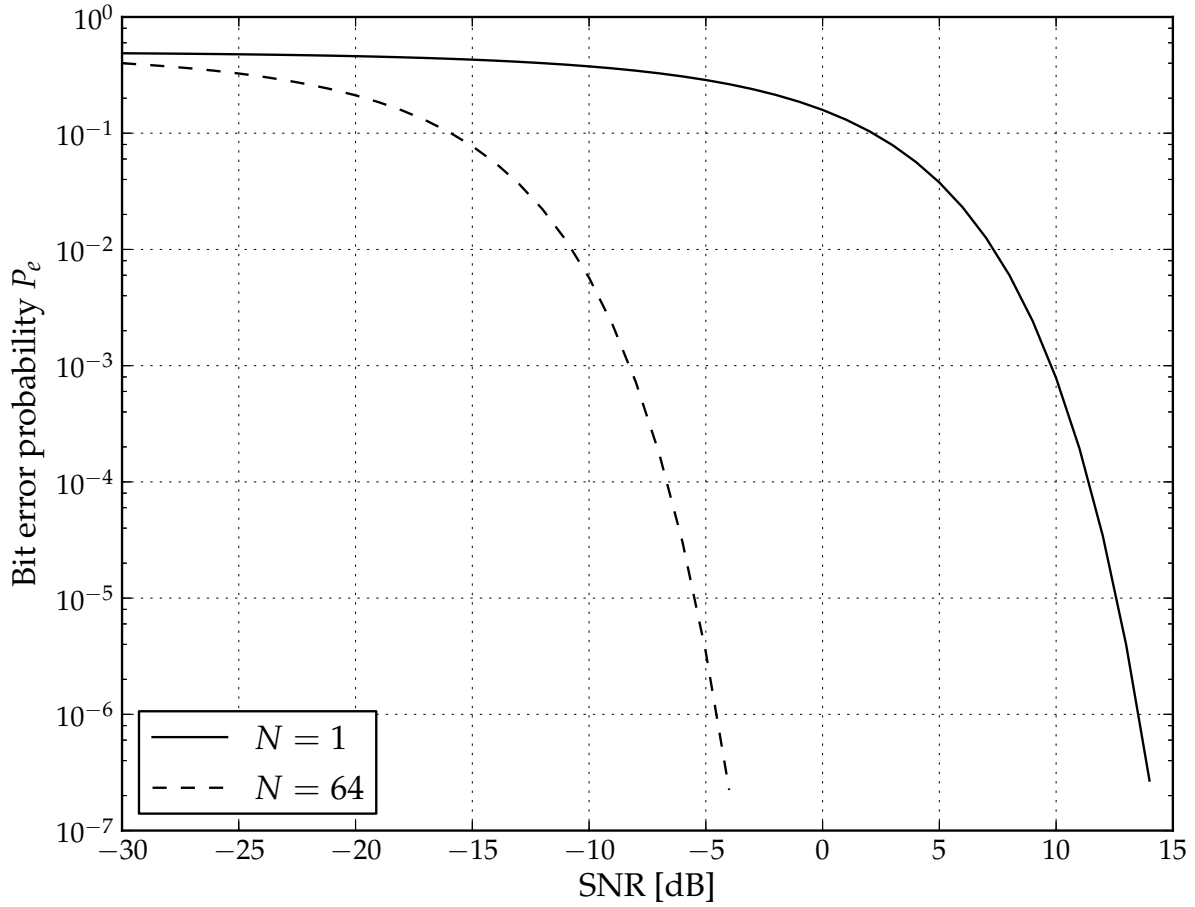


Figure 4.1: Bit error probability for non-spread ($N = 1$) and spread ($N = 64$) BPSK DSSS in AWGN.

4.2.2 Intercept receiver

The detection algorithms presented in this paper are envisaged to be implemented on a system with a receiver architecture similar to the one presented in [112] (see also Appendix A). The received RF signal is filtered using a wideband front-end filter, mixed down and then sampled with a high-speed (greater than twice the surveillance bandwidth) analogue-to-digital converter. The resultant baseband signal is then analysed without further filtering stages. For this receiver architecture, the signal model presented in [77] can be used to develop the detection algorithms. For simplicity of analysis, perfect timing and carrier synchronisation are assumed such that one sample is used to represent a single chip of the spreading sequence in the intercept receiver. The intercepted discrete DSSS signal can then be expressed as

$$\mathbf{y}(nT_c) = \sigma_x \mathbf{d}(nT_c)\mathbf{c}(nT_c) + \sigma_w \mathbf{w}(nT_c) \tag{4.3}$$

with the chip number $n = 1, 2, \ldots, N$ and T_c the chip interval. $\mathbf{c}$ is the length-N ($N >> 1$) pseudo-noise code sequence with period $T_{sym} = NT_c$ and $\mathbf{d}$ the data sequence assumed to be invariant over T_{sym}. Since the target communication system is a BPSK DSSS system,

both $\mathbf{c}$ and $\mathbf{d}$ are sequences with values ± 1. The noise sequence is represented by $\mathbf{w} \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$. The code, data and noise sequences are also assumed independent of each other. The constants σ_x and σ_w are included to scale the signal and noise sequences respectively in order to obtain different SNR values, using

$$\text{SNR} = \frac{\sigma_x^2}{\sigma_w^2} \tag{4.4}$$

which is the SNR before despreading.

Before the detection feature can be extracted from the intercepted signal, some signal formatting is required. The intercepted signal is firstly divided into non-overlapping segments of length equal to N. Secondly, these segments are stacked to form the $N \times N$ detection matrix given by

$$\mathbf{Y} = \sigma_x \mathbf{X} + \sigma_w \mathbf{W} \tag{4.5}$$

with $\mathbf{X}$ the data and $\mathbf{W}$ the noise matrices. The detection technique presented in this paper then takes $\mathbf{Y}$ as input and decides whether a DSSS signal is present or not.

4.3 Feature extraction

The detection feature is the largest eigenvalue sequence of the SCM of the intercepted signal stacked in the detection matrix $\mathbf{Y}$. The feature is extracted by performing PCA on $\mathbf{Y}$ while cyclically shifting the matrix. This section considers the noise-free scenario with $\sigma_x > 0$ and $\sigma_w = 0$ such that $\mathbf{Y} = \sigma_x \mathbf{X}$ (see also Appendix F). The scenario with $\sigma_w > 0$ is considered in Section 4.4. Although this section considers real binary data and spreading sequences, the result can easily be extended to the complex case.

4.3.1 Constructing the data matrix

The spreading sequence $\mathbf{c} = [c_1, c_2, \ldots, c_N]$ containing N chips is used to spread the data sequence $\mathbf{d} = [d_1, d_2, \ldots, d_N]$ containing N bits. If the receiver started intercepting at the beginning of a new sequence, the data matrix can be denoted as

$$\mathbf{X}_0 = \begin{bmatrix} d_1c_1 & d_1c_2 & \ldots & d_1c_{N-1} & d_1c_N \\ d_2c_1 & d_2c_2 & \ldots & d_2c_{N-1} & d_2c_N \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{N-1}c_1 & d_{N-1}c_2 & \ldots & d_{N-1}c_{N-1} & d_{N-1}c_N \\ d_Nc_1 & d_Nc_2 & \ldots & d_Nc_{N-1} & d_Nc_N \end{bmatrix} \tag{4.6}$$

which will be referred to as the aligned data matrix since the code sequence elements are aligned to the columns of $\mathbf{X}_0$; i.e. c_1 is in column 1, c_2 in column 2, and so on.

4.3.2 Principal component analysis

To perform PCA on the aligned data matrix, the covariance matrix of $\mathbf{X}_0$ is computed and then eigen decomposition is performed. The $N \times N$ covariance matrix is defined as (see [73, 76, 102])

$$\mathbf{COV}(\mathbf{X}_0) = E\left[\mathbf{X}_0^T \mathbf{X}_0\right] \tag{4.7}$$

with $[\cdot]^T$ the matrix transpose. The covariance matrix can be estimated using the SCM

$$\mathbf{R}(\mathbf{X}_0) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_n^T \mathbf{x}_n = \frac{1}{N}\mathbf{X}_0^T \mathbf{X}_0 \tag{4.8}$$

with $\mathbf{x}_n$ the n^{th} row of $\mathbf{X}_0$. The SCM of the aligned data matrix can therefore be expressed as

$$\mathbf{R}(\mathbf{X}_0) = \left[\frac{1}{N}\sum_{n=1}^{N} d_n^2\right] \mathbf{c}^T \mathbf{c} = \mathbf{c}^T \mathbf{c} \tag{4.9}$$

which is a positive-semidefinite (and by definition symmetric) matrix [75]. The simplification in (4.9) follows from the fact that $d_n^2 = 1$ for all values of n. By performing elementary row operations on $\mathbf{R}(\mathbf{X}_0)$ it can be shown that the row echelon form $\mathbf{R}_{\text{ech}}(\mathbf{X}_0)$ is an $N \times N$ matrix with only the first row being non-zero. $\mathbf{R}(\mathbf{X}_0)$ therefore has a rank of one and thus only one non-zero eigenvalue [5]. This can be expected since $\mathbf{X}_0$ has N linearly dependent rows. The non-zero eigenvalue of $\mathbf{R}(\sigma_x \mathbf{X}_0)$ can therefore be expressed as

$$\begin{aligned} \lambda_{X,1} &= \text{tr}(\mathbf{R}(\sigma_x \mathbf{X}_0)) \\ &= \sigma_x^2 \left[\sum_{n=1}^{N} c_n^2\right] \\ &= \sigma_x^2 N \end{aligned} \tag{4.10}$$

since the trace of a matrix equals the sum of its eigenvalues [75] and $c_n^2 = 1$. Furthermore, $\lambda_{X,1}$ is non-negative and real, since the eigenvalues of positive-semidefinite matrices are non-negative and real [5].

4.3.3 Cyclic shifting

Shifting is performed by discarding the first received sample and appending a new sample at the end of the sample sequence. The data matrix $\mathbf{X}_0$ after $k \in [0, N)$ shifts can then

be written as

$$\mathbf{X}_k = \begin{bmatrix} d_1c_{k+1} & \dots & d_1c_N & d_2c_1 & \dots & d_2c_k \\ d_2c_{k+1} & \dots & d_2c_N & d_3c_1 & \dots & d_3c_k \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ d_{N-1}c_{k+1} & \dots & d_{N-1}c_N & d_Nc_1 & \dots & d_Nc_k \\ d_Nc_{k+1} & \dots & d_Nc_N & d_{N+1}c_1 & \dots & d_{N+1}c_k \end{bmatrix}. \tag{4.11}$$

$\mathbf{X}_k$ for $k \geq N$ can be obtained in a similar way by noting that $\mathbf{X}_k$ will be in aligned form for any non-negative multiple of N shifts starting from $\mathbf{X}_0$. $\mathbf{X}_N$ will therefore have the same form as $\mathbf{X}_0$ though the contents of the entire matrix are shifted up one row with the top row removed and the bottom row replaced by a new sequence. The effect of a single sample shift on the data matrix is a cyclic left shift of all the columns followed by a cyclic upwards shift of the last column and finally replacing the bottom right element with the new sample. The resultant effect on the SCM is a shift diagonally upwards (in a North West direction) such that the first row and column are removed and the last row and column are replaced by new values due to the new data sample in the last row and column position of the data matrix. The SCM of $\mathbf{X}_k, k \in [0, N)$ can be given as

$$\mathbf{R}(\mathbf{X}_k) = \mathbf{A}_k \circ \left[\mathbf{c}_k^T \mathbf{c}_k\right] \tag{4.12}$$

where $\circ$ denotes the Hadamard product operator. $\mathbf{c}_k$ is the spreading code $\mathbf{c}$ cyclically left-shifted k times given as

$$\mathbf{c}_k = [c_{k+1}, c_{k+2}, \dots c_N, c_1, \dots c_k] \tag{4.13}$$

such that the first $N_k = N - k$ elements in $\mathbf{c}_k$ correspond to the last N_k elements in $\mathbf{c}$. $\mathbf{A}_k$ is an $N \times N$ coefficient matrix

$$\mathbf{A}_k = \begin{bmatrix} \alpha_{11}\mathbf{J}_{N_k,N_k} & \alpha_{12}\mathbf{J}_{N_k,k} \\ \alpha_{12}\mathbf{J}_{k,N_k} & \alpha_{22}\mathbf{J}_{k,k} \end{bmatrix} \tag{4.14}$$

consisting of four submatrices where $\mathbf{J}_{m,n}$ denotes the $m \times n$ matrix of ones. The coefficient values are

$$\begin{aligned} \alpha_{11} &= \frac{1}{N}\sum_{n=1}^{N} d_n^2 = 1 \\ \alpha_{12} &= \frac{1}{N}\sum_{n=1}^{N} d_n d_{n+1} \\ \alpha_{22} &= \frac{1}{N}\sum_{n=2}^{N+1} d_n^2 = 1 \end{aligned}$$

since $d_n^2 = 1$. The form of the coefficient matrix $\mathbf{A}_k$ arises from the fact that the first N_k columns of each row of $\mathbf{X}_k$ have the same data bit value as the corresponding row of $\mathbf{X}_0$. The last k columns of each row of $\mathbf{X}_k$ contain the data bit that overflowed from the row beneath (e.g. the last k columns of the first row of $\mathbf{X}_k$ contain d_2).

By performing elementary row operations on $\mathbf{R}(\mathbf{X}_k)$ it can be shown that the row echelon form $\mathbf{R}_{\text{ech}}(\mathbf{X}_k)$ is an $N \times N$ matrix with only the first two rows being non-zero (except when the data matrix is aligned - only the first row is then non-zero). The second row of $\mathbf{R}_{\text{ech}}(\mathbf{X}_k)$ contains N_k zeros followed by k non-zero elements and therefore $\mathbf{R}_{\text{ech}}(\mathbf{X}_k)$ contains a maximum of only two non-zero rows for all values of k. $\mathbf{R}(\mathbf{X}_k)$ therefore contains a maximum of two non-zero (also non-negative real) eigenvalues and the data matrix $\mathbf{X}_k$ consequently has a maximum of only two principal components.

4.3.4 Largest eigenvalues

Since $\mathbf{R}(\sigma_x \mathbf{X}_k)$ has a maximum of only two non-zero eigenvalues, its trace can be expressed in terms of the eigenvalues as

$$\text{tr}(\mathbf{R}(\sigma_x \mathbf{X}_k)) = \sum_{p=1}^{2} \lambda_{X,p} = \lambda_{X,1} + \lambda_{X,2} \tag{4.15}$$

with the largest eigenvalue $\lambda_{X,1} \geq \lambda_{X,2}$. Note that $\lambda_{X,p}$ depends on the shift parameter k but this dependence is omitted for the sake of simplifying notation. Using (4.12)-(4.14) the trace can be expressed in terms of the matrix elements as

$$\begin{aligned} \text{tr}\left(\mathbf{R}(\sigma_x \mathbf{X}_k)\right) &= \sigma_x^2 \sum_{n=1}^{N} \left[\mathbf{R}(\mathbf{X}_k)\right]_{nn} \\ &= \sigma_x^2 \left(\alpha_{11} \sum_{n=1}^{N_k} \left[\mathbf{c}_k^T \mathbf{c}_k\right]_{nn} + \alpha_{22} \sum_{n=N_k+1}^{N} \left[\mathbf{c}_k^T \mathbf{c}_k\right]_{nn} \right) \\ &= \sigma_x^2 N \end{aligned} \tag{4.16}$$

with $[\cdot]_{nn}$ denoting the diagonal entries of each matrix. The simplification in (4.16) follows from the fact that $\alpha_{11} = \alpha_{22} = 1$ and $c_n^2 = 1$. The sum of the eigenvalues therefore has a constant value irrespective of the value of k, assuming the signal power σ_x^2 remains constant. (4.15) can thus be stated as

$$\text{tr}(\mathbf{R}(\sigma_x \mathbf{X}_k)) = \lambda_{X,1} + \lambda_{X,2} = \sigma_x^2 N \ \ \forall\ k. \tag{4.17}$$

Whenever $\mathbf{X}_k$ is aligned, $\lambda_{X,1} = \sigma_x^2 N$ as in (4.10) and $\lambda_{X,2} = 0$. $\lambda_{X,1}$ therefore exhibits a pattern with period N as $\mathbf{X}_k$ is cyclically shifted, since $\lambda_{X,1} \geq \lambda_{X,2}$ for all values of k. $\lambda_{X,1}$ therefore reaches its maximum value $\sigma_x^2 N$ once during every complete cycle of shifting.

4.3.4.1 Eigenvalue bounds

Though it is possible to develop analytic expressions for $\lambda_{X,p}$, such equations would not be tractable since they are functions of all the elements of $\mathbf{c}$ and $\mathbf{d}$. Instead, in order to describe the behaviour of the largest eigenvalue, bounds on its variation are provided in this section. Since $\mathbf{R}(\mathbf{X}_k)$ has a maximum rank of two, its eigenvalues are the roots of a quadratic polynomial. By evaluating the eigenvalues of $\mathbf{R}(\sigma_x\mathbf{X}_k)$ with $d_n^2 = c_n^2 = 1$, it can be shown that the two non-zero eigenvalues can be expressed in the form (see also Appendix G)

$$\lambda_X = \frac{N^2 \pm \sqrt{\Delta}}{2N}\sigma_x^2 \tag{4.18}$$

with the discriminant of the quadratic polynomial $\Delta \in [0, N^4]$. From (4.18) the largest eigenvalue of $\mathbf{R}(\sigma_x\mathbf{X}_k)$ is bounded according to

$$\frac{N\sigma_x^2}{2} \leq \lambda_{X,1} \leq N\sigma_x^2. \tag{4.19}$$

Fig. 4.2 shows an example of the temporal variation of $\lambda_{X,1}$ when BPSK data is spread using the Walsh $N = 64$ code with $\sigma_x^2 = 1$. The bounds in Fig. 4.2 are given in (4.19). Note that the minimum bound for $\lambda_{X,1}$ is not necessarily reached during every cycle; the actual minimum value depends on the input data bits. From (4.17) the maximum bound for $\lambda_{X,1}$ is however reached during every cycle irrespective of the input data bit values.

4.4 Analysis in noise

This section considers the effect of noise on the feature extraction technique. The detection matrix can be written as

$$\mathbf{Y}_k = \sigma_x\mathbf{X}_k + \sigma_w\mathbf{W}_k \tag{4.20}$$

with the noise matrix in a similar form to $\mathbf{X}_k$ given by

$$\mathbf{W}_k = \begin{bmatrix} w_{1,k+1} & \dots & w_{1,N} & w_{2,1} & \dots & w_{2,k} \\ w_{2,k+1} & \dots & w_{2,N} & w_{3,1} & \dots & w_{3,k} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{N-1,k+1} & \dots & w_{N-1,N} & w_{N,1} & \dots & w_{N,k} \\ w_{N,k+1} & \dots & w_{N,N} & w_{N+1,1} & \dots & w_{N+1,k} \end{bmatrix} \tag{4.21}$$

containing real independent and identically distributed (i.i.d.) zero mean unity variance Gaussian samples.

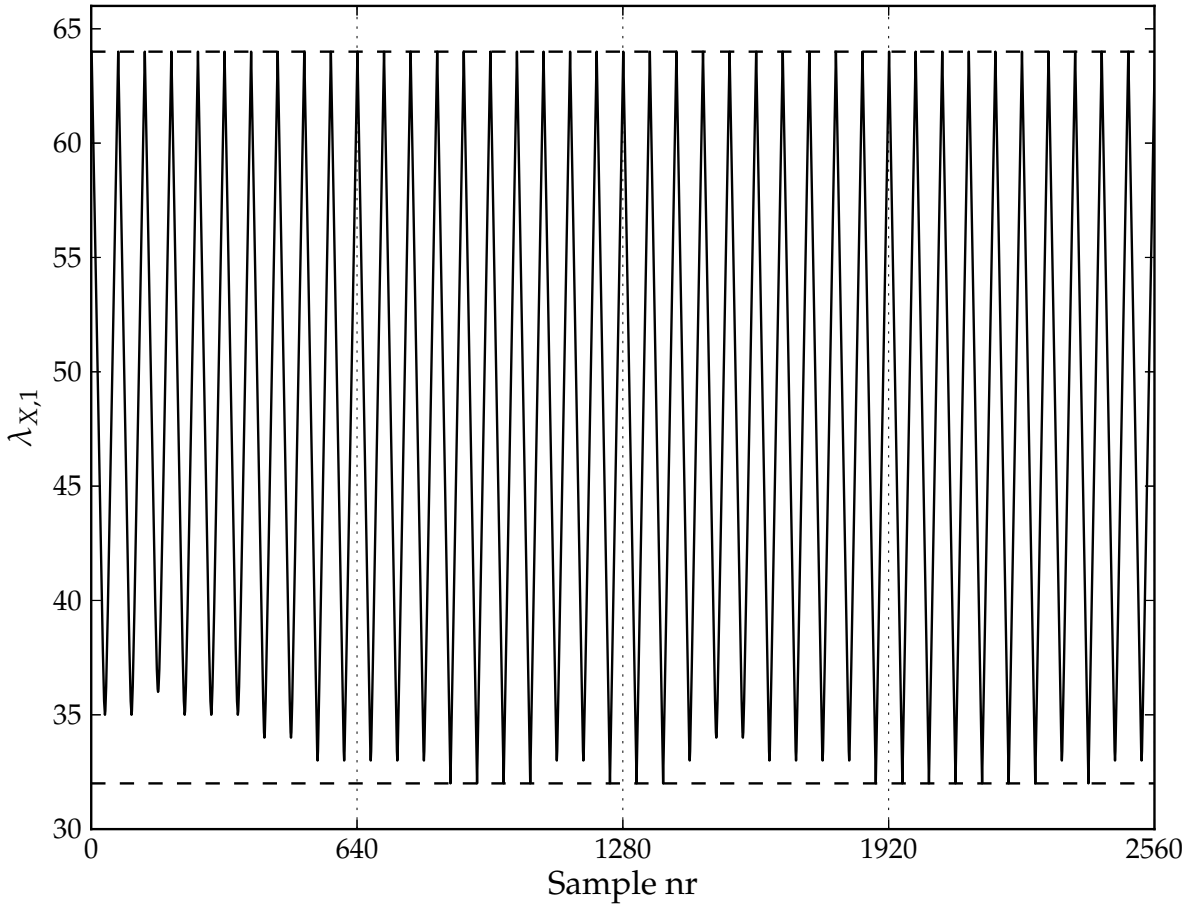


Figure 4.2: Example temporal representation of $\lambda_{X,1}$ for $\sigma_x^2 = 1$ and $N = 64$.

4.4.1 Noise-only scenario

If $\sigma_x = 0$, the intercepted signal will contain only noise. For any cyclic shift performed on $\mathbf{W}_k$, the SCM $\mathbf{R}(\sigma_w \mathbf{W}_k)$ is a Wishart matrix. The normalised largest eigenvalue of a Wishart matrix is distributed according to the TW law [102], which can be approximated using the Gamma distribution [68, 76]. The distribution of the largest eigenvalue $\lambda_{W,1}$ of $\mathbf{R}(\sigma_w \mathbf{W}_k)$ can therefore be approximated using the Gamma PDF [76]

$$\gamma(x) = \frac{(x - x_0)^{\alpha-1}}{\theta^\alpha \Gamma(\alpha)} \exp\left[\frac{-(x - x_0)}{\theta}\right] \tag{4.22}$$

with support region $[x_0, x_0 + 2\alpha\theta]$. The parameter values of (4.22) for $\mathbf{W}_k$ with real elements are

$$\begin{aligned} \alpha &= 46.5651 \\ \theta &= 0.1850\ \sigma_w^2\ \sigma_N / N \\ x_0 &= \sigma_w^2 \left(\mu_N - 9.8209 \sigma_N\right) / N \end{aligned}$$

with the centre and scaling parameters

$$\begin{aligned}\mu_N &= \left(\sqrt{N-1}+\sqrt{N}\right)^2 \\ \sigma_N &= \sqrt{\mu_N}\left(\frac{1}{\sqrt{N-1}}+\frac{1}{\sqrt{N}}\right)^{\frac{1}{3}}.\end{aligned}$$

Although $\mathbf{R}(\sigma_w\mathbf{W}_k)$ is always a Wishart matrix, all SCMs of $\mathbf{W}_k$ for shifts from k to $k+N-1$ (starting from any value of k) will contain some identical elements, though not in identical positions. This is due to the diagonal shift on the SCM as explained in Section 4.3.3. It can therefore be expected that a pattern with period N will arise in the eigenvalue sequence of the SCM as the noise matrix is cyclically shifted.

Fig. 4.3 shows an illustration of the temporal variation of $\lambda_{W,1}$ as $\mathbf{W}_k$ is cyclically shifted for $\sigma_w^2 = 1$ and $N = 64$. The bounds of $\lambda_{W,1}$ afforded by the support region

$$x_0 \leq \lambda_{W,1} \leq x_0 + 2\alpha\theta \tag{4.23}$$

and the measured mean value of $\lambda_{W,1}$ are also shown in Fig. 4.3. It can be shown that the PDF of $\lambda_{W,1}$ illustrated in Fig. 4.3 is accurately predicted by the Gamma approximation of (4.22) (see [76]).

4.4.2 Signal and noise scenario

The SCM of $\mathbf{Y}_k$ can be expressed as

$$\begin{aligned}\mathbf{R}(\mathbf{Y}_k) &= \mathbf{R}\left(\sigma_x\mathbf{X}_k + \sigma_w\mathbf{W}_k\right) \\ &= \sigma_x^2\mathbf{R}(\mathbf{X}_k) + \sigma_w^2\mathbf{R}(\mathbf{W}_k) + \mathbf{E}_k\end{aligned} \tag{4.24}$$

with the error matrix expressed as

$$\mathbf{E}_k = \frac{\sigma_x\sigma_w}{N}\left[\mathbf{X}_k^T\mathbf{W}_k + \mathbf{W}_k^T\mathbf{X}_k\right]. \tag{4.25}$$

This section considers bounds on the largest eigenvalue $\lambda_{Y,1}$ of $\mathbf{R}(\mathbf{Y}_k)$ using the Weyl inequalities [113–115]. Although it is often assumed that $\mathbf{E}_k = \mathbf{0}$ (see for example [113]), the effect of the error matrix is also considered here.

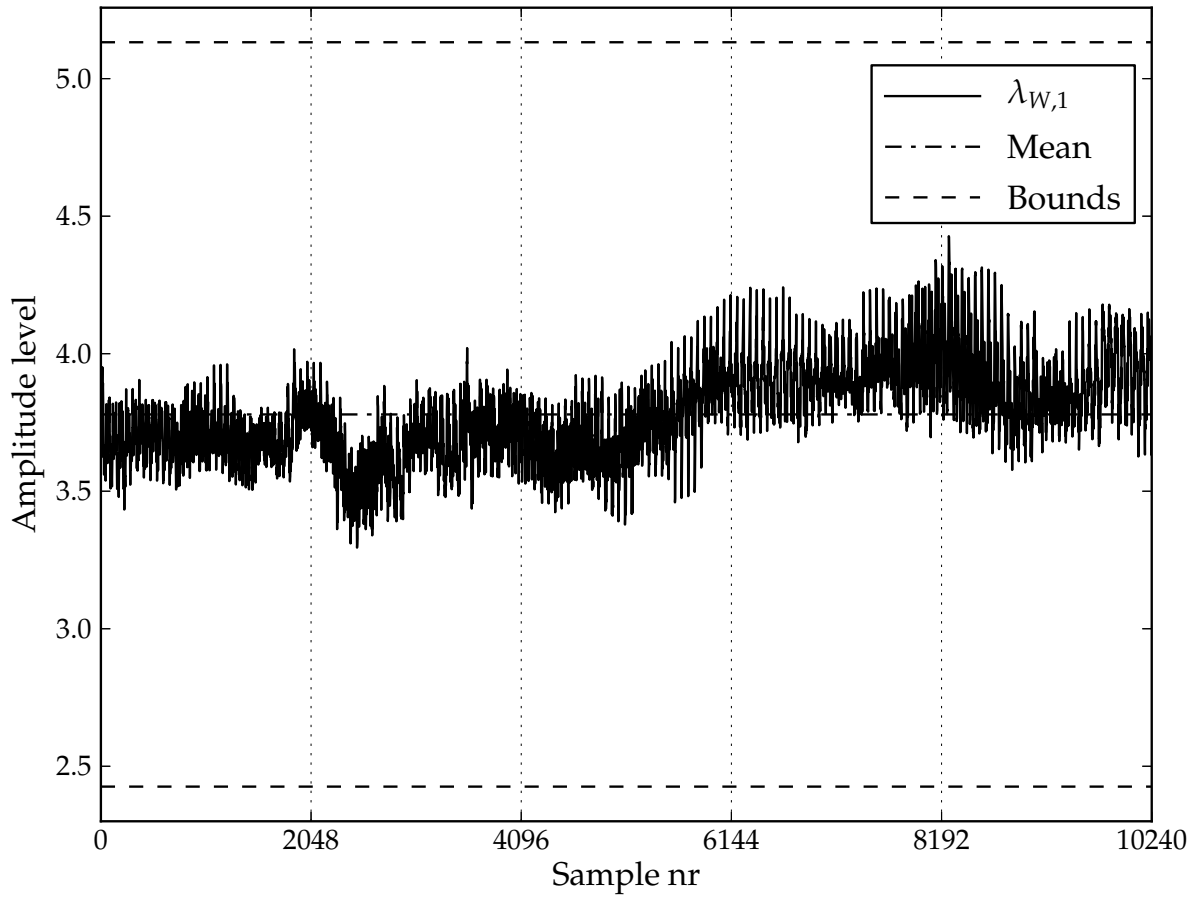


Figure 4.3: Example temporal representation of $\lambda_{W,1}$ for $\sigma_w^2 = 1$ and $N = 64$.

4.4.2.1 Weyl inequalities

The eigenvalue spectrum of the sum of two $N \times N$ Hermitian (or real symmetric) matrices $\mathbf{A}$ and $\mathbf{B}$ can be described using the Weyl inequalities. Arranging all eigenvalues in decreasing order ($\lambda_1 \geq \cdots \geq \lambda_N$) the p^{th} largest eigenvalue of the matrix sum $\mathbf{A} + \mathbf{B}$ is bounded according to

$$L_p \leq \lambda_p(\mathbf{A} + \mathbf{B}) \leq U_p \tag{4.26}$$

with the upper and lower bounds respectively given by [116]

$$U_p = \min\{\lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}) \ : \ i + j = p + 1\} \tag{4.27}$$

$$L_p = \max\{\lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}) \ : \ i + j = p + N\} \tag{4.28}$$

with the eigenvalue index values $i, j \in [1, N]$. Note that any eigenvalue sum combination in (4.27) and (4.28) is respectively a valid upper and lower bound. Taking the minimum in (4.27) and the maximum in (4.28) will however result in the tightest bounds. The upper bound for the largest eigenvalue ($p = 1$) can be obtained from (4.27) as

$$U_1 = \lambda_1(\mathbf{A}) + \lambda_1(\mathbf{B}) \tag{4.29}$$

since there is only one solution ($i = j = 1$) to the index equation $i + j = 2$. The lower bound for $p = 1$ can be obtained from (4.28) as

$$L_1 = \max\{[\lambda_1(\mathbf{A}) + \lambda_N(\mathbf{B})], [\lambda_2(\mathbf{A}) + \lambda_{N-1}(\mathbf{B})], \ldots, [\lambda_N(\mathbf{A}) + \lambda_1(\mathbf{B})]\} \tag{4.30}$$

since $i + j = 1 + N$. In order to use the Weyl inequalities, it can easily be verified mathematically that the matrices $\sigma_x^2\mathbf{R}(\mathbf{X}_k)$, $\sigma_w^2\mathbf{R}(\mathbf{W}_k)$, $\sigma_x^2\mathbf{R}(\mathbf{X}_k) + \sigma_w^2\mathbf{R}(\mathbf{W}_k)$ and $\mathbf{E}_k$ are all symmetric.

4.4.2.2 Eigenvalue bounds of $\mathbf{E}_k$

When N is small or correlation exists between the signal and noise sequences, the error matrix $\mathbf{E}_k$ given in (4.25) cannot be neglected in (4.24). This section considers the upper and lower limits, respectively, of the largest eigenvalue $\lambda_{E,1}$ and the smallest eigenvalue $\lambda_{E,N}$ of $\mathbf{E}_k$ in order to derive bounds for $\lambda_{Y,1}$. Each element of $\mathbf{X}_k$ in (4.11) equals ± 1 and each element of $\mathbf{W}_k$ given in (4.21) is independently drawn from $\mathcal{N}(0, 1)$. Each element of $\mathbf{X}_k^T\mathbf{W}_k$ (and $\mathbf{W}_k^T\mathbf{X}_k$) is therefore the sum of N random variables drawn independently from $\mathcal{N}(0, 1)$ such that each element is distributed according to $\mathcal{N}(0, N)$. The sum $\mathbf{X}_k^T\mathbf{W}_k + \mathbf{W}_k^T\mathbf{X}_k$ will therefore result in a matrix with main diagonal elements distributed according to $\mathcal{N}(0, 4N)$ and all other elements distributed according to $\mathcal{N}(0, 2N)$. Since $\mathbf{E}_k$ is a real symmetric matrix with all entries symmetrically distributed around 0, the eigenvalue limits can be given as [117]

$$\lambda(\mathbf{E}_k) \in [-Nb, Nb] \tag{4.31}$$

with the entries of $\mathbf{E}_k$ in the range $[-b, b]$. Since the elements of $\mathbf{E}_k$ are normally distributed and [117] requires a fixed range, the support region of the matrix entries should be truncated. The loss in tail probability mass of $\mathcal{N}(0, \sigma^2)$ due to truncation can be described in terms of the Q-function [26] as

$$Q\left(\frac{b}{\sigma}\right) = Q(\beta) \tag{4.32}$$

with

$$b = \beta\sigma\{(\mathbf{E}_k)_{ij}\} \tag{4.33}$$

where β is the factor of the standard deviation at which $\mathcal{N}(0, \sigma^2)$ is truncated. The standard deviation in (4.33) is denoted by $\sigma\{(\mathbf{E}_k)_{ij}\}$ with $i = j$ for the diagonal entries and $i \neq j$ for the off-diagonal entries. Although all entries of $\mathbf{E}_k$ are confined to the interval of the diagonal entries ($(\mathbf{E}_k)_{jj}$ has twice the variance of $(\mathbf{E}_k)_{ij}$), using $i = j$ in (4.33) will result in very conservative bounds. It can be shown that the lower and upper bounds in (4.31) would only be attained if all entries of $\mathbf{E}_k$ were, respectively, $-b$ and b

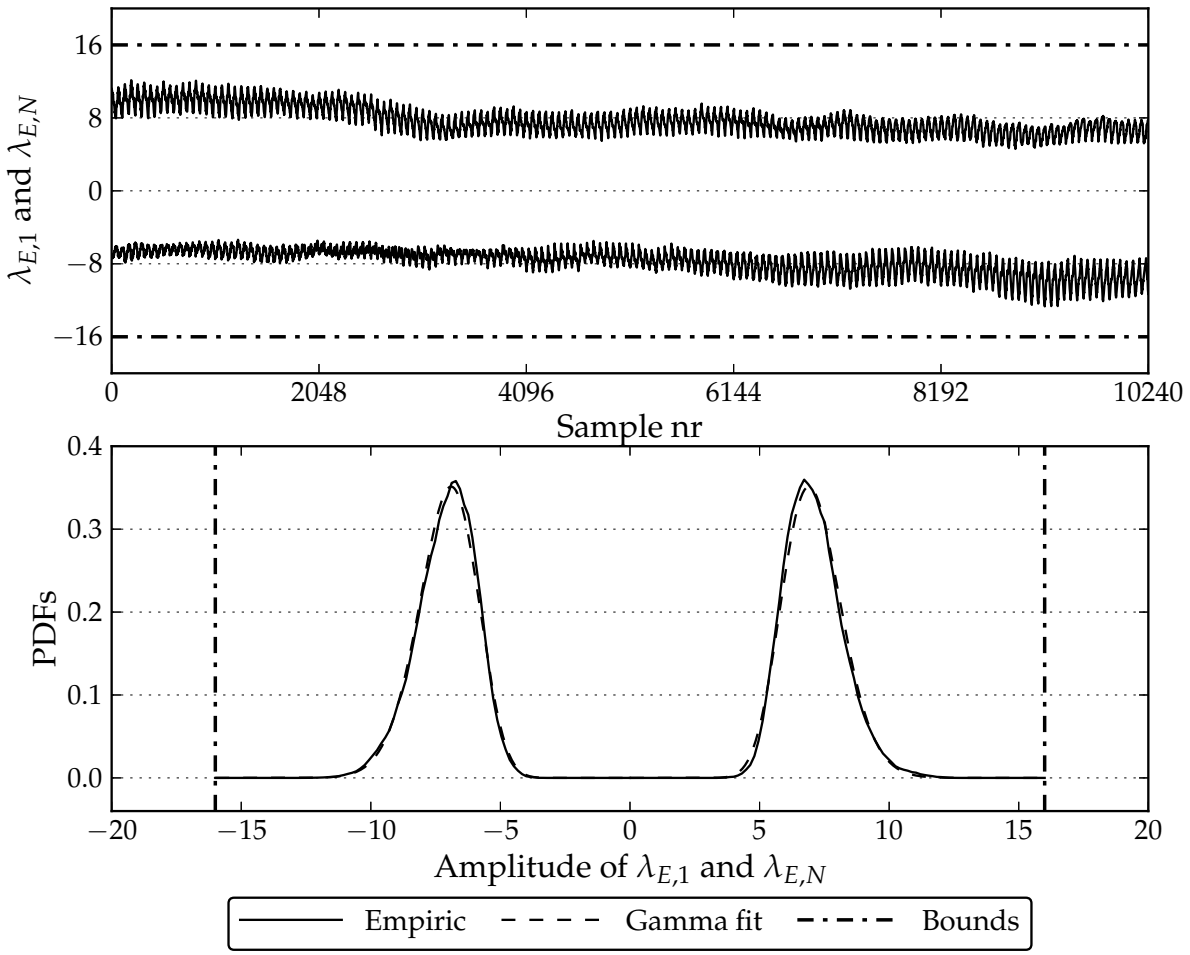


Figure 4.4: Largest and smallest eigenvalues of $\mathbf{E}_k$ and the corresponding PDF of each for $\sigma_x^2 = \sigma_w^2 = 1$ and $N = 64$.

(see [118]). The standard deviation of the majority of the matrix elements (i.e. the off-diagonal entries) is often used to calculate eigenvalue bounds (see for example [119, 120] for the case where all matrix elements are i.i.d.). Using (4.25) and the standard deviation of the off-diagonal entries in (4.33), the eigenvalue bounds of $\mathbf{E}_k$ can therefore be expressed using (4.31) with

$$Nb = \beta\sigma_x\sigma_w\sqrt{2N}. \tag{4.34}$$

The tightness of the bounds further depends on the value of the factor β in (4.34). Empirical results indicate that choosing $\beta = \sqrt{2}$ results in valid bounds such that

$$\lambda(\mathbf{E}_k) \in [-2\sigma_x\sigma_w\sqrt{N}, 2\sigma_x\sigma_w\sqrt{N}]. \tag{4.35}$$

Simulated sequences of the largest and smallest eigenvalues of $\mathbf{E}_k$ for $\sigma_x^2 = \sigma_w^2 = 1$ and $N = 64$ and the bounds of (4.35) are shown in Fig. 4.4. The PDF of each sequence is also shown in the figure; $\lambda_{E,1}$ seems to have a Gamma PDF and the two PDFs are mirror images.

4.4.2.3 Bounds of $\lambda_{Y,1}$ when $\mathbf{E}_k = \mathbf{0}$

This section considers the bounds of the largest eigenvalue of $\mathbf{R}(\mathbf{Y}_k)$ for the scenario where the data and noise are uncorrelated (when N is large) such that (4.24) can be written as

$$\mathbf{R}(\mathbf{Y}_k) = \sigma_x^2 \mathbf{R}(\mathbf{X}_k) + \sigma_w^2 \mathbf{R}(\mathbf{W}_k). \tag{4.36}$$

The eigenvalue bounds of $\mathbf{R}(\mathbf{Y}_k)$ can be expressed using (4.26) to (4.28) with $\mathbf{A} = \sigma_x^2 \mathbf{R}(\mathbf{X}_k)$ and $\mathbf{B} = \sigma_w^2 \mathbf{R}(\mathbf{W}_k)$. According to (4.29) the upper bound U_1 of $\lambda_{Y,1}$ is

$$\begin{aligned} U_1 &= \lambda_{X,1} + \lambda_{W,1} \\ &= N\sigma_x^2 + x_0 + 2\alpha\theta \end{aligned} \tag{4.37}$$

obtained by summing the upper bounds of (4.19) and (4.23). According to (4.30) the lower bound L_1 can be expressed as

$$\begin{aligned} L_1 &= \max\{[\lambda_{X,1} + \lambda_{W,N}], [\lambda_{X,2} + \lambda_{W,N-1}], [\lambda_{W,N-2}], \ldots, [\lambda_{W,1}]\} \\ &= \max\{\lambda_{X,1}, \lambda_{W,1}\} \end{aligned} \tag{4.38}$$

using the fact that $\lambda_{X,i} = 0$ for $i > 2$, since $\mathbf{R}(\sigma_x \mathbf{X}_k)$ has a maximum of only two non-zero eigenvalues (see Section 4.3.4) and the fact that the smallest eigenvalue $\lambda_{W,N}$ approaches zero (even for small values of N [113,121]) since the detection matrix is square. (From [121] the asymptotic lower bound of the eigenvalue spectrum of the SCM of an $M \times N$ noise matrix is $b_- = \left(1 - \sqrt{N/M}\right)^2$. In the case considered here $N = M$, such that $b_- = 0$.) By substituting the lower bounds of (4.19) and (4.23) into (4.38), the lower bound can be expressed as

$$L_1 = \max\left\{\frac{N\sigma_x^2}{2}, x_0\right\} \tag{4.39}$$

with $N\sigma_x^2/2 > x_0$ if SNR $> 2(\mu_N - 9.8209\sigma_N)/N^2$ using the value of x_0 from Section 4.4.1.

4.4.2.4 Bounds of $\lambda_{Y,1}$ when $\mathbf{E}_k \neq \mathbf{0}$

In this section new bounds L_{1E} and U_{1E} for the largest eigenvalue of $\mathbf{R}(\mathbf{Y}_k)$ in (4.24) are derived for the case when the error matrix is non-zero. The same method used in Section 4.4.2.3 is applied, though with $\mathbf{A} = \sigma_x^2 \mathbf{R}(\mathbf{X}_k) + \sigma_w^2 \mathbf{R}(\mathbf{W}_k)$ and $\mathbf{B} = \mathbf{E}_k$. Using (4.29) and the upper bounds given in (4.35) and (4.37) the new upper bound can be written as

$$U_{1E} = U_1 + 2\sigma_x \sigma_w \sqrt{N}. \tag{4.40}$$

Similarly, using (4.30) and the lower bounds given in (4.35) and (4.39) the new lower bound can be shown to be

$$L_{1E} = L_1 - 2\sigma_x \sigma_w \sqrt{N}. \tag{4.41}$$

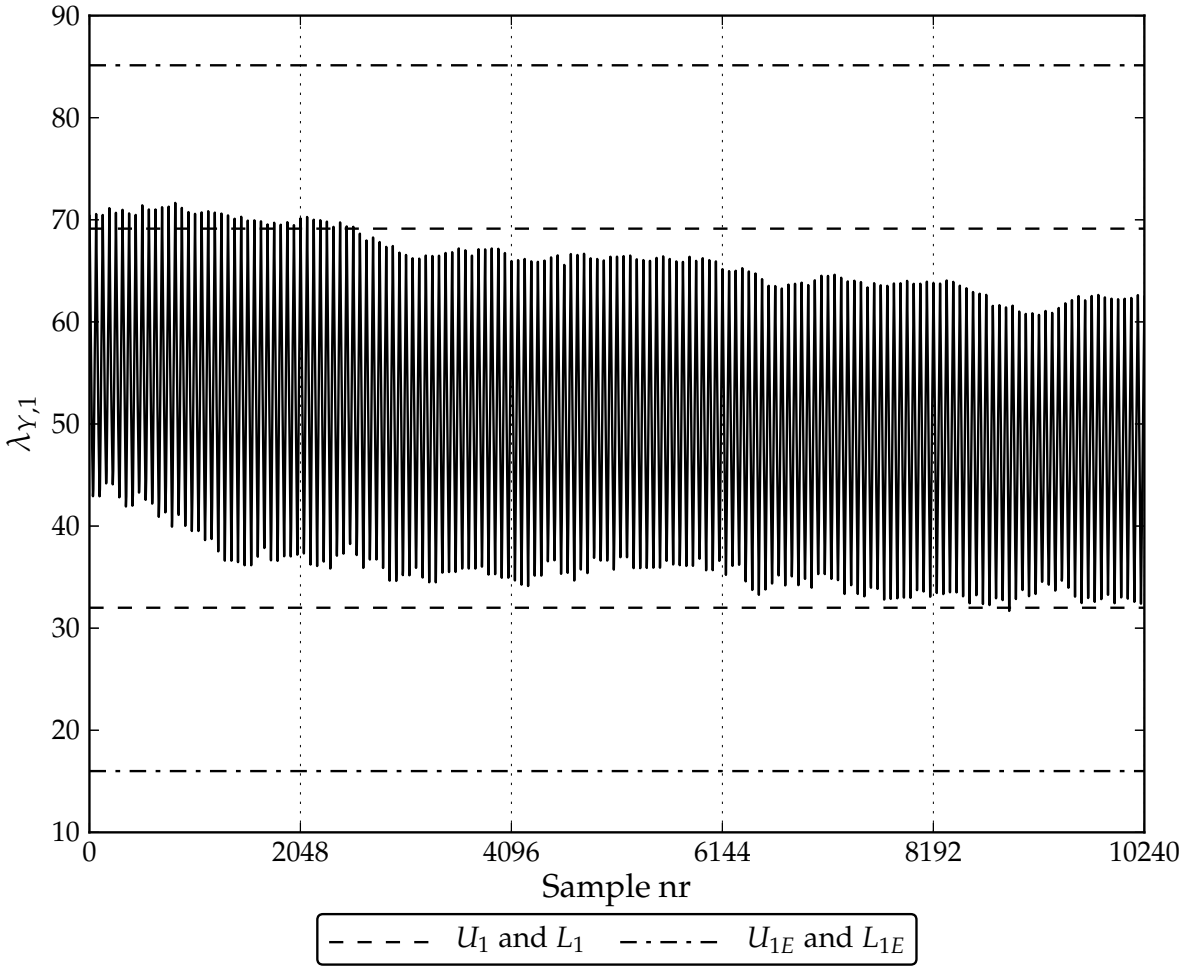


Figure 4.5: Example temporal representation and bounds of $\lambda_{Y,1}$ for $\sigma_x^2 = \sigma_w^2 = 1$ and $N = 64$.

Fig. 4.5 shows an example of the temporal variation of $\lambda_{Y,1}$ with the bounds from (4.37), (4.39), (4.40) and (4.41) for $\sigma_x^2 = \sigma_w^2 = 1$ and $N = 64$. The bounds U_1 and L_1 are exceeded in the figure since they neglect the error matrix.

4.5 Detection techniques

The detection of an unknown DSSS signal can be cast as a binary hypothesis testing problem

$$\begin{aligned} \mathcal{H}_0: \quad \mathbf{y}(nT_c) &= \sigma_w \mathbf{w}(nT_c) \\ \mathcal{H}_1: \quad \mathbf{y}(nT_c) &= \sigma_x \mathbf{x}(nT_c) + \sigma_w \mathbf{w}(nT_c) \end{aligned}$$

using the intercepted signal of (4.3) with the unknown deterministic signal $\mathbf{x}(nT_c) = \mathbf{d}(nT_c)\mathbf{c}(nT_c)$. $\mathcal{H}_0$ signifies the signal absent (noise-only) case and $\mathcal{H}_1$ the signal present case. The task of the detection receiver is to decide whether the signal of interest is present or not, based on a detection threshold derived from a test statistic under the $\mathcal{H}_0$ hypothesis.

This section considers three techniques that can be used to detect DSSS signals. The techniques are classic ED and two new detection techniques based on eigen analysis presented in Sections 4.3 and 4.4. For ED the theoretical false alarm probability P_{FA} and detection probability P_D are derived in terms of the threshold. For the other two techniques, P_{FA} is derived in terms of the threshold and P_D is then determined using computer simulation with the results presented in Section 4.7.

4.5.1 Energy detection

Using the generalised likelihood ratio test, it can be shown that hypothesis $\mathcal{H}_1$ should be chosen if [5]

$$T(\mathbf{y}) = \sum_{n=1}^{V} \left[\mathbf{y}(nT_c)\right]^2 > \zeta \tag{4.42}$$

with $T(\cdot)$ signifying the test statistic function calculated over V samples and ζ the detection threshold which can be calculated from the required P_{FA} according to the NP criteria. Clearly, (4.42) signifies ED. Assuming perfect knowledge of the unknown signal, a performance upper bound of the NP detector can be determined [5]. Assuming the signal of interest is BPSK with $\mathbf{x}(nT_c) = \pm 1$, the hypothesis test can be stated as

$$\begin{aligned} \mathcal{H}_0: \quad T(\mathbf{y}) &= \sum_{n=1}^{N} \left[\sigma_w \mathbf{w}(nT_c)\right]^2 \\ \mathcal{H}_1: \quad T(\mathbf{y}) &= \sum_{n=1}^{N} \left[\pm\sigma_x + \sigma_w \mathbf{w}(nT_c)\right]^2 . \end{aligned}$$

It can then be shown that

$$\begin{aligned} \mathcal{H}_0: \quad \frac{T(\mathbf{y})}{\sigma_w^2} &\sim \mathcal{X}_N^2 \\ \mathcal{H}_1: \quad \frac{T(\mathbf{y})}{\sigma_w^2} &\sim \mathcal{X}_N'^2(p_{nc}) \end{aligned}$$

with $\mathcal{X}_N^2$ the central and $\mathcal{X}_N'^2$ the non-central Chi-square distributions with non-centrality parameter

$$p_{nc} = \frac{N\sigma_x^2}{\sigma_w^2} = N \times \text{SNR}.$$

The false alarm probability can then be expressed as

$$\begin{aligned} P_{FA} &= p\left\{T(\mathbf{y}) > \zeta; \mathcal{H}_0\right\} \\ &= p\left\{\frac{T(\mathbf{y})}{\sigma_w^2} > \frac{\zeta}{\sigma_w^2}; \mathcal{H}_0\right\} \\ &= Q_{\mathcal{X}_N^2}\left(\zeta'\right) \end{aligned} \tag{4.43}$$

with $Q_{\mathcal{X}_N^2}(\zeta')$ the right tail probability of $\mathcal{X}_N^2$ and the scaled threshold $\zeta' = \zeta/\sigma_w^2$. Likewise, the detection probability can be expressed as

$$P_D = Q_{\mathcal{X}_N'^2}(\zeta'). \tag{4.44}$$

4.5.2 Eigen detection technique 1

The first eigen detection algorithm is based on the feature extraction technique presented in Section 4.3. The algorithm contains three phases. Firstly, the intercepted signal is segmented and stacked to form the detection matrix. Secondly, the principal component of the detection matrix is calculated and thirdly the detection matrix is cyclically shifted. The principal component is calculated for every cycle, such that the largest eigenvalue sequence $\lambda_{Y,1}$ is generated. A detection is declared if $\lambda_{Y,1}$ exceeds a given threshold ζ. This threshold can be calculated from the right tail probability

$$P_{FA} = Q_g(\zeta) = \int_{\zeta}^{\infty} \gamma(x)dx \tag{4.45}$$

with $\gamma(x)$ the PDF of $\lambda_{W,1}$ given in (4.22). The threshold is therefore

$$\zeta = Q_g^{-1}(P_{FA}). \tag{4.46}$$

The SNR limit at which perfect detection ($P_{FA} = 0$ and $P_D = 1$) is achieved can be calculated using the upper bound on $\lambda_{W,1}$ given in (4.23) and the lower bound on $\lambda_{Y,1}$ given in (4.41). For SNR values above this limit, $\lambda_{W,1}$ will always be less than $\lambda_{Y,1}$ such that

$$x_0 + 2\alpha\theta < L_1 - 2\sigma_x\sigma_w\sqrt{N} \tag{4.47}$$

with $L_1 = N\sigma_x^2/2$ from (4.39). The SNR limit can be calculated by solving the quadratic inequality for σ_x/σ_w that results when the parameter values x_0, α and θ from Section 4.4.1 are substituted into (4.47). For example, perfect detection for $N = 64$ using technique 1 can be shown to be possible if SNR > -2.828 dB. For non-zero values of P_{FA}, $P_D = 1$ will be achieved below this SNR limit.

4.5.3 Eigen detection technique 2

As illustrated in Sections 4.3 and 4.4, $\lambda_{Y,1}$ exhibits a strong pattern with period N (or T_{sym}) as the detection matrix is cyclically shifted (if a DSSS signal is present). Detection could therefore be based on the frequency content $I(f_0)$ of $\lambda_{Y,1}$ at $f_0 = 1/T_{sym}$. Eigen detection technique 2 uses sinusoidal detection to isolate $I(f_0)$ and declares a detection if

a given threshold ζ is exceeded. The test statistic is therefore given by

$$T(\lambda_{Y,1}) = I(f_0) = \left[\frac{1}{\sqrt{V}}\sum_{n=1}^{V}\lambda_{Y,1}(nT_s)\cos(2\pi f_0 nT_s)\right]^2 + \left[\frac{1}{\sqrt{V}}\sum_{n=1}^{V}\lambda_{Y,1}(nT_s)\sin(2\pi f_0 nT_s)\right]^2 \tag{4.48}$$

which is the periodogram equation [5] for sinusoidal detection of unknown amplitude and phase over V samples. In order to calculate $I(f_0)$ for $\lambda_{Y,1}$ formed by the content of a single detection matrix, the number of samples should be $V = N^2$. V can however be chosen even longer, since the repetition pattern continues beyond N^2. Under $\mathcal{H}_0$, the PDF of $I(f_0)$ is related to the central Chi-square distribution with two degrees of freedom [5], since each term of the sum in (4.48) is normally distributed before squaring. The false alarm probability can therefore be given as

$$P_{FA} = Q_{\mathcal{X}_2^2}(\zeta') = \exp\left(-\frac{\zeta'}{2}\right) \tag{4.49}$$

similar to (4.43), from which the threshold can determined using

$$\zeta' = -2\,\ln(P_{FA})\,. \tag{4.50}$$

4.6 Computational complexity

Computational complexity can be measured in terms of the total number of elementary arithmetic operations ($+$, $-$, $\times$, $\div$) required to execute an algorithm. The complexity of other operations can also be expressed in terms of the elementary operations. It will be assumed that the complexity of square rooting and division is the same as that of multiplication [122]. Likewise, addition and subtraction will be assumed equivalent. The complexity of an algorithm is typically dominated by the number of multiplications required, since multiplication is more complex than addition. This section considers the computational complexity of the three algorithms presented in Section 4.5 in terms of the number of arithmetic operations required to extract the test statistic from the intercepted data. Additional processing, such as memory allocation and movement of data values, is not taken into account. The equivalent number of multiplications and additions required for each algorithm are summarised in Table 4.1 and measured execution times are given Section 4.7.3.

4.6.1 Energy detection

ED is implemented in this paper using (4.42) with a fixed size window moving over the intercepted data. The window size V is equal to the size of the detection matrix N^2. The energy in a single window can therefore be calculated using N^2 multiplications and N^2-1 additions.

Table 4.1: Comparative computational complexities of the three detection algorithms.

Algorithm	Multiplications	Additions
ED	N^2	$N^2 - 1$
Tech 1	$N^3 + M(N^2 + 4N + 2)$	$N^3 - N^2 + M(N^2 + 3N - 3)$
Tech 2	$2V + 3$	$2V - 1$

4.6.2 Eigen detection technique 1

The most computationally intensive step of the first eigen detection algorithm (described in Section 4.5.2) is the calculation of the principal component, which involves calculating a SCM and its largest eigenvalue. The computational complexity of this step will subsequently be considered for a single cyclic shift of the detection matrix.

4.6.2.1 Calculation of the SCM

Equation (4.8) indicates that the SCM is calculated by performing matrix multiplication on two $N \times N$ matrices and scaling the resultant matrix by N. It might be more economical to rather scale the detection threshold value (see [76]) than to scale each of the N^2 elements of the resultant matrix (if the hardware platform allows large numbers). For the analysis presented here, it will be assumed that calculation of the SCM is equivalent to a single matrix multiplication operation. The matrix multiplication operation $\mathbf{F} = \mathbf{GH}$ can be implemented through the conventional approach [75] using

$$f_{ij} = \sum_{k=1}^{N} g_{ik} h_{kj} \tag{4.51}$$

where $[\cdot]_{ij}$ is the ij^{th} matrix element. Calculation of each element of $\mathbf{F}$ therefore requires N multiplications and $N - 1$ additions. Since there are N^2 elements to be calculated, the overall matrix multiplication operation requires N^3 multiplications and $N^2(N - 1)$ additions.

4.6.2.2 Calculation of the largest eigenvalue

Eigenvalues are usually approximated using iterative numerical methods, such as the power method [75, 123]. The power method is used to determine the dominant eigenvalue (or simply the largest eigenvalue, if the eigenvalues are non-negative real) and associated eigenvector of the $N \times N$ matrix $\mathbf{F}$. The algorithm takes as input a non-zero column vector $\mathbf{v}_0$ (with unit Euclidean norm) which could be an approximation to the associated eigenvector or a random vector. If $\mathbf{F}$ is symmetric, the algorithm can be simplified and

the largest eigenvalue λ_m and eigenvector $\mathbf{v}_m$ can then approximated using [123]

$$\begin{aligned} \mathbf{u}_m &= \mathbf{F}\mathbf{v}_m && (4.52)\\ \lambda_m &= \mathbf{v}_m^T\mathbf{u}_m && (4.53)\\ \mathbf{v}_{m+1} &= \frac{\mathbf{u}_m}{||\mathbf{u}_m||_2} && (4.54)\\ \varepsilon &= ||\mathbf{v}_{m+1} - \mathbf{v}_m||_2 && (4.55) \end{aligned}$$

for each $m^{\text{th}}, m \geq 0$ iteration, with $||\cdot||_2$ signifying the Euclidean norm. The convergence of the algorithm depends on $\mathbf{F}$ and the choice of $\mathbf{v}_0$. Although bad choices of $\mathbf{v}_0$ may result in slow convergence or division by zero errors (the algorithm will then not converge and will need to be restarted with a new $\mathbf{v}_0$), the algorithm will converge for almost all initial guesses of $\mathbf{v}_0$ if $\mathbf{F}$ has a single dominant eigenvalue (see [124]). The algorithm should therefore check at each iteration whether $||\mathbf{u}_m||_2 = 0$. The number of iterations can be limited to a specified maximum or the algorithm can stop when the tolerance ε decreases below a predefined value. The computational complexity of the symmetric power method is considered below, assuming the algorithm will converge.

Equation (4.52) is a matrix-vector multiplication which is the same as (4.51) with $j = 1$. Since only N elements are calculated, (4.52) requires N^2 multiplications and $N(N-1)$ additions. (4.53) is a dot-product operation and requires N multiplications and $N-1$ additions. (4.54) requires N divisions and the calculation of the Euclidean norm, which requires N multiplications, $N-1$ additions and one square root calculation. Lastly, (4.55) requires N subtractions and again the calculation of an Euclidean norm. Assuming M iterations are required to reach a certain ε, calculation of the largest eigenvalue therefore requires the equivalent processing time of $M(N^2+4N+2)$ multiplications and $M(N^2+3N-3)$ additions.

4.6.2.3 Possible simplifications

A more efficient (though with reduced numerical stability) method to perform matrix multiplication is Strassen's algorithm [75, 125]. The algorithm recursively splits the matrices to be multiplied into smaller submatrices and performs matrix multiplication using less multiplications, but more additions than the conventional approach. The multiplicative cost of the Strassen algorithm is $O(N^{\log_2 7}) \approx O(N^{2.807})$, instead of $O(N^{\log_2 8}) = O(N^3)$ of the conventional approach.

An alternative method to approximate the largest eigenvalue is the trace method [126] using

$$\text{tr}\,(\mathbf{F}^{\,r}) = \sum_{i=1}^{N} \lambda_i^r. \qquad (4.56)$$

If $\lambda_1 >> \lambda_2$ it follows from (4.56) that $\mathrm{tr}\left(\mathbf{F}^{\,r}\right) \approx \lambda_1^r$ such that

$$\lambda_1 \approx \left[\mathrm{tr}\left(\mathbf{F}^{\,r}\right)\right]^{\frac{1}{r}}. \tag{4.57}$$

The larger the difference between the largest and other eigenvalues, the better the approximation. Computation of (4.57) requires $r-1$ matrix multiplications, $N-1$ additions to calculate the trace and the calculation of the r^{th} root of a real number. If $\mathbf{F}$ could have negative eigenvalues, r should be odd in order to retain the correct eigenvalue sign and the minimum value is then $r = 3$. However, when $\mathbf{F}$ is positive-semidefinite (all eigenvalues are non-negative real) the minimum value is $r = 2$. Although the trace method will not produce accurate eigenvalues for the matrices considered in this paper, it can be used to produce a periodic eigenvalue sequence.

4.6.3 Eigen detection technique 2

Technique 2 calculates the frequency content of $\lambda_{Y,1}$ produced by technique 1. The complexity can be calculated by writing (4.48) as

$$T(\lambda_{Y,1}) = \frac{1}{V}\left[\left(\sum_{n=1}^{V}\lambda_{Y,1}(nT_s)f_{\cos}(nT_s)\right)^2 + \left(\sum_{n=1}^{V}\lambda_{Y,1}(nT_s)f_{\sin}(nT_s)\right)^2\right] \tag{4.58}$$

with $f_{\cos}$ and $f_{\sin}$ representing cosine and sine function values which can be calculated once and stored in memory. Each term within the square brackets of (4.58) requires $V+1$ multiplications and $V-1$ additions. The total complexity of (4.58) is therefore equivalent to $2V+3$ multiplications and $2V-1$ additions. To calculate the total processing time to produce a single value of $T(\lambda_{Y,1})$, the time required to produce the V samples of $\lambda_{Y,1}$ used in (4.58) should also be included.

4.7 Simulation study and results

The performances of the three detection techniques were evaluated through Monte Carlo computer simulation with the target communication system a baseband BPSK DSSS system employing a length $N = 64$ binary Walsh spreading code (taking row 32 of the Walsh matrix as code sequence). The intercept receiver used 1 sample per code chip such that $T_s = T_c$ to perform detection. The number of samples used to calculate a single test statistic value (for a single cyclic shift) for each detection technique was $N_s = N^2 = 64^2 = 4096$ corresponding to the size of the detection matrix. Since eigen technique 2 calculates the frequency content of the eigenvalue sequence and is not confined to the size of the detection matrix, the value $10N^2 = 40960$ was also considered for this technique. This section presents the simulated false alarm and detection performance results where

Table 4.2: Measured false alarm probability values.

Set P_{FA}	ED	Tech 1	Tech 2
0.001	0.0010205	0.0012686	0.0009245
0.003	0.0030093	0.0034905	0.0027469
0.005	0.0050614	0.0056225	0.0046470
0.010	0.0101522	0.0107803	0.0099211
0.030	0.0302207	0.0307656	0.0309257
0.050	0.0504181	0.0505104	0.0514241
0.100	0.1007815	0.0997421	0.1010004
0.300	0.3005873	0.2995871	0.2995768
0.500	0.5012642	0.5045618	0.4978770

10^6 data bits were considered per simulation run. A single simulation run consisted of calculating the probability of detection for a given SNR value. Measured average execution times for each technique are also presented.

4.7.1 Calculation of false alarm rates

The false alarm performance of each technique was measured by calculating the theoretical threshold ζ from set values of P_{FA} using the equations in Section 4.5. The false alarm or detection probability for the noise-only scenario was then obtained by setting $\sigma_x = 0$ and $\sigma_w = 1$ in (4.3). The results are given in Table 4.2. To obtain improved accuracies of these values or to measure even lower probabilities, larger values of N_s would be required. The measured values in Table 4.2 however correspond closely to the set values, confirming that the threshold equations in Section 4.5 are correct.

4.7.2 Probability of detection performance

Using the theoretical threshold values calculated from the set values of P_{FA} and setting $\sigma_w = 1$ in (4.3) and σ_x according to (4.4) for a given SNR value, simulated values for P_D were obtained. The simulated detection performances for $P_{FA} = 0.1$ and $P_{FA} = 10^{-6}$ are shown respectively in Figs. 4.6 and 4.7. At $P_{FA} = 10^{-6}$ technique 1 exhibits a maximum performance gain of approximately 3 dB over ED. Technique 2 exhibits additional gains over technique 1 of 1.5 dB for $V = N^2$ and 4 dB for $V = 10N^2$.

Fig. 4.7 shows that for $P_{FA} = 10^{-6}$, technique 1 achieves $P_D = 1$ at SNR $= -10$ dB. The SNR value at which $P_D = 1$ is reached increases with decreasing P_{FA}. Technique 1's theoretic limit for perfect detection (see Section 4.5.2) however predicts the value at

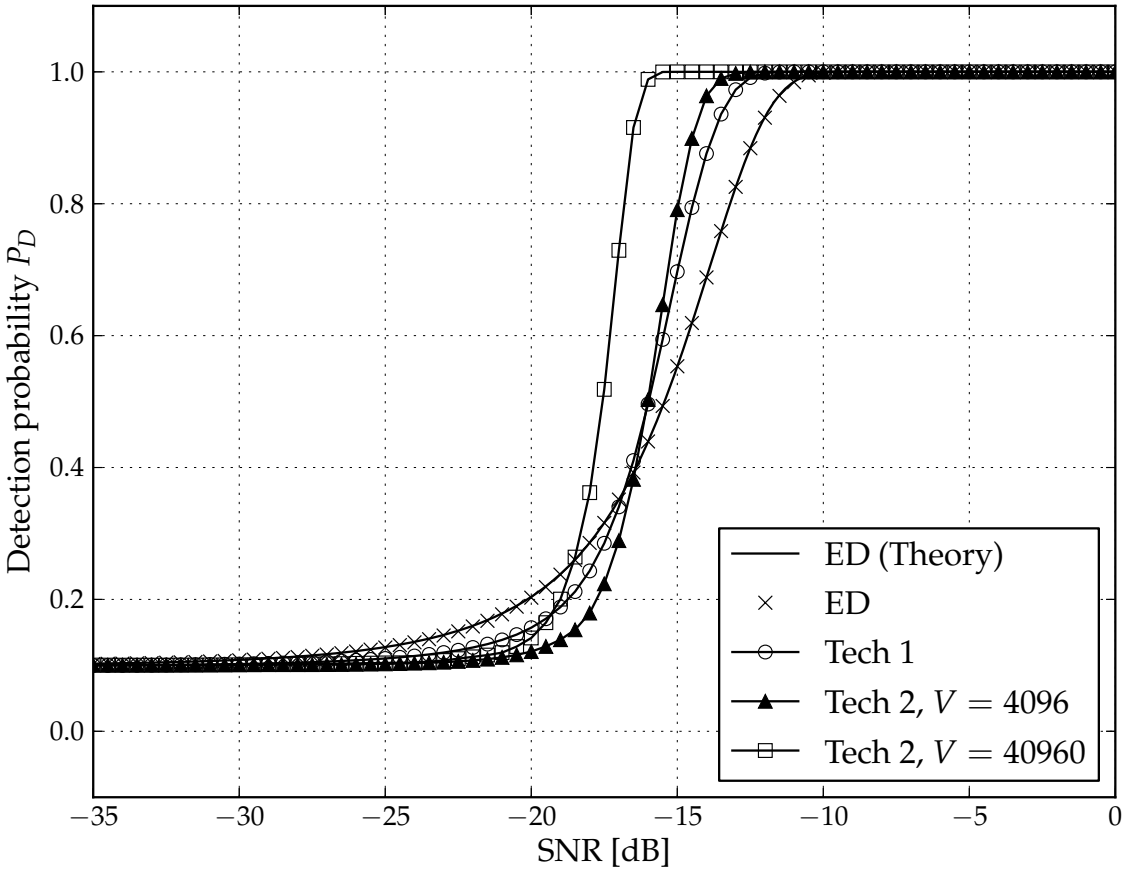


Figure 4.6: Simulated detection performance for $P_{FA} = 0.1$.

which $P_D = 1$ is reached will not increase beyond SNR $= -2.828$ dB.

The ROCs, displaying the P_D vs. P_{FA} performance, for SNR $= -14$ dB and the reference 45° line are shown in Fig. 4.8. For values of P_{FA} approaching one, the three techniques have the same detection performances, though for values of P_{FA} approaching zero both eigen detection techniques are superior to ED. At SNR $= -14$ dB, technique 2 ($V = 40960$) exhibits perfect detection. As the SNR value decreases, the ROC curves moves closer to the 45° reference line.

4.7.3 Evaluation of execution time

Although the number of arithmetic operations required to execute an algorithm (as derived in Section 4.6) predicts execution time, several other factors should also be considered. Two most important factors determining execution speed are hardware platform specifications (processing speed, memory size, etc.) and the efficiency of the implementation (how well the code is written to exploit the hardware).

The measured average execution times matching Table 4.1 are given in Table 4.3 for

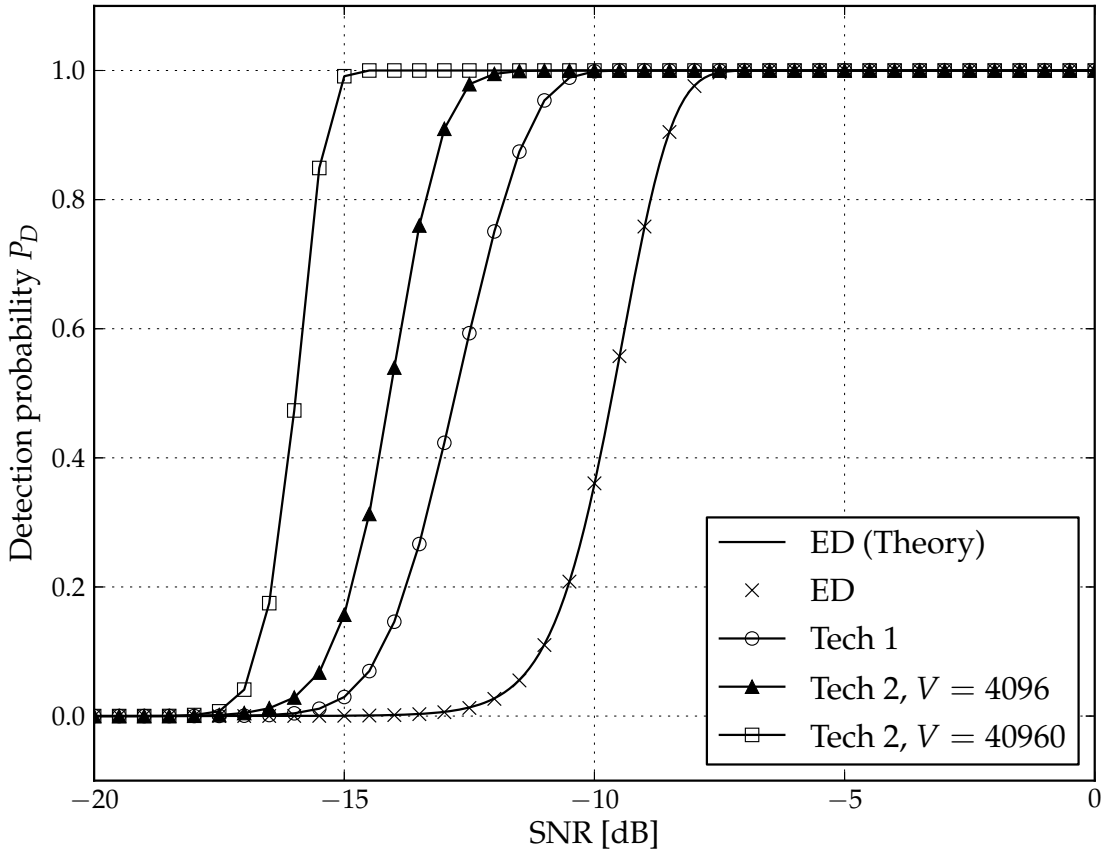


Figure 4.7: Simulated detection performance for $P_{FA} = 10^{-6}$.

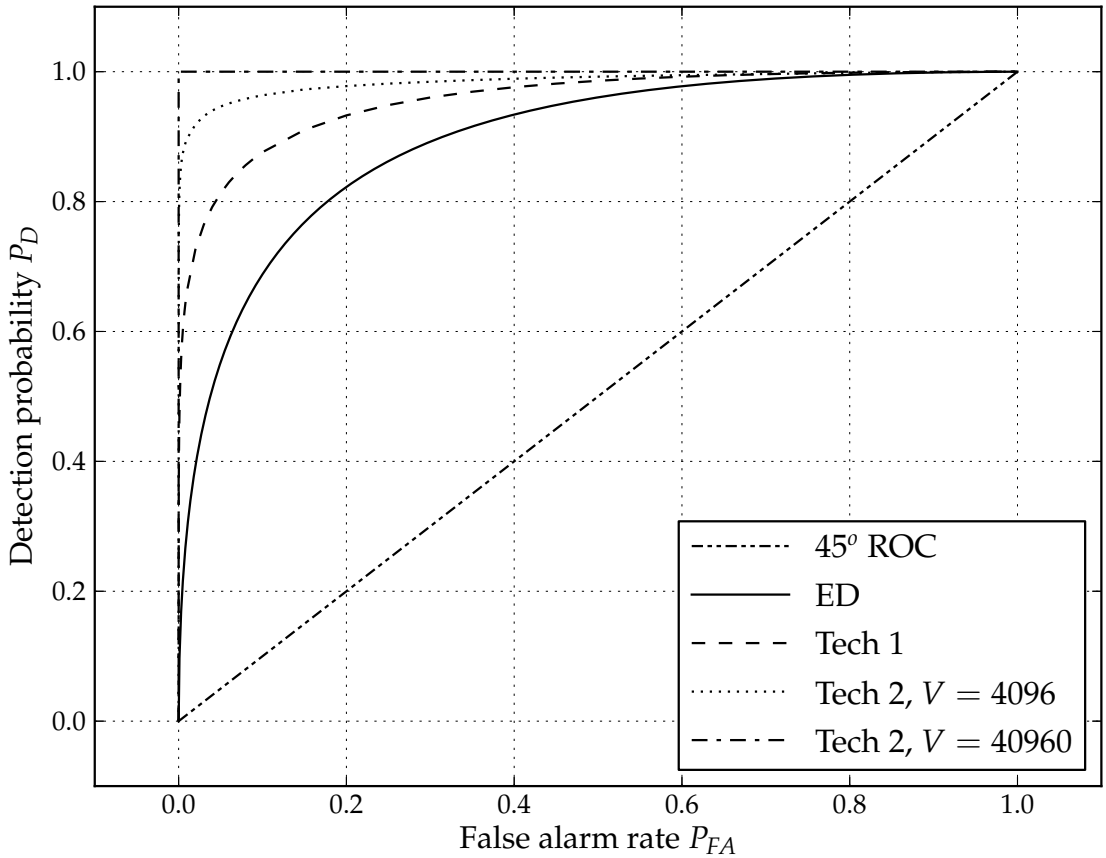


Figure 4.8: Simulated receiver operating curves for SNR = -14 dB.

Table 4.3: Average execution time in seconds of the detection algorithms for different values of N.

Algorithm	$N = 32$	$N = 64$	$N = 96$	$N = 128$
ED (C)	3.2928×10^{-6}	1.2870×10^{-5}	2.9118×10^{-5}	5.1695×10^{-5}
Tech 1 (C, -5 dB)	2.4120×10^{-4}	1.4825×10^{-3}	4.9252×10^{-3}	1.2397×10^{-2}
Tech 1 (GSL, -5 dB)	7.7575×10^{-5}	3.5204×10^{-4}	1.1936×10^{-3}	2.5524×10^{-3}
Tech 1 (C, -25 dB)	4.8001×10^{-4}	2.8274×10^{-3}	1.0530×10^{-2}	2.5837×10^{-2}
Tech 1 (GSL, -25 dB)	1.8533×10^{-4}	8.1181×10^{-4}	3.0881×10^{-3}	6.5832×10^{-3}
Tech 2 (C, $V = N^2$)	5.7688×10^{-6}	2.3089×10^{-5}	5.2298×10^{-5}	9.2508×10^{-5}
Tech 2 (C, $V = 10N^2$)	5.7705×10^{-5}	2.3179×10^{-4}	5.2073×10^{-4}	9.2447×10^{-4}

different code lengths N. The three techniques were implemented using ANSI C on a computer platform with an Intel Core 2 Duo T9600 (2.80 GHz) processor. The average execution time was measured by counting the number of processing cycles required by the section of C code that calculates the test statistic. The implementation of technique 1 was also enhanced using BLAS [127] through the C GSL [128].

For ED, the complexity only depends on N and the values in Table 4.3 clearly indicates that the execution time is related to N^2 as predicted in Table 4.1. The complexity of technique 1 depends both on N and the number of iterations M required by the symmetric power method. M is a function of SNR and the tolerance value ε; the power method converges slower for smaller values of both SNR and ε. The detection performance results presented in this paper can be obtained using $\varepsilon = 10^{-4}$ (smaller values do not improve performance) which is reached on average after $M = [11, 70, 111]$ iterations respectively for SNR $= [-5, -15, -25]$ dB.

Table 4.3 contains execution time results for technique 1 for SNR $= -5$ and -25 dB. The results of both the C and C with GSL implementations are shown. All four technique 1 cases indicate a growth in execution time greater than N^2. However, using GSL reduces the execution times and growth rates. Lastly, the execution times of technique 2 are strongly related to N^2 in each case (as predicted in Table 4.1). By increasing the window length N^2 by a factor 10, the execution times increase by the same factor.

Comparing the different techniques using Tables 4.1 and 4.3, it is clear that ED is the least complex. Although technique 2 is faster than technique 1, technique 2 needs the output sequence produced by technique 1.

4.8 Conclusion

This paper considered the detection of weak DSSS signals using PCA. Two new semi-blind detection techniques with superior performance to classic ED were presented. If the carrier frequency is known, the techniques only require the sequence length of the spreading code to construct a detection matrix which forms the basis of detection. It was shown that a DSSS signal can be detected using a function of the largest eigenvalue of the shifted detection matrix as test statistic. The performance improvement is due to the large difference between the largest eigenvalues of the detection matrix under the $\mathcal{H}_0$ and $\mathcal{H}_1$ hypotheses even at low SNR values. The computational complexity of each algorithm was also evaluated. Although this paper focussed on real signals, the techniques can also be applied to complex signals since the eigenvalues of the SCM of a complex detection matrix will also be real.

4.9 Acknowledgement

This work was supported by the Armaments Corporation of South Africa (Armscor) under contract no. KT521896. The authors would like to thank the anonymous reviewers for their valuable inputs.

CHAPTER 5

PUBLICATION 3

Title	Blind sequence-length estimation of low-SNR cyclostationary sequences
Authors	J.D. Vlok and J.C. Olivier
Journal	IET Communications
Publication date	12 June 2014
Volume	8
Issue	9
Pages	1578–1588
doi	10.1049/iet-com.2013.0616
Print ISSN	1751-8628
Online ISSN	1751-8636

The submission timeline is given below.

Original manuscript submitted	15 July 2013
Decisioned: Revision requested	25 November 2013
Revised manuscript submitted	23 December 2013
Decisioned: Accepted	17 February 2014
Publication	12 June 2014

Blind sequence-length estimation of low-SNR cyclostationary sequences

J.D. Vlok[1] and *J.C. Olivier*[2]

[1]Defence, Peace, Safety & Security (DPSS), Council for Scientific and Industrial Research (CSIR), Pretoria 0001, South Africa

[2]School of Engineering, University of Tasmania, Hobart 7005, Australia

E-mail: jvlok@csir.co.za

Abstract: Several existing DSSS detection and estimation algorithms assume prior knowledge of the symbol period or sequence length, although very few sequence-length estimation techniques are available in the literature. This paper presents two techniques to estimate the sequence length of a baseband DSSS signal affected by AWGN. The first technique is based on a known autocorrelation technique which is used as reference, and the second technique is based on PCA. Theoretical analysis and computer simulation show that the second technique can correctly estimate the sequence length at a lower SNR than the first technique. The techniques presented in this paper can estimate the sequence length blindly which can then be fed to semi-blind detection and estimation algorithms.

5.1 Introduction

In non-cooperative reception, blind detection and estimation techniques are required as the parameters used by the communication transmitter are in general not known by the receiver. The normal approach used in cooperative communication systems, such as optimal correlation or matched filtering techniques [26], is therefore not applicable in non-cooperative communication receiver systems. Typical applications of such systems are to be found in spectrum surveillance and electronic interception, but blind estimation techniques are important in their own right also [15], and hence this paper considers some aspects of this problem in depth.

Signal detection and parameter estimation are usually performed separately and independently in communication problems [129]. Detection is performed to determine whether the signal of interest is present or absent given observed data that is corrupted by noise. When it has been determined that the signal of interest is present, estimation is performed to determine the signal parameter values. This paper is concerned with blind estimation (assuming the signal is present) of the period of cyclostationary sequences used in DSSS communication systems.

The spreading codes used in DSSS are cyclostationary, since the mean and autocorrelation of the transmitted signal are periodic with the same period [6]. This periodicity distinguishes the signal from noise and can be exploited to perform detection and estimation especially when the signals are weak. This is in fact the case for received DSSS, especially for a non-cooperative receiver which does not know the spreading code and cannot therefore take advantage of the processing gain. Furthermore, modern communication systems use feedback to limit transmit power levels. The resultant effect is that intercepted DSSS signals have very low SNR levels and sophisticated algorithms are required to perform detection and estimation reliably.

We wish to emphasise that blind estimation of the sequence or code length (or symbol period) of hidden DSSS transmissions is essential since semi-blind techniques often assume knowledge of the sequence length which is generally not known *a priory*. Such semi-blind techniques include sequence estimation techniques [77, 86, 93] and detection algorithms [16].

A few techniques that may be used to estimate the sequence length (or related parameters from which the sequence length can be determined) of DSSS transmissions have been suggested in the literature. These include cyclic-feature analysis to determine the chip rate through spectral-line regeneration [4], the related Fourier analysis of cyclic correlation to estimate the bit rate [86], and higher-order statistical analysis where unique bispectrum and triple correlation patterns reveal characteristics of m-sequences that may be used to determine the sequence length and generator polynomials [111]. These techniques however require relatively high SNRs to perform parameter estimation. Another technique based on autocorrelation suggested in [54, 55] has the potential to estimate the sequence length at lower SNR values.

In this paper, we propose two new methods to estimate the sequence length of an intercepted DSSS signal at low SNR, and we present some novel results based on these new methods. The first method is based on the autocorrelation technique [54, 55] mentioned above, and the second method on a PCA detection technique [16]. The two techniques are compared in an AWGN environment in terms of the probability of correct estimation of the sequence length.

The paper is organised as follows. Section 5.2 presents the communication and intercept systems, and defines the SNR regime to which the estimation is applied. The two new methods are introduced in Sections 5.3 and 5.4. Section 5.5 presents numerical results, and Section 5.6 concludes the paper. Possible future research areas are identified in Section 5.7. A Barker code (length $N = 11$) and an m-sequence ($N = 63$) are considered as test cases throughout the paper.

5.2 Communication and intercept systems

The target communication system and intercept receiver platform used in this study are identical to those used in our previous study [16] and are briefly reviewed here. The target communication system is a BPSK DSSS system employing a length-N spreading code, such that the intercepted signal can be written as

$$\mathbf{y}(nT_c) = \sigma_x \mathbf{d}(nT_c)\mathbf{c}(nT_c) + \sigma_w \mathbf{w}(nT_c) \tag{5.1}$$

with the chip number $n = 1, 2, \ldots, N$ and T_c the chip interval. One sample is used to represent a single chip in the intercept receiver and therefore the sampling period $T_s = T_c$. $\mathbf{c}$ is the length-N ($N \gg 1$) pseudo-noise code sequence with period $T_{sym} = NT_c$ and $\mathbf{d}$ the data sequence assumed to be invariant over T_{sym}. Since the target communication system is a BPSK DSSS system, both $\mathbf{c}$ and $\mathbf{d}$ are sequences with values ± 1. The noise sequence is assumed to be a realisation of a standard normal RV represented by $\mathbf{w} \sim \mathcal{N}(0, 1)$ which contains i.i.d. samples. The code, data and noise sequences are also assumed independent of each other. The constants σ_x and σ_w are included to scale the signal and noise sequences respectively in order to obtain different SNR values, using

$$\text{SNR} = \frac{\sigma_x^2}{\sigma_w^2} \tag{5.2}$$

which is the SNR before despreading. The SNR at which the intercept receiver must be able to operate is dictated by the SNR required by the intended or target DSSS communication receiver system, which in turn is determined by the maximum tolerable bit error rate (BER). The error probability or BER and the SNR are related by [16] (see also Appendix A.3)

$$P_e = Q\left(\sqrt{N_s \text{ SNR}}\right) \tag{5.3}$$

with N_s the number of samples used to represent one transmitted bit. If one spreading sequence chip is represented by one sample in the receiver, then $N_s = N$. Fig. 5.1 shows the BER curves for the non-spread BPSK case ($N = 1$) and two spread cases ($N = 11$ and $N = 63$) under AWGN conditions. Spreading affords a processing gain $P_G = 10 \log_{10} N$ or SNR advantage to the intended DSSS receiver over the intercept receiver. For example, the receiver of a DSSS communication system using $N = 63$ ($P_G \approx 18$ dB) can despread a received signal at SNR $= -3$ dB to SNR $= 15$ dB to achieve $P_e \approx 10^{-8}$. An intercept receiver that does not know the spreading code will have to deal with the SNR $= -3$ dB signal, assuming the intended and intercept receivers lie on a circle with the target transmitter in the centre with an omnidirectional antenna. Identical channel conditions between the transmitter and each receiver are also assumed. If the transmitter uses a directional antenna and/or the intercept receiver is located at a further distance (which is typical in electronic interception scenarios), an even lower SNR will result.

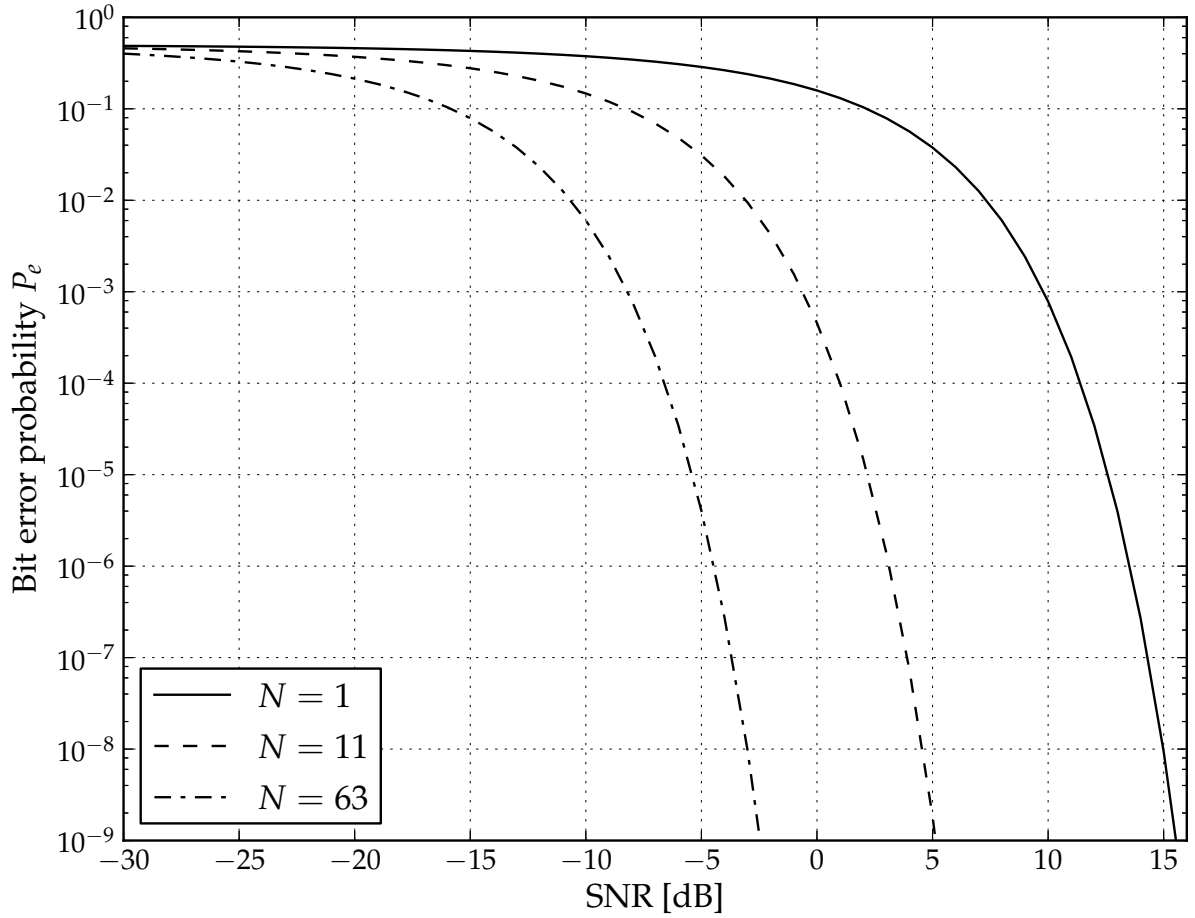


Figure 5.1: Bit error probability for non-spread ($N = 1$) and spread ($N = 11$ and $N = 63$) BPSK DSSS in AWGN.

To compete with the target DSSS communication system, the intercept receiver must be equipped with powerful techniques to detect the signals and estimate their parameter values at very low SNR levels. The BER that can be tolerated in a system depends on the application; it has been reported that in wireless multimedia transmission, voice packets can tolerate maximum BER levels of 10^{-3}, while data packets require a BER less than 10^{-9} [130]. Using Fig. 5.1, these BER values translate to minimum SNR levels of approximately -8.2 dB (-0.6 dB) for voice and -2.4 dB (5.1 dB) for data if $N = 63$ ($N = 11$) is used. The lower the SNR value at which a detection or estimation technique can function, the larger the detection range or intercept distance will be. In electronic intercept applications, SNR values less than the values given above will typically be required.

5.3 Estimation technique 1: Autocorrelation

It has been suggested that the time interval between autocorrelation spikes of an intercepted DSSS signal can be used to estimate the symbol period T_{sym} [54]. However, this section proposes a new estimation technique with detailed mathematical analysis, which is based on the concept that the correlation can be performed such that the index value of the first peak corresponds to the sequence length.

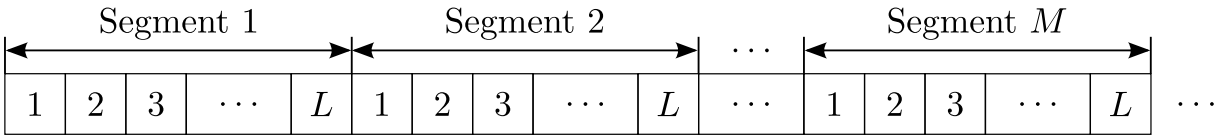


Figure 5.2: Segmented section of the intercepted signal consisting of ML samples.

5.3.1 Mean-square correlation

The intercepted signal of (5.1) can be expressed as

$$\mathbf{y} = \sigma_x d\mathbf{c} + \sigma_w \mathbf{w} \tag{5.4}$$

with the data bit value $d = \pm 1$ constant over a single spreading code. The first ML samples of $\mathbf{y}$ are split into M segments, such that each segment contains L samples as shown in Fig. 5.2.

A sliding correlation is then calculated between the m^{th} ($m = 1, 2, \ldots, M$) segment and the neighbouring section to the right within $\mathbf{y}$ using

$$R_{yy}^{(m)}(k) = \frac{1}{\sqrt{L}} \sum_{n=1}^{L} y_n \; y_{n+k} \tag{5.5}$$

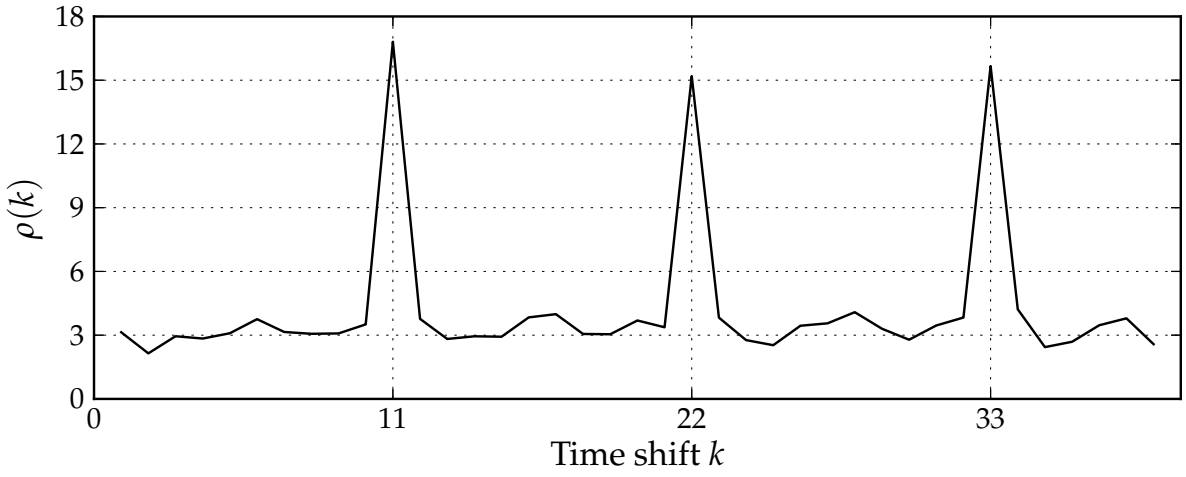
with the time-shift parameter $k = 1, 2, \ldots, K$ and the scale factor $\sqrt{L}$ chosen to simplify the mathematical analysis presented in Section 5.3.3. The samples of $\mathbf{y}$ in (5.5) are numbered such that y_n is the n^{th} sample of the m^{th} segment in each case. For example, when calculating $R_{yy}^{(1)}(k)$, sample y_{L+2} will refer to sample 2 of segment 2 in Fig. 5.2. However, when calculating $R_{yy}^{(2)}(k)$ this same sample will be referred to as y_2. The scalar product of a segment with itself is not needed and therefore $k = 0$ is excluded from (5.5). K is the value of the maximum time shift and determines the number of samples required beyond the segmented ML samples.

Using the M correlation sequences defined by (5.5), the mean-square correlation sequence can be calculated as

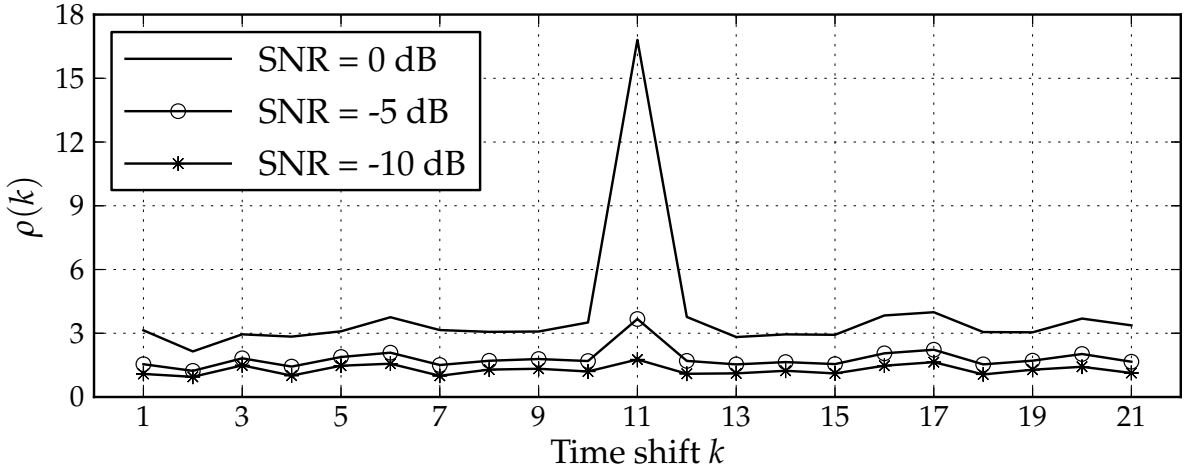
$$\rho(k) = \frac{1}{M} \sum_{m=1}^{M} \left[R_{yy}^{(m)}(k) \right]^2 \tag{5.6}$$

which is similar to the correlation estimators of [54]. An example mean-square correlation sequence for the Barker $N = 11$ spreading code is shown in Fig. 5.3(a) with parameters $K = 39$, $L = 11$, $M = 100$ and $\sigma_x^2 = \sigma_w^2 = 1$ such that the SNR is 0 dB.

The number of correlation spikes (or peaks) depends on the maximum time-shift parameter K, while the visibility of the peaks within the noise (or correlation sidelobe values) depends on the segment length L, number of segments M, the position of spreading se-



(a) $K = 39$ and SNR $= 0$ dB

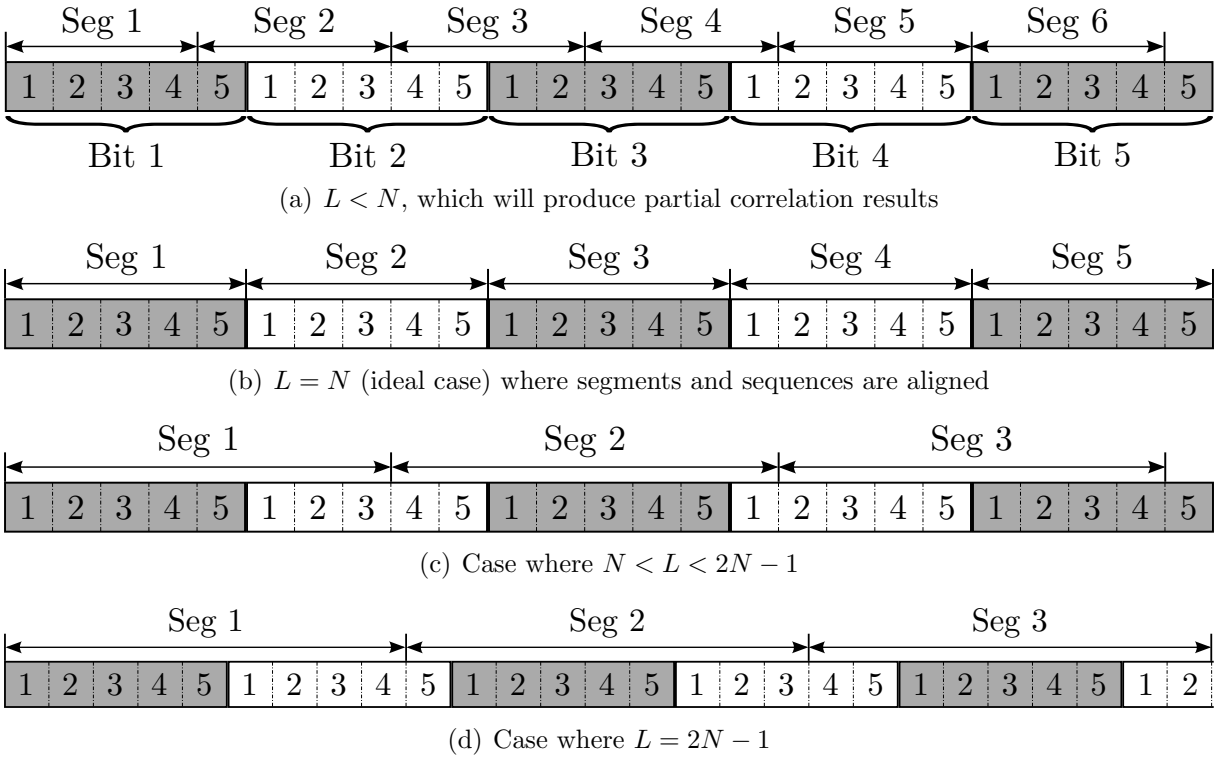


(b) $K = 21$ and SNR $= \{-10, -5, 0\}$ dB

Figure 5.3: Simulated mean-square correlation sequences for $L = N = 11$ and $M = 100$ for the Barker-11 code.

quences within segments, and the SNR. Assuming the spikes are detectable, there will be at least one in $\rho(k)$ if $K \geq N$, two if $K \geq 2N$, three if $K \geq 3N$, and so on. Fig. 5.3(a) shows three clear spikes since $K \geq 3N$ and the SNR of 0 dB is relatively high. For increasing values of M and SNR, the peak values generally become more visible as long as L is chosen correctly and a complete spreading code is located within a segment. Fig. 5.3(a) was obtained by choosing the actual spreading sequences as segments, which is the ideal case for which the correlation peaks will be maximum.

Due to partial correlation, the sum in (5.5) will produce smaller correlation peaks if $L < N$. Similarly, for $L > N$ partial correlation caused by incomplete spreading codes within each segment will add to or subtract from the correlation between complete spreading codes, depending on the data bit values. Also, if $L \neq N$ the spreading codes and segments are misaligned, such that the starting positions of each spreading code will differ in neigh-



(a) $L < N$, which will produce partial correlation results

(b) $L = N$ (ideal case) where segments and sequences are aligned

(c) Case where $N < L < 2N - 1$

(d) Case where $L = 2N - 1$

Figure 5.4: Example of a length $N = 5$ spreading sequence to illustrate the effect of the segment length L on the correlation process.

bouring segments. Even if $L > N$, some segments might not contain a single complete sequence and smaller correlation peaks than possible with $L = N$ will be formed. To ensure that a segment will always contain at least one complete spreading code, either $L = N$ (with sequences and segments aligned) or $L \geq 2N - 1$.

To illustrate the effects of the value of L and the alignment between segments and spreading sequences on the correlation process, an example is shown in Fig. 5.4. The sequence length $N = 5$, and the first sample of the first segment is also the starting position of a spreading code in each of the four cases shown. Each bit is represented by 5 samples (chips) and has a unique value (± 1) independent of the other bit values.

The case where $L < N$ is shown in Fig. 5.4(a) where the segments and sequences are misaligned. Using (5.5) will therefore result in partial correlation - and smaller peak values. The ideal case is shown in Fig. 5.4(b), where $L = N$ and the alignment is retained throughout the intercepted signal. When (5.5) is applied to this scenario, two complete spreading sequences will be correlated, producing a large peak for each value of m. The case where $N < L < 2N - 1$ is shown in Fig. 5.4(c). The first two segments each contains a complete sequence, which will provide a peak value when (5.5) is calculated. The incomplete sequences within these two segments will however affect this peak value depending on the bit values. Segment 3 in Fig. 5.4(c) does not contain a complete sequence

which will negatively affect the correlation process. When $L \geq 2N-1$ as illustrated in Fig. 5.4(d), all segments will however contain at least one complete spreading sequence. The incomplete sequences within each segment will also negatively affect the correlation peaks for this scenario, and the result will be worse compared with $L = N$ as in Fig. 5.4(b).

5.3.2 Method of estimation

Fig. 5.3(a) illustrates that if the parameters are chosen correctly and the SNR is sufficiently high, all correlation peaks will be located at index values k, which are equal to multiples of the code length N. For time shifts $k = aN$, with a a positive integer, the spreading code(s) within the m^{th} segment will align with spreading code(s) in $\mathbf{y}$, irrespective of the starting position of the spreading code(s) within the segment or $\mathbf{y}$. The correlation peak values will however depend on the alignment between segments and sequences as discussed in Section 5.3.1 above. The code length may therefore be determined from the index values of peaks within $\rho(k)$. In this paper we propose to use the index of the first peak ($a = 1$) to determine N.

5.3.3 Mathematical analysis

The intercepted signal of (5.4) can be written in discrete form as

$$y_n = \sigma_x d_i c_n + \sigma_w w_n \tag{5.7}$$

with $d_i = \pm 1$ the i^{th} data bit value and c_n the n^{th} ($n = 1, 2, \ldots, N$) chip of the spreading code. The ideal case corresponding to Fig. 5.4(b) will be analysed here and compared with simulation results of non-ideal cases in Section 5.5. The segment size will therefore be chosen as $L = N$, such that each segment in Fig. 5.2 will contain one complete spreading code. Furthermore, the maximum time shift is assumed to be bounded according to $N \leq K \leq 2N-1$, such that a single correlation peak is produced. Fig. 5.3(b) shows the resultant mean-square correlation sequence $\rho(k)$ for $K = 2N-1$ and SNR $= \{-10, -5, 0\}$ dB with $\sigma_w^2 = 1$. A clear peak value is shown at $k = N$ if the SNR is sufficiently high.

5.3.3.1 Signal-only analysis

This section will consider the characteristics of $R_{yy}^{(m)}(k)$ and $\rho(k)$ defined respectively in (5.5) and (5.6), using (5.7) with $\sigma_w = 0$.

Peak value

When $k = N$, the spreading codes within each segment align, such that (5.5) can be written using (5.7) as

$$\begin{aligned} R_{yy}^{(m)}(N) &= \frac{1}{\sqrt{L}} \sum_{n=1}^{L} y_n \; y_{n+N} \\ &= \frac{1}{\sqrt{L}} \sum_{n=1}^{L} \left(\sigma_x d_m c_n\right) \left(\sigma_x d_{m+1} c_{n+N}\right) \end{aligned} \tag{5.8}$$

where $d_m = \pm 1$ represents the data bit value associated with the m^{th} segment. Since $c_n = c_{n+N} = \pm 1$, (5.8) can be simplified as

$$R_{yy}^{(m)}(N) = \pm \sigma_x^2 \sqrt{L}. \tag{5.9}$$

The mean-square correlation peak value can then be expressed using (5.6) as

$$\rho(N) = \sigma_x^4 L. \tag{5.10}$$

Sidelobe values

When $k \neq N$, the correlation between misaligned spreading codes can be expressed using (5.5) and (5.7) as

$$R_{yy}^{(m)}(k) = \frac{\sigma_x^2}{\sqrt{L}} \sum_{n=1}^{L} \left(d_p c_n\right) \left(d_q c_{n+k}\right) \tag{5.11}$$

with d_p and d_q antipodal bit values depending on the segment in which c_n and c_{n+k} are respectively located. The correlation values produced by (5.11) depend on the specific spreading code $\mathbf{c}$ and the data bit values. If $d_p = d_q$, (5.11) resembles the periodic autocorrelation function, for which a number of bounds have been derived [131], which can be used to evaluate the sidelobe levels.

Barker codes and maximum-length or m-sequences are of particular interest, since the unscaled periodic autocorrelation for each time shift ($k \neq N$) is ± 1. Even when $d_p \neq d_q$, the sum in (5.11) equals ± 1 if $\mathbf{c}$ is a Barker code of length $N = [5, 7, 11, 13]$. In these conditions, (5.11) can be simplified to

$$R_{yy}^{(m)}(k) = \pm \frac{\sigma_x^2}{\sqrt{L}} \tag{5.12}$$

such that the mean-square value can be written by substituting (5.12) into (5.6) as

$$\rho(k) = \frac{\sigma_x^4}{L}. \tag{5.13}$$

Equation (5.12) will subsequently be used to derive the theoretical performance bound of technique 1.

5.3.3.2 Noise-only analysis

When $\sigma_x = 0$, the correlation sequence can be written by combining (5.5) and (5.7) as

$$R_{ww}^{(m)}(k) = \frac{\sigma_w^2}{\sqrt{L}} \sum_{n=1}^{L} w_n w_{n+k}. \tag{5.14}$$

Since $w_n \sim \mathcal{N}(0,1)$ is a sample within a sequence of i.i.d. samples, w_n and w_{n+k} will be independent when $k \neq 0$. The product $w_n w_{n+k}$ will therefore have a normal product distribution [132] with zero mean and unity variance (see Appendix H). According to the central limit theorem [110], the sum in (5.14) will approach the normal distribution with zero mean and variance approaching L, as L increases. $R_{ww}^{(m)}(k)$ will therefore approach the normal distribution with zero mean and variance

$$\begin{aligned} \text{var}\left(R_{ww}\right) &= \left(\frac{\sigma_w^2}{\sqrt{L}}\right)^2 L \\ &= \sigma_w^4. \end{aligned} \tag{5.15}$$

The distribution of the mean-square correlation $\rho(k)$ can then be obtained by combining (5.6), (5.14) and (5.15) as

$$\begin{aligned} \rho(k) &= \frac{1}{M} \sum_{m=1}^{M} \left[R_{ww}^{(m)}(k)\right]^2 \\ &= \frac{1}{M} \sum_{m=1}^{M} \left[\sigma_w^2 \tilde{R}_{ww}^{(m)}(k)\right]^2 \\ &= \frac{\sigma_w^4}{M} \sum_{m=1}^{M} \left[\tilde{R}_{ww}^{(m)}(k)\right]^2 \end{aligned} \tag{5.16}$$

with $\tilde{R}_{ww}^{(m)}(k)$ normalised such that it has unity variance. The mean-square correlation can then be written from (5.16) as

$$\frac{M}{\sigma_w^4}\rho(k) = \sum_{m=1}^{M} \left[\tilde{R}_{ww}^{(m)}(k)\right]^2 \tag{5.17}$$

which has a central Chi-squared distribution [5] with M degrees of freedom. Note that, although $\tilde{R}_{ww}^{(m)}(k)$ contains correlated samples over k for any given value of m, the sum in (5.17) is calculated over m (and not over k) such that the summands are independent.

It can be confirmed in simulation that (5.14) and (5.17) approach, respectively, a normal distribution (with variance given by (5.15)) and a central Chi-squared distribution (with M degrees of freedom) as L increases.

5.3.3.3 Signal-and-noise analysis

When a signal is present within the noise, the correlation can be expressed by combining (5.5) and (5.7) as

$$R_{yy}^{(m)}(k) = \frac{1}{\sqrt{L}} \sum_{n=1}^{L} \{(\sigma_x d_p c_n + \sigma_w w_n) \times (\sigma_x d_q c_{n+k} + \sigma_w w_{n+k})\} \tag{5.18}$$

with d_p and d_q bit values depending on the value of k as in (5.11). Equation (5.18) can further be developed as

$$R_{yy}^{(m)}(k) = \frac{\sigma_x^2}{\sqrt{L}} \sum_{n=1}^{L} d_p d_q c_n c_{n+k} + \frac{\sigma_x \sigma_w}{\sqrt{L}} \sum_{n=1}^{L} d_p c_n w_{n+k} + \frac{\sigma_x \sigma_w}{\sqrt{L}} \sum_{n=1}^{L} d_q c_{n+k} w_n + \frac{\sigma_w^2}{\sqrt{L}} \sum_{n=1}^{L} w_n w_{n+k}. \tag{5.19}$$

The first term of (5.19) is non-random while the remaining terms are random since they contain the noise sample $w \sim \mathcal{N}(0,1)$. The second and third terms of (5.19) are both normally distributed with zero mean and variance

$$\sigma^2 = \sigma_x^2 \sigma_w^2 \tag{5.20}$$

since $cd = \pm 1$ has no effect on the statistics of each term separately, and the sum of L i.i.d. standard normal samples has a variance equal to L. The last term of (5.19) is the same as the noise-only scenario described in (5.14), and is therefore also normally distributed with zero mean and variance σ_w^4 as in (5.15). By assuming that the terms of (5.19) are independent, $R_{yy}^{(m)}(k)$ can be described as a normally distributed RV with mean $\mu_R^{(m)}$ equal to the first term of (5.19), and variance given by

$$\sigma_R^2 = 2\ \sigma_x^2 \sigma_w^2 + \sigma_w^4 \tag{5.21}$$

which is the sum of the variances of the three random terms in (5.19). The independence assumption is based on the fact that $k \neq 0$, the noise samples are i.i.d., and the spreading chips and data bits are independent. (Independence can further be ensured using a dual-channel receiver structure as in [53].)

By substituting (5.19) into (5.6), $\rho(k)$ becomes the sum of squares of M non-zero mean Gaussian RVs. The distribution of $\rho(k)$ can therefore be determined using the non-central Chi-squared distribution $\mathcal{X}_M'^2$ [5]. By normalising the variance of $\rho(k)$ it can then be shown that

$$\frac{M}{\sigma_R^2} \rho(k) \sim \mathcal{X}_M'^2 \tag{5.22}$$

with non-centrality parameter

$$p_{nc} = \frac{1}{\sigma_R^2} \sum_{m=1}^{M} \left[\mu_R^{(m)} \right]^2 \tag{5.23}$$

since $R_{yy}^{(m)}(k)$ must be divided by σ_R to normalise the variance.

Peak value

When the spreading sequences align, the first term of (5.19) equals (5.8), and the mean value $\mu_R^{(m)}$ therefore equals (5.9). Using (5.9) as the mean value in (5.23), the non-centrality parameter can then be determined as

$$p_{nc} = \frac{\sigma_x^4 LM}{\sigma_R^2}. \tag{5.24}$$

It can be confirmed in simulation that the peak value $\rho(k = N)$ scaled according to (5.22) will have a non-central Chi-squared distribution with non-centrality parameter given in (5.24).

Sidelobe values

When the spreading sequences are misaligned, the first term of (5.19) equals (5.11), and the mean value $\mu_R^{(m)}$ therefore equals (5.12). Using (5.12) and (5.23), the non-centrality parameter can therefore be calculated as

$$p_{nc} = \frac{\sigma_x^4 M}{\sigma_R^2 L}. \tag{5.25}$$

It can also be confirmed in simulation that the sidelobe values $\rho(k \neq N)$ scaled according to (5.22) will have a non-central Chi-squared distribution with non-centrality parameter given in (5.25).

5.3.4 Estimation performance bound

Estimation algorithm 1 takes the index k of the maximum value of the mean-square correlation sequence $\rho(k)$ as the estimated sequence length N_{est}. The sequence length will therefore be estimated correctly if the peak of $\rho(k)$ is located at $k = N$. The estimation performance bound will be expressed in terms of the probability of correct estimation P_{ce}, which is the probability that the value $\rho(k)$ located at $k = N$ exceeds all other values

located at $k \neq N$, defined as

$$P_{\text{ce}} = p\left\{\rho(k = N) > \rho(k \neq N)\right\} \tag{5.26}$$

with $k = 1, 2, \ldots, K$ chosen such that a single distinct peak will be present within $\rho(k)$, assuming the SNR is sufficiently high as in Fig. 5.3(b). By defining the RVs $\rho_{\text{peak}} = \rho(k = N)$ and the largest sidelobe contender $\rho_{\text{max}} = \max[\rho(k \neq N)]$, the performance bound can further be developed from (5.26) as

$$\begin{aligned} P_{\text{ce}} &= p\left(\rho_{\text{peak}} - \rho_{\text{max}} > 0\right) \\ &= \int_0^\infty f_{\text{diff}}(z)\ dz \end{aligned} \tag{5.27}$$

with $f_{\text{diff}}(z)$ the PDF of the RV difference $\rho_{\text{peak}} - \rho_{\text{max}}$, which can be calculated using [110]

$$f_{\text{diff}}(z) = \int_{-\infty}^{\infty} f_{\text{peak}}(z + v)\ f_{\text{max}}(v)\ dv \tag{5.28}$$

which is simply the convolution between $f_{\text{peak}}(z)$ and $f_{\text{max}}(z)$, respectively the PDFs of ρ_{peak} and ρ_{max}. $f_{\text{peak}}(z)$ is the non-central Chi-squared PDF with non-centrality parameter given in (5.24), and $f_{\text{max}}(z)$ is the PDF of the maximum of the $K - 1$ sidelobe values, where each one has a non-central Chi-squared PDF with non-centrality parameter given in (5.25). $f_{\text{max}}(z)$ can therefore be expressed as [110]

$$f_{\text{max}}(z) = (K - 1) f_{\text{side}}(z)\left[F_{\text{side}}(z)\right]^{K-2} \tag{5.29}$$

with $f_{\text{side}}(z)$ and $F_{\text{side}}(z)$ respectively the PDF and CDF of each of the i.i.d. sidelobe values. By evaluating (5.27) to (5.29) numerically, the performance bound in terms of P_{ce} over a range of SNR values (with L and M fixed) can be obtained as is done in Section 5.5.

5.3.5 Choice of parameter values

It is important to note that the bound derived in Section 5.3.4 is the optimal estimation performance for technique 1. The bound is a function of the parameter values K (or range of k), L, M and the SNR, under the assumptions that the segments and sequences align as in Fig. 5.4(b), and that a single correlation peak is present as in Fig. 5.3(b).

The actual estimation performance depends on the choice or assumptions made regarding these parameter values. The range of k constrains the estimated sequence length N_{est}, and technique 1 can therefore only provide the correct answer as long as $k = N$ is considered within the range of k. The positions of segments within the intercepted signal and the

value of L will also influence the performance as described in Section 5.3.1 (see Fig. 5.4). The number of segments M required to estimate N depends on the SNR, although the number of segments available may be less than required, depending on the number of samples (containing the DSSS signal) intercepted. It can be shown that a lower SNR value will require more segments to maintain a given P_{ce} value, as more segments will be required to reduce or average out the noise.

Although the estimation technique presented here is blind, parameter values for K, L and M must be chosen correctly in order to determine N. The ranges of k and L may be set up according to known or expected DSSS sequence lengths or an exhaustive search may be required to find an autocorrelation peak. Although real-time application of the algorithm is possible in high-SNR scenarios, the typical low-SNR scenario considered in this paper will necessitate off-line analysis on a HPC platform as large values of M (and large ranges of k and L) would be required to perform estimation.

5.4 Estimation technique 2: Eigen analysis

Eigenvalues are used in several signal analysis techniques, including signal detection and parameter estimation. PCA [73] and SVD [75] are two such related approaches where the principal components (dominant eigenvectors) or singular values (square roots of eigenvalues) are extracted from a matrix constructed from the intercepted signal. An example PCA technique to estimate spreading sequences by concatenating the first two eigenvectors of the covariance matrix of the intercepted signal is presented in [91]. Several SVD methods used to estimate the parameters of sinusoids in noise are available in the literature. Examples include estimation of signal parameters via rotational invariance techniques (ESPRIT) [133] and matrix pencil algorithms [134], where generalised eigenvalues of matrix pencils are extracted to estimate the parameters of interest [135].

More recently, eigenvalue techniques have been suggested to perform spectrum sensing in cognitive radio applications [69]. The presence of a primary user can be detected by using test statistics based on eigenvalues of a fusion matrix constructed from signal samples collected cooperatively from distributed sensors [68]. A similar DSSS detection technique that uses the largest eigenvalue of the covariance matrix of the intercepted signal as test statistic is presented in [16].

In this section we wish to show that the detection technique described in [16] can be adapted to determine the sequence length of a hidden DSSS transmission. Subsequently, the portion of the technique presented in [16] which is required to develop the sequence estimation algorithm is reviewed, and the estimation technique itself is then presented.

5.4.1 Largest eigenvalue sequence

The detection technique of [16] consists of two stages. During the first stage, the baseband intercepted signal of (5.1) is divided into non-overlapping segments containing D samples each, and the segments are then stacked to form the $D \times D$ detection matrix given by

$$\mathbf{Y} = \sigma_x \mathbf{X} + \sigma_w \mathbf{W} \tag{5.30}$$

with $\mathbf{X}$ the data matrix (containing the spread data) and $\mathbf{W}$ the AWGN matrix with i.i.d. elements.

During the second stage, the largest eigenvalue $\lambda_{Y,1}$ of the SCM of $\mathbf{Y}$, denoted as [73]

$$\mathbf{R}(\mathbf{Y}) = \frac{\mathbf{Y}^T \mathbf{Y}}{D} \tag{5.31}$$

is calculated. A sequence of largest eigenvalues of $\mathbf{R}(\mathbf{Y})$ is then formed by cyclically shifting the elements of $\mathbf{Y}$ to the left and upwards, such that the first element in each row moves to the last element of the row above it. The top left element of $\mathbf{Y}$ is removed, and the lower right element takes on a new sample value. For each cyclic or time shift τ of $\mathbf{Y}$, the largest eigenvalue is calculated, to form a sequence $\lambda_{Y,1}(\tau)$.

Fig. 5.5 shows the largest eigenvalue sequences formed when the Barker-11 code is considered for different values of D, for the signal-only scenario ($\sigma_x = 1$ and $\sigma_w = 0$ in (5.30)). Clearly, when $D = N = 11$, the eigenvalue sequence exhibits a more regular pattern and has a larger mean (and variance) compared with the other values.

5.4.2 Method of estimation

The sequence length can be determined by analysing the mean of the largest eigenvalue sequence $\lambda_{Y,1}(\tau)$ over a range of values of D as shown in Fig. 5.6(a). Similar to Fig. 5.3(a), the mean of $\lambda_{Y,1}(\tau)$ exhibits peak values at integer multiples of the spreading sequence length N.

The sequence length can therefore be determined by identifying the index of the peak in the graph shown in Fig. 5.6(a). Since the mean value increases along with D, the peak will be located at N only if the SNR is sufficiently high. The peak value can therefore be determined more reliably by identifying a decrease (or negative slope) within the mean value graph across the range of D. The technique presented here is therefore to estimate the sequence length using the minimum value of the derivative of the mean, shown in Fig. 5.6(b).

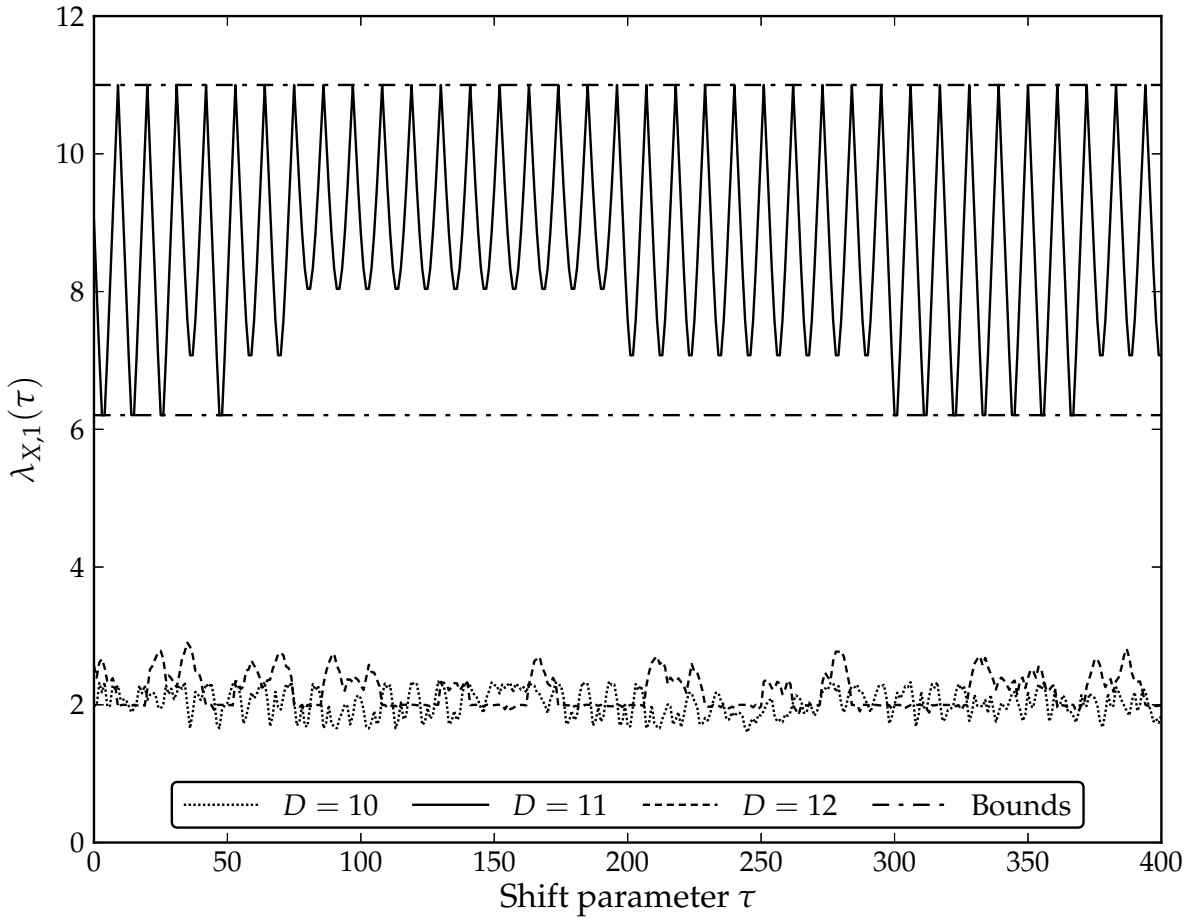


Figure 5.5: Simulated largest eigenvalue sequences for the Barker-11 code for the signal-only scenario.

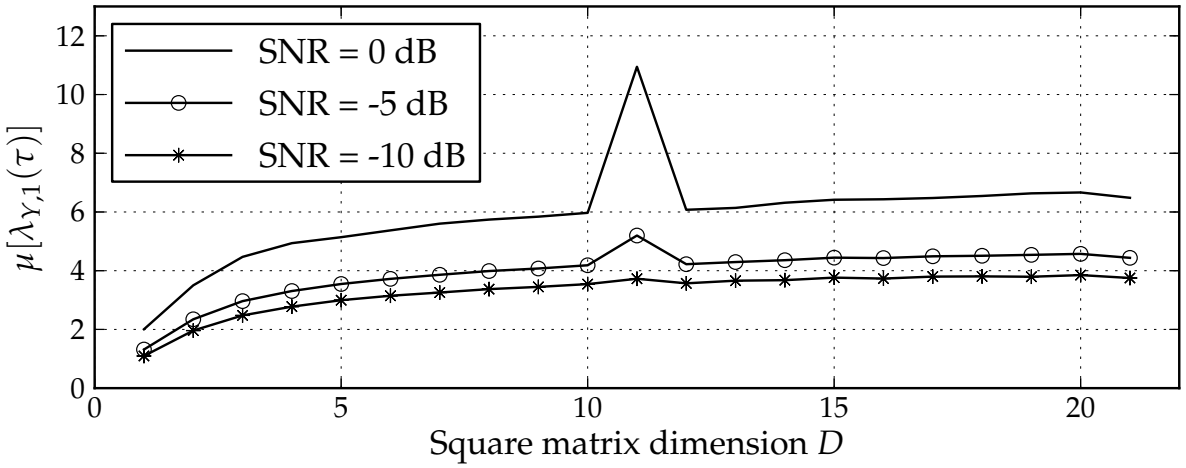
5.4.3 Mathematical analysis

Though it is possible to develop an analytic expression for the largest eigenvalue sequence, such an expression will be intractable since $\lambda_{X,1}$ depends on all the elements of $\mathbf{c}$ and $\mathbf{d}$ in (5.1). Instead, bounds on the variation of $\lambda_{Y,1}(\tau)$ will be developed in this section in order to describe its behaviour. The performance of the estimation technique will then be obtained through simulation which is presented in Section 5.5.

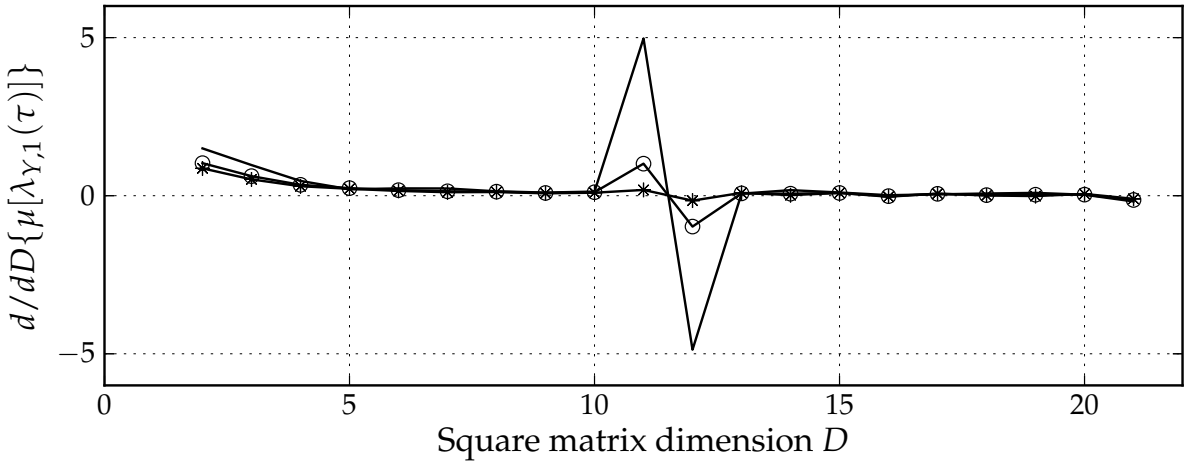
5.4.3.1 Signal-only analysis

If the square matrix dimension D matches the spreading code length N, and $\sigma_w = 0$ in (5.30), the SCM

$$\mathbf{R}(\mathbf{Y}) = \sigma_x^2 \mathbf{R}(\mathbf{X}) \tag{5.32}$$



(a) Mean of largest eigenvalue sequence



(b) Derivative of mean of largest eigenvalue sequence

Figure 5.6: Simulated functions of the largest eigenvalue sequence over square matrix dimension D for SNR $= \{-10, -5, 0\}$ dB.

from (5.31) has a maximum rank of two, such that its eigenvalues are the roots of a quadratic polynomial, which can be expressed in the form [16] (see also Appendix G)

$$\lambda_X = \frac{N^2 \pm \sqrt{\Delta}}{2N}\sigma_x^2 \tag{5.33}$$

since the spreading code and data bits have values ± 1. Also note that both eigenvalues will be non-negative, as the SCM $\mathbf{R}(\cdot)$ is positive semidefinite [5]. As illustrated in Fig. 5.5 for $D = N$, $\lambda_{X,1}(\tau)$ exhibits a pattern with period N as $\mathbf{X}$ is cyclically shifted. By considering all possible combinations of code and data values for any cyclic shift of $\mathbf{X}$, it can be shown that the discriminant in (5.33) has ranges

$$\Delta \in \begin{cases} [0, N^4] & (N \text{ even}) \\ \\ [2N^2 - 1, N^4] & (N \text{ odd}) \end{cases} \tag{5.34}$$

and the largest eigenvalue sequence is therefore bounded according to

$$\lambda_{X,1} \in \begin{cases} \left[\frac{N}{2}, N\right] \sigma_x^2 & (N \text{ even}) \\ \left[\frac{N^2+\sqrt{2N^2-1}}{2N}, N\right] \sigma_x^2 & (N \text{ odd}) \end{cases} \tag{5.35}$$

which are the bounds shown (for N odd) in Fig. 5.5.

5.4.3.2 Noise-only analysis

If $\sigma_x = 0$ in (5.30) such that $\mathbf{Y} = \sigma_w \mathbf{W}$, the SCM

$$\mathbf{R}(\mathbf{Y}) = \sigma_w^2 \mathbf{R}(\mathbf{W}) \tag{5.36}$$

from (5.31) is a Wishart matrix [102]. The distribution of the largest eigenvalue of a Wishart matrix can be described using the TW law [104], which can be approximated using the Gamma distribution [76]. Using functions of distributions [110], the distribution of the largest eigenvalue $\lambda_{W,1}$ of (5.36) can be expressed using the Gamma PDF, given by [76]

$$\gamma(z) = \frac{(z - z_0)^{\alpha-1}}{\theta^\alpha \Gamma(\alpha)} \exp\left[\frac{-(z - z_0)}{\theta}\right] \tag{5.37}$$

with $\Gamma(\cdot)$ the Gamma function, z_0 the location parameter, α the shape, and θ the scale, respectively given by

$$\begin{aligned} z_0 &= \frac{\sigma_w^2 \left(\mu_c - 9.8209\ \sigma_c\right)}{D} \\ \alpha &= 46.5651 \\ \theta &= \frac{0.1850\ \sigma_w^2\ \sigma_c}{D} \end{aligned}$$

with centre and scaling parameters [102]

$$\begin{aligned} \mu_c &= \left(\sqrt{D-1} + \sqrt{D}\right)^2 \\ \sigma_c &= \sqrt{\mu_c} \left(\frac{1}{\sqrt{D-1}} + \frac{1}{\sqrt{D}}\right)^{\frac{1}{3}}. \end{aligned}$$

Furthermore, the support region of (5.37) can be expressed as [76]

$$z \in [z_0, z_0 + 2\alpha\theta]. \tag{5.38}$$

5.4.3.3 Signal-and-noise analysis

When both $\sigma_x > 0$ and $\sigma_w > 0$ in (5.30), the SCM of (5.31) can be written as

$$\mathbf{R}(\mathbf{Y}) = \sigma_x^2 \mathbf{R}(\mathbf{X}) + \sigma_w^2 \mathbf{R}(\mathbf{W}) + \mathbf{E} \tag{5.39}$$

with the cross-term or error matrix

$$\mathbf{E} = \frac{\sigma_x \sigma_w}{D} \left[\mathbf{X}^T \mathbf{W} + \mathbf{W}^T \mathbf{X} \right] \tag{5.40}$$

which is zero under the assumption that the signal and noise are uncorrelated. Under this assumption, $\mathbf{R}(\mathbf{Y})$ is a linear function of $\mathbf{R}(\mathbf{X})$ and $\mathbf{R}(\mathbf{W})$ as indicated by (5.39), although the eigenvalues of $\mathbf{R}(\mathbf{Y})$ are nonlinear functions of the eigenvalues of $\mathbf{R}(\mathbf{X})$ and $\mathbf{R}(\mathbf{W})$ [113]. According to the Weyl inequalities [113–115], the largest eigenvalue of $\mathbf{R}(\mathbf{Y})$ is however bounded according to

$$\lambda_{\min} \leq \lambda_{Y,1} \leq \lambda_{\max} \tag{5.41}$$

with the upper and lower bounds given by [16]

$$\lambda_{\max} = [\lambda_{X,1}]_{\max} + [\lambda_{W,1}]_{\max} \tag{5.42}$$

$$\lambda_{\min} = \max \left\{ [\lambda_{X,1}]_{\min}, [\lambda_{W,1}]_{\min} \right\}. \tag{5.43}$$

Note that both $[\lambda_{X,1}]_{\min}$ and $[\lambda_{W,1}]_{\min}$ are less than or equal to $\lambda_{Y,1}$, though the tightest lower bound is obtained by taking the maximum of the two. The upper bound can be written from (5.42) as [76]

$$\lambda_{\max} = N\sigma_x^2 + z_0 + 2\alpha\theta \tag{5.44}$$

using the upper bounds given in (5.35) and (5.38). The lower bound can similarly be obtained from (5.43) using the lower bounds of (5.35) and (5.38). For N odd, the lower bound can therefore be expressed as

$$\lambda_{\min} = \max \left\{ \frac{N^2 + \sqrt{2N^2 - 1}}{2N} \sigma_x^2, z_0 \right\} \tag{5.45}$$

from which it can easily be shown that $\lambda_{\min} = z_0$ for small SNR values.

A simulated example of the largest eigenvalue sequence for $\sigma_x = \sigma_w = 1$ (SNR of 0 dB) is shown in Fig. 5.7 for the Barker-11 code for different values of D similar to Fig. 5.5. The eigenvalue bounds given in (5.44) and (5.45), and the measured mean value λ_μ of $\lambda_{Y,1}(\tau)$ are also shown. When $D = N$, the mean λ_μ clearly exceeds the means of the eigenvalue sequences associated with $D \neq N$, which is also illustrated in Fig. 5.6(a).

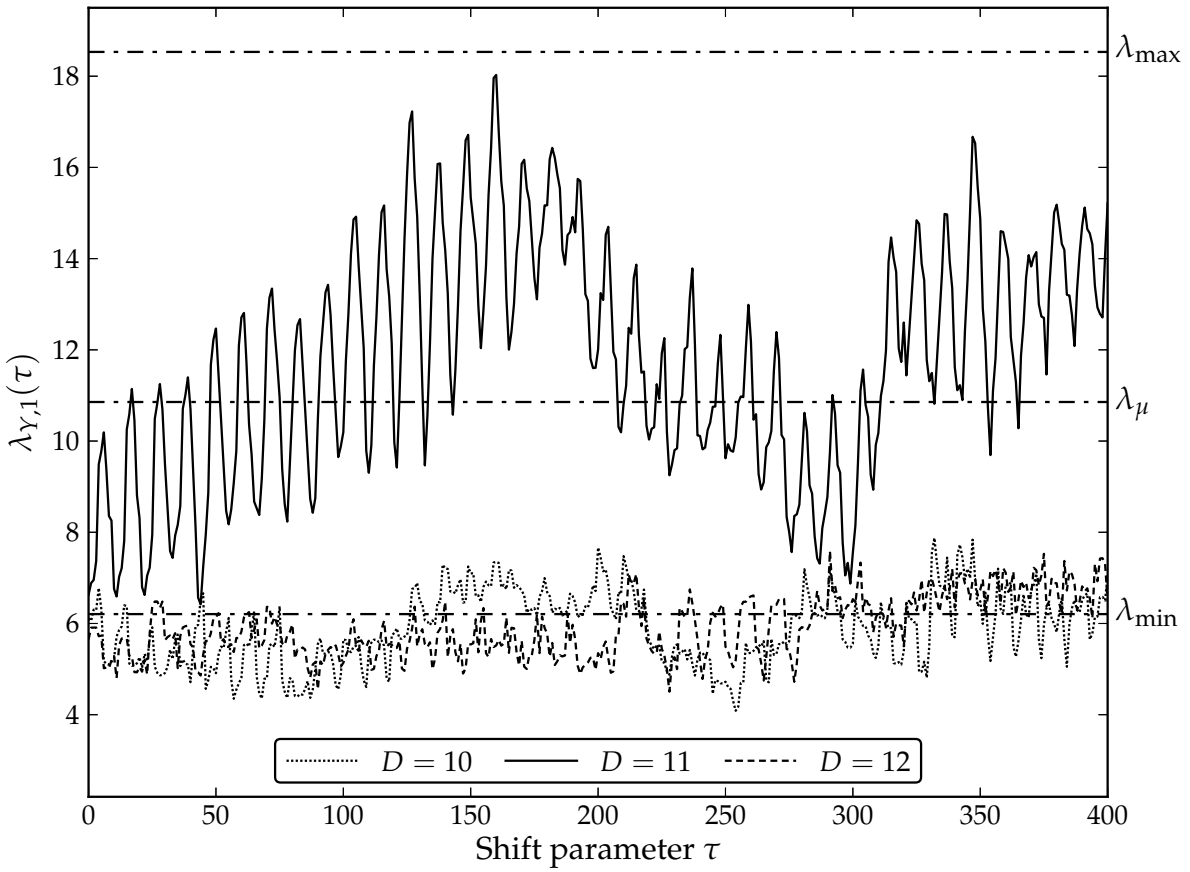


Figure 5.7: Simulated largest eigenvalue sequences for the Barker-11 code for the signal-and-noise scenario.

5.4.4 Choice of parameter values

The discussion of Section 5.3.5 is also relevant to technique 2. The range over which the matrix dimension D should be evaluated and the number of matrices (maximum time shift τ) to be considered to calculate the average eigenvalue must be chosen, similar to the suggestions made for the parameters of technique 1. As the calculation of eigenvalues are computationally expensive [16], technique 2 will also typically require off-line analysis.

5.5 Simulation results

This section provides Monte Carlo simulation results obtained by implementing the communication and intercept models of Section 5.2 in software. The performance of the two estimation techniques presented in Sections 5.3 and 5.4 were evaluated against a Barker code ($N = 11$) and m-sequence ($N = 63$) with generator polynomial $g(X) = X^6 + X + 1$. The output of each estimation technique is the estimated sequence length N_{est}, which is compared with the actual sequence length N in order to evaluate the performance of each estimation technique.

5.5.1 Probability of estimation

Fig. 5.8 shows the probability of estimation P_{est} obtained over a range of values for N_{est} when the Barker-11 code is considered. Fig. 5.8(a) shows the results obtained using technique 1 for segment lengths $L = \{5, 11\}$ and SNR $= \{-15, -12\}$ dB. The sidelobe values (for $N_{\text{est}} \neq N$) are shown to be uniformly distributed, while a peak is observable (depending on L and the SNR) at $N_{\text{est}} = N$. The peak value increases while the sidelobe values decrease, as the SNR increases. Furthermore, the value at $N_{\text{est}} = N$ is the highest for $L = N$, provided that the SNR is sufficiently high.

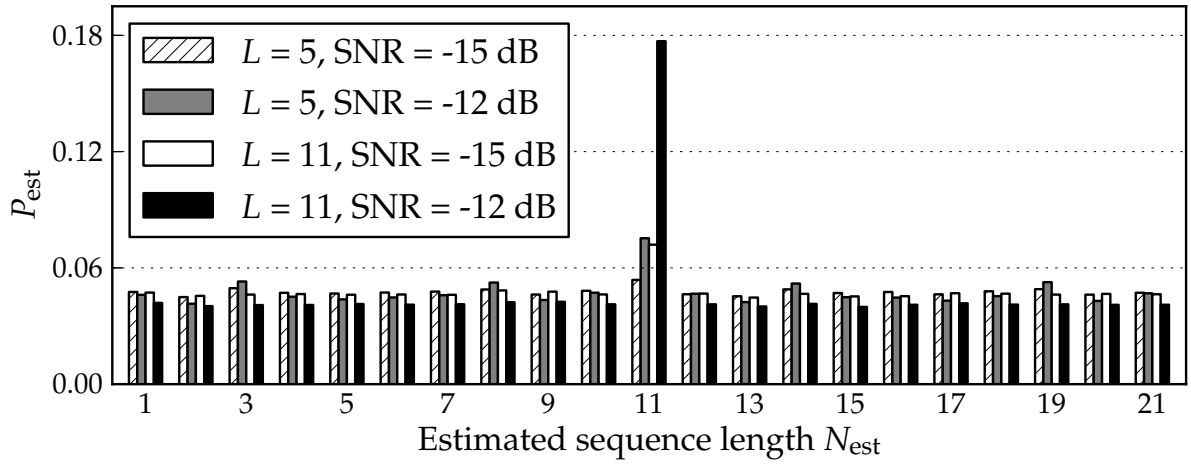
Similarly, Fig. 5.8(b) shows the results obtained using technique 2 for different SNR values. The distribution of the sidelobe values can be explained from Fig. 5.6; the derivative of the mean of $\lambda_{Y,1}(\tau)$ is positive for small values of the square matrix dimension D, and decreases as D increases. The minimum value of the derivative will therefore typically (depending on the SNR) be located at $N_{\text{est}} \geq N$. As the SNR increases, the probability that $N_{\text{est}} = N$ will also increase.

5.5.2 Probability of correct estimation

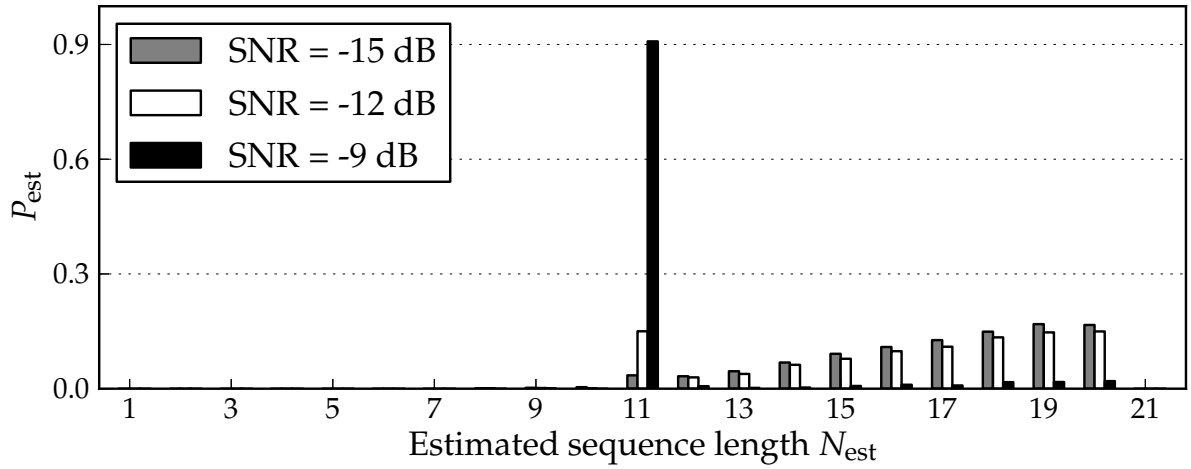
As discussed in Section 5.3.4, the performance of the estimation techniques can be evaluated in terms of the probability of correct estimation P_{ce}, which is the probability that N_{est} will equal N. Fig. 5.9 shows the simulated P_{ce} over a range of SNR values for the two estimation techniques against the Barker-11 code.

For technique 1, 10^5 runs of $M = 1000$ segments (of length L each) were simulated for each SNR point shown, with the maximum time shift $K = 2L - 1$. Different segment sizes L were considered, and when L matches $N = 11$, the best performance is obtained. When $L < N$, the performance degrades due to partial correlation of incomplete spreading sequences as explained in Section 5.3.1. When $L > N$, the performance improves since some segments contain complete spreading codes, though the performance is still worse compared with $L = N$, since fragments of spreading codes within segments reduce the correlation peaks (also due to partial correlation). The theoretical bound for technique 1 given by (5.27) is also shown, which is nearly attained by the $L = 11$ curve. (The $L = 11$ curve does not match the theoretical curve *exactly* since L or N is not sufficiently long.)

For technique 2, the mean of the largest eigenvalue sequence λ_μ was calculated by shifting 1000 bits through the detection matrix $\mathbf{Y}$ as explained in Section 5.4. The same data sequence was used to calculate λ_μ for a single simulation run, during which the range of matrix dimensions $D = 1, \ldots, 2N - 1$ were evaluated for each SNR point. A total of 10^4 simulation runs was completed per SNR value. Fig. 5.9 indicates that technique



(a) Results obtained using technique 1



(b) Results obtained using technique 2

Figure 5.8: Normalised histograms to indicate P_{est} for both estimation techniques against the Barker-11 code.

2 outperforms the best possible performance of technique 1 for SNR values exceeding approximately -11 dB. When the segment size of technique 1 is chosen as $L = 5$, technique 2 outperforms technique 1 by up to 4 dB.

Similar to Fig. 5.9, the simulated P_{ce} performances of the two estimation techniques against the length-63 m-sequence are shown in Fig. 5.10. The theoretical performance bound for technique 1 predicted in (5.27) is approached for $L = N$, and worse performances are shown for both $L < N$ and $L > N$ as in Fig. 5.9. Furthermore, technique 2 exceeds the best possible performance of technique 1 by approximately 4 dB. By comparing Figs. 5.9 and 5.10, it is clear that the estimation performance improves for a larger value of N.

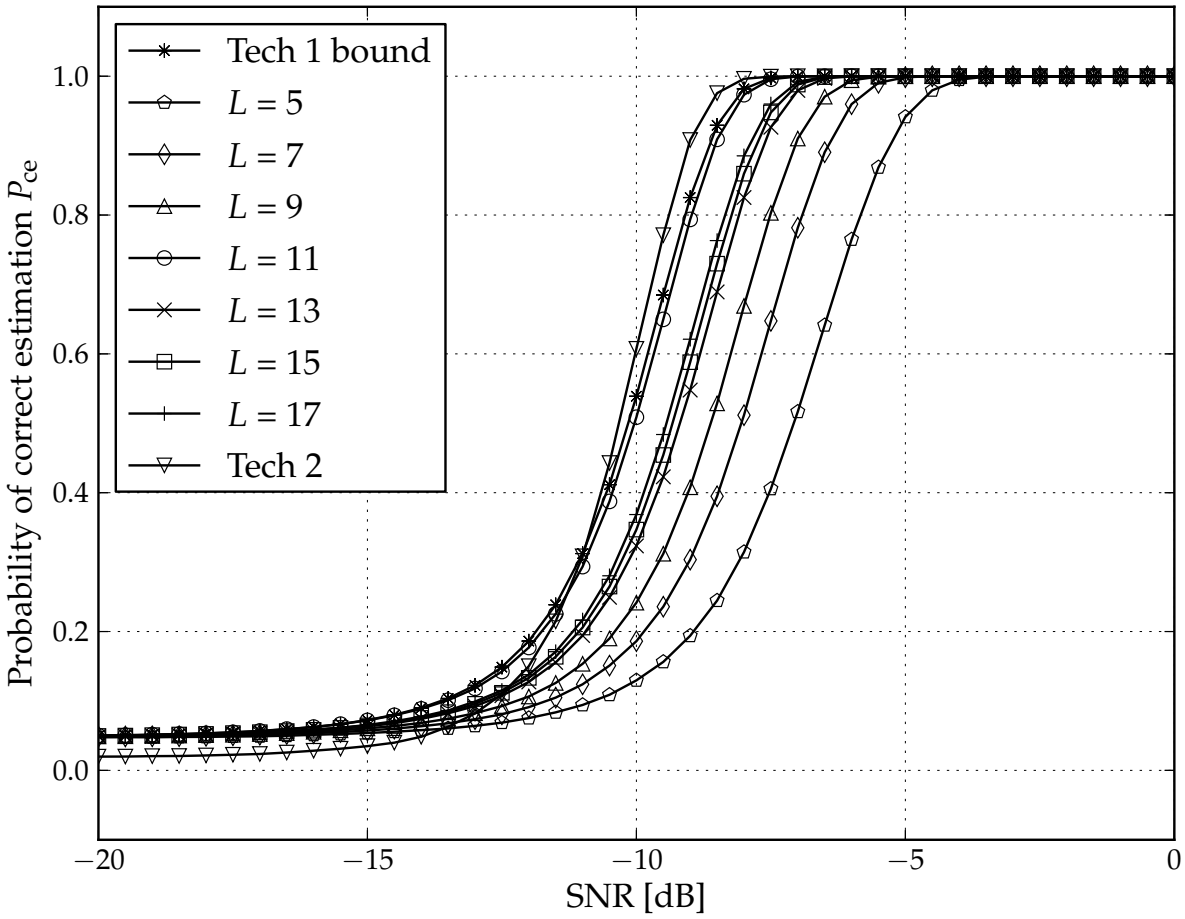


Figure 5.9: Estimation performance of the two techniques against the Barker-11 code.

Figs. 5.9 and 5.10 also indicate that both techniques 1 and 2 will be able to correctly estimate the sequence lengths (with $P_{ce} = 1$) of DSSS transmissions at the SNR levels given in terms of the maximum tolerable BER and channel requirements described in Section 5.2.

5.6 Conclusion

Two novel techniques based on autocorrelation and PCA were presented to blindly estimate the sequence length N of an intercepted DSSS transmission hidden within noise. Mathematical analyses and results of a simulation study for each technique were given.

The autocorrelation technique computes the mean-square correlation between segments of the intercepted signal and takes the index value of the first peak as the sequence length. The performance of this technique depends on the choice of the segment length L, with best performance if $L = N$. Furthermore, the range of k over which the correlation peak is searched must include $k = N$ and the number of segments M must be sufficient to

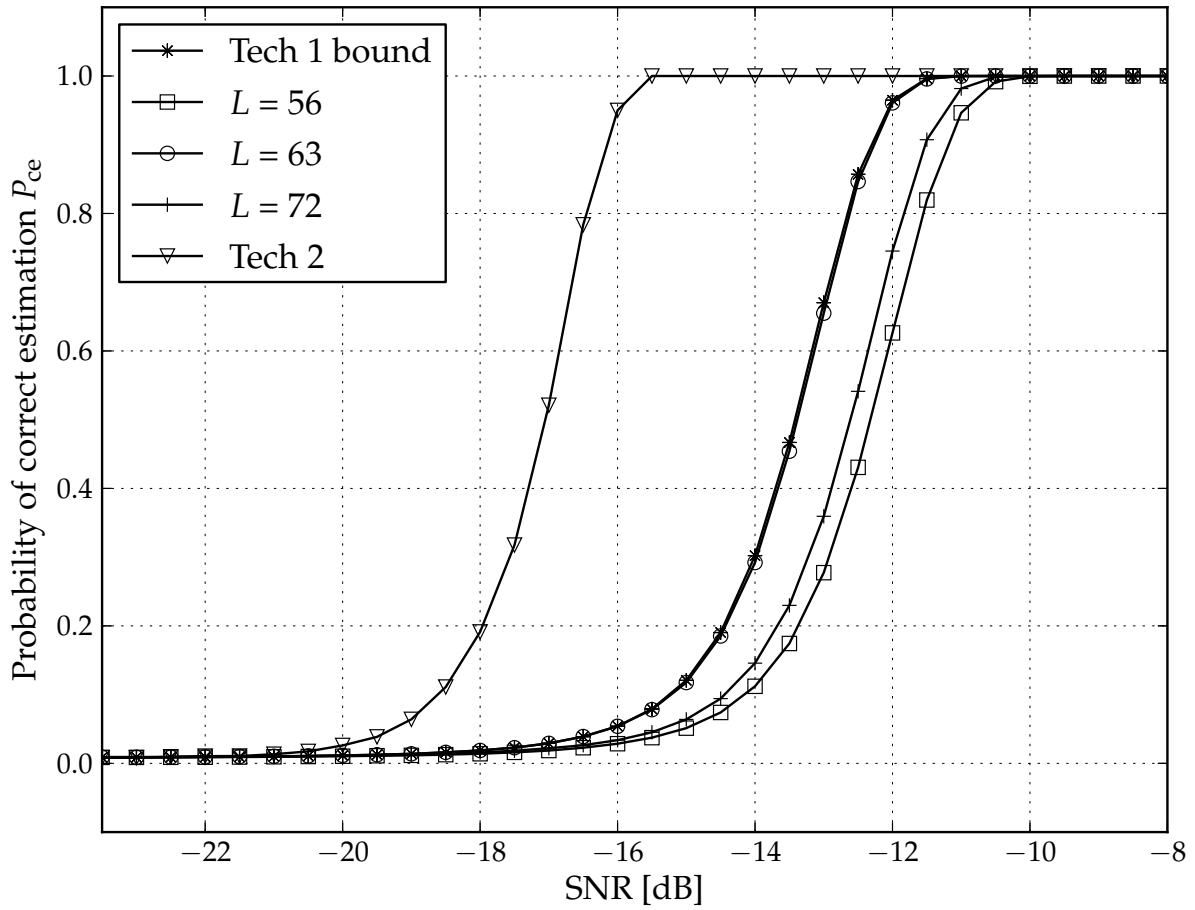


Figure 5.10: Estimation performance of the two techniques against the length-63 m-sequence.

suppress the noise. This technique is similar to the correlation spike spacing estimation algorithm suggested in [54], and was used to establish a reference performance in terms of probability of correct estimation P_{ce} over SNR.

The eigen-analysis technique computes the mean value λ_μ of the largest eigenvalue sequence of the intercepted signal for a range of square data matrix dimensions D. It was shown in this paper that λ_μ is much larger when D equals the spreading sequence length N. The sequence length can therefore be determined as the value of D at which λ_μ versus D has a peak value. The eigen-analysis technique, adapted from the detection technique of [16], was also shown to have superior estimation performance compared with the reference autocorrelation technique.

5.7 Future work

In this paper it was assumed that the signal of interest was detected first such that the estimation algorithm was applied to a signal that is surely present within noise. However,

it cannot always be ascertained that the signal of interest *is* present before attempting to perform estimation. In certain applications, detection and estimation can be formulated as a single problem which can provide improved results [136]. The estimation techniques presented in this paper can be adapted to function as detection algorithms by using the estimated parameters as detection test statistics. A confidence level that the sequence length is estimated correctly can also be established by evaluating the consistency of the estimated value.

Furthermore, only spreading codes with noise-like autocorrelation characteristics, including Barker codes and m-sequences, were considered in this paper. How to blindly estimate the sequence length of orthogonal codes, such as Walsh codes which have multiple sidelobe correlation peaks, remains an open question.

5.8 Acknowledgement

This work was supported by the Armaments Corporation of South Africa (Armscor) under contract no. KT521909. The authors would like to thank the anonymous reviewers for their valuable inputs.

CHAPTER 6

DISCUSSION

This chapter presents a summary and considers the major findings of the work presented in this thesis. Conclusions drawn from studying the existing detection and estimation techniques, developing new techniques and evaluating the simulated performance results are included. The challenge and importance of the research are highlighted, and ethical concerns are addressed. The work is also critically evaluated and possible areas for future investigation are briefly explored.

6.1 Summary and major findings of publications

A short summary and major findings of the research work presented in the three publications (Chapters 3 to 5) are provided below.

6.1.1 Publication 1

The first publication presented analytic expressions that approximate the distribution of the largest eigenvalue of white Wishart matrices, which can be used to predict the performance of eigenvalue-based detection techniques in AWGN. It was shown that the desired analytic expression follows from an approximation to the TW distribution in terms of the Gamma distribution. The parameters of the analytic expression were found by scaling and shifting the Gamma distribution, such that the difference between the Gamma and TW distribution was minimised. The approximation offers largely simplified computation and also provides statistics, including the mean value and support region, of the largest eigenvalue distribution. The approximation is given in simple closed-form equations that do not require numerical evaluation.

Publication 1 illustrated (through Monte-Carlo computer simulation and analysis of the distributions) that the TW and associated distributions (which can be used to predict the performance of eigenvalue-based techniques) can be accurately approximated using a simplified Gamma distribution.

6.1.2 Publication 2

The second publication presented a semi-blind DSSS detection technique based on eigen analysis. The received signal is segmented and the samples packed in a square detection matrix, which is cyclically shifted when a new sample is received. The largest eigenvalue is calculated for each cyclic shift of the detection matrix. Test statistics based on the resulting eigenvalue sequence (including the amplitude and frequency content) are then used to make a decision (using binary hypothesis testing) as to whether the DSSS signal is present or not. The eigenvalue-based test statistics show large differences under the $\mathcal{H}_0$ (signal absent) and $\mathcal{H}_1$ (signal present) hypotheses. The two hypotheses can thus easily be distinguished, such that detection in low-SNR conditions is made possible.

Publication 2 illustrated that detection of DSSS signals is possible at low-SNR values, if the spreading sequence length is known. Detection performance improvements of up to 7 dB were shown to be achievable (using Monte-Carlo computer simulation) when comparing the new algorithms with classic ED. However, achieving detection in lower SNR scenarios comes at a price of increased computational complexity.

6.1.3 Publication 3

The third publication presented two techniques to blindly estimate the sequence length of a cyclostationary sequence, such as a DSSS signal. The first technique splits the received signal into segments and then calculates the mean-square correlation between the segments, such that the first correlation peak is located at the index value corresponding to the sequence length. The second technique uses the eigenvalue test statistic presented in Publication 2 to estimate the sequence length. The mean of the eigenvalue sequence is calculated over a range of the dimensions of the square detection matrix. The matrix dimension that exhibits the maximum mean eigenvalue is then taken as the estimated sequence length.

Publication 3 illustrated that the sequence length of a DSSS signal hidden within noise can be estimated blindly. Blind estimation of the sequence length is important, as several DSSS detection and estimation techniques are available in the literature, which assume prior knowledge of the sequence length.

6.1.4 Connections between three publications

The following two important connections exist between the three publications.

- The analytic approximation presented in Publication 1 is used in Publication 2 to describe the distribution of the eigenvalue sequence under $\mathcal{H}_0$. The threshold level is also calculated (for a given false alarm rate) using the right-tail of the approximated distribution.
- The eigenvalue-based detection algorithm presented in Publication 2 is adapted and presented in Publication 3 as a sequence-length estimation algorithm.

6.2 Challenge of DSSS detection

For the unintended receiver, DSSS signals are usually weak and indistinguishable from noise. Advanced and processor-intensive algorithms are typically required to extract weak signals from noise, especially when the signals to be detected are unknown. Real-time blind signal detection is however required on tactical platforms, which have limited processing capability [10]. These platforms can only accommodate relatively low-complexity algorithms, such as ED which lack the performance to detect the typically weak DSSS signal blindly.

It is however possible to record the signals intercepted by tactical units and pass them on to strategic platforms for analysis on HPC facilities [11]. Processor-intensive algorithms can then be used to perform detection and estimation, but this approach only allows off-line processing. By developing algorithms with improved detection performances, and then reducing their complexities through simplifications (which may in turn result in a reduction in detection performance) may result in the realisation of low-complexity detection techniques with sufficient performance to offer a real-time solution for certain scenarios.

Development of high-complexity techniques should however continue, as future hardware platforms will most likely be capable of executing increasingly complex algorithms in real time. The aim of the newly developed detection and estimation techniques presented in this thesis was therefore on improving theoretical performance, even though real-time processing and analysis of signals (as they are being transmitted) may not be currently possible using these new techniques.

6.3 Importance of DSSS detection research

DSSS communication signals are designed to avoid detection by unintended receivers, and are therefore used in covert communication with a potentially hostile nature. New and improved detection and estimation techniques will benefit the communications intelligence community by providing an enhanced capability to enact countermeasures against these hidden transmissions. Exploiting or taking action against weak and unknown DSSS signals may include direction finding, message interception and extraction, and communication denial through jamming.

Improved weak signal detection techniques may also be applicable to the wider communications community, especially for receiver design and cognitive radio spectrum sensing applications [14]. Incorporating some of the blind detection concepts into non-blind systems may result in improved detection performance, extended communication ranges, and lower transmit powers.

6.4 Ethical issues surrounding the research topic

Ethical research issues concerning the work conducted and reported in this thesis are limited to indirect effects, as detailed below. Only theoretical research, including the analysis and development of signal detection and estimation techniques, statistical modelling, and computer simulation was conducted. However, when test and evaluation move beyond theoretical research, there are a number of ethical research issues that may become relevant, depending on how the research outputs are applied in hardware. These potential issues are highlighted below.

6.4.1 Lawful and unlawful interception

The legal implications of interception of communication signals, and the possibility of infringing on privacy rights are complex issues, with different laws and regulations applicable in different countries and scenarios [137, 138]. In certain circumstances, authorised law enforcement agencies may be allowed access to wireless communication transmissions by network operators (which are often required by law to implement lawful intercept strategies or systems in their networks), subject to interception regulations [139].

6.4.2 Communication denial

Although this thesis did not focus on jamming of communication signals [9], it is an important part of EW which depends on electronic support activities. A jamming device may include a detection system to intelligently deny active communication transmissions. The denial of licensed or authorised communication transmissions, and the sale or marketing of jamming devices, are prohibited by law in many countries [140].

6.4.3 Hardware test platforms

To practically evaluate detection and estimation techniques, it may be required to develop or use an experimental communication system as configurable hardware test platform. Such a platform would be required to transmit communication signals, which may cause interference (or jamming) in the licensed RF spectrum if configured incorrectly. Strategies to avoid such a scenario include transmitting in one of the unlicensed ISM bands [141], or performing evaluation in an isolated location such as an anechoic chamber.

6.5 Future research

The gaps in the literature identified in Section 2.5 can be considered further by continuing the development, evaluation and improvement of detection and estimation techniques. Realising low-complexity, high-performance algorithms is the ideal, as computational complexity and detection performance are usually traded off. Furthermore, future development of fully-blind detection and estimation techniques is also important as signal parameters are unknown in the non-cooperative context. Some areas of possible future research into the detection and estimation of DSSS signals are subsequently highlighted.

6.5.1 Reduced-complexity detection techniques

The most promising weak signal detection technique considered in this thesis is the PCA-based technique presented in Chapter 4. Detection based on nonlinear methods and chaos theory have also been reported to show promising performance [83]. Methods to simplify the underlying algorithms in these approaches should be developed to realise practical detection techniques. Simplifying the calculation of eigenvalues (for PCA) can possibly be achieved using the trace method [126] or the NN structure [93] discussed in Section 2.4.4.

6.5.2 Multi-channel receiver architectures

Most detection techniques considered in this thesis were designed for single-channel receiver systems. Methods to expand these techniques for application in multi-channel receivers should be investigated as improved performance can thus be expected. Using multiple receiver nodes is also the philosophy behind cooperative spectrum sensing (see Section 1.1.4 and the footnote on p. 7) as used in cognitive radio applications [14,72]. Some aspects of this approach have already been applied to DSSS detection using a dual-channel correlator structure [53], indicating a performance advantage.

6.5.3 Joint detection and estimation approaches

Detection and estimation are closely related and interdependent in many applications. Estimation may be performed as part of detection and vice versa [5]. Performing detection and estimation separately is not always optimal, and joint approaches may improve the performance [136]. Furthermore, the consistency (or fluctuation) of an estimated parameter value may also serve as confidence measure that the parameter is estimated correctly and that the signal of interest is indeed present.

An example of a joint approach include the detection of DSSS transmissions using the sequence-length estimate from the correlation technique as test statistic [54]. Conversely, the sequence length may be determined using the eigen-detection techniques as discussed in Chapter 5. Such joint approaches should be developed and evaluated further.

6.5.4 Algorithms for different spreading codes

Different spreading codes have unique properties [18, 28], and detection and estimation techniques will therefore perform differently on different spreading codes. For example, Walsh codes and m-sequences have respectively large and small sidelobe correlation values [28, 142]. The correlation-based, sequence-length estimation technique presented in Chapter 5 will only work on sequences with noise-like autocorrelation functions, such as m-sequences, and not on Walsh codes. Determining which algorithm will work on which type of spreading code is an area of possible future investigation. Developing techniques for specific spreading sequences, such as sequence-length estimation for Walsh codes, should also be considered.

6.5.5 Effect of parameter value uncertainty

The theoretical performance bound of a given detection technique can be determined by assuming perfect knowledge of all the unknown parameters [5]. The practical detection performance will typically be worse, since not all parameters can be estimated perfectly. However, some detection algorithms such as the cognitive radio techniques [68] do not require perfect knowledge of all parameter values. The effect of inaccurate parameter estimates on detection performance should be investigated. The performance and complexity of algorithms insensitive to parameter value uncertainty should also be evaluated and compared with algorithms that depend more heavily on parameter value accuracy.

6.6 Conclusion

A communication signal can be detected in noise by differentiating between the statistical properties of the signal and the noise [5]. This thesis considered the blind detection of DSSS communication signals, which were designed to resemble noise in order to avoid detection. The detection of such covert signals therefore presents a significant challenge to the intercept receiver, which can only be addressed by implementing sophisticated detection algorithms [16].

Sophisticated algorithms typically have high computational complexities, and real-time detection might not always be possible. Simpler techniques are favoured in practical applications, due to limited processing power [10]. This presents another challenge as intercepted DSSS signals typically have very low SNRs, beyond the detection capability of low-complexity algorithms. The ultimate aim is therefore to realise low-complexity, high-performance detection methods, which can be implemented in field-deployable systems. It is however still worthwhile to investigate computationally complex algorithms, as intercepted signals can be analysed off-line on HPC clusters [11].

Several detection and estimation algorithms, which are available in the literature, were reviewed in this thesis. A number of new algorithms were also developed and published, which are presented in Chapters 3 to 5. All these algorithms can broadly be categorised as energy-based [49] and correlation-based [55, 60, 72] techniques. ED has low computational requirements, but needs a relatively high SNR to perform detection. Correlation-based detection methods require more processing power, but show improved detection potential in low-SNR scenarios [16]. The two most promising detection methods are the PCA or eigen-detection techniques [16, 68, 77], and the nonlinear techniques [80, 83]. These techniques should be investigated further as suggested in Section 6.5.

Parameter estimation is also important and forms part of the detection process, as an unknown intercepted signal should be classified in order to confirm that the intercepted signal is in fact the signal of interest [15]. Furthermore, semi-blind detection and estimation techniques depend on prior knowledge of certain signal parameters. For example, the DSSS detection techniques presented in [16, 54] require knowledge of the spreading sequence length in order to detect the presence of the signal. Likewise, the estimation techniques presented in [77, 86, 90, 93] require the sequence length in order to estimate the actual chip sequence. If the spreading sequence can be estimated, the intercepted DSSS transmission can be despread, and the transmitted information can be accessed. The estimation techniques reviewed in this thesis focus on the extraction of the spreading sequence (in addition to other RF signal parameters) from the intercepted signal. Techniques to estimate the spreading sequence length are presented in Chapter 5 and published in [17].

The detection of both military and non-military communication systems were considered in this thesis, as DSSS signaling is widely used in both domains. Furthermore, the algorithms presented in this thesis are applicable to the detection of any type of weak signal, as long as it has a periodic pattern. It is this feature, the cyclostationary nature of communication signals such as DSSS transmissions, which make them detectable even if the signal parameters are not known a priori.

REFERENCES

[1] R. H. Pettit, *ECM and ECCM techniques for digital communication systems.* Belt-mont, CA, USA: Lifetime Learning Publications, 1982.

[2] D. L. Adamy, *EW103 Tactical battlefield communications electronic warfare.* Nor-wood, MA, USA: Artech House, Inc., 2009.

[3] W. A. Gardner, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Transactions on Communications*, vol. 36, no. 8, pp. 897–906, Aug. 1988.

[4] W. A. Gardner and C. M. Spooner, "Signal interception: Performance advantages of cyclic-feature detectors," *IEEE Transactions on Communications*, vol. 40, no. 1, pp. 149–159, Jan. 1992.

[5] S. M. Kay, *Fundamentals of statistical signal processing: Detection theory.* Upper Saddle River, NJ, USA: Prentice Hall, 2009, vol. 2.

[6] D. J. Torrieri, *Principles of spread-spectrum communication systems*, 2nd ed. New York, NY, USA: Springer, 2011.

[7] J. G. Proakis, *Digital communications*, 4th ed. Boston, MA, USA: McGraw-Hill, 2001.

[8] D. L. Adamy, *EW101 A first course in electronic warfare.* Norwood, MA, USA: Artech House, Inc., 2001.

[9] R. A. Poisel, *Modern communications jamming principles and techniques.* Nor-wood, MA, USA: Artech House, Inc., 2004.

[10] Grintek Ewation Technologies, "Tactical wideband search, direction find-ing, monitoring and jamming EW system," no. GMA-500-105-SG, Aug. 2012, Accessed: 21 Aug. 2013. [Online]. Available: http://www.gew.co.za/GenericTacticalSystemBrochureGMA-500-105-SGv0.05lores.pdf

[11] Grintek Ewation Technologies, "Strategic wideband search, direction find-ing and monitoring electronic warfare system," no. GMA-500-104-SG, Sep. 2012, Accessed: 21 Aug. 2013. [Online]. Available: http://www.gew.co.za/GenericStrategicSystemBrochureGMA-500-104-SG0.02.pdf

[12] US National Security Agency, "COMINT and COMSEC: The tactics of 1914-1918," Accessed: 21 Aug. 2013. [Online]. Available: http://www.nsa.gov/public_info/_files/cryptologic_spectrum/Comint_and_Comsec_Pt_1.pdf

[13] C. H. Sterling, *Military communications: From ancient times to the 21st century.* Santa Barbara, CA, USA: ABC-CLIO, Inc., 2008.

[14] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[15] S. M. Kay, *Fundamentals of statistical signal processing: Estimation theory.* Upper Saddle River, NJ, USA: Prentice Hall, 1993, vol. 1.

[16] J. D. Vlok and J. C. Olivier, "Non-cooperative detection of weak spread-spectrum signals in additive white Gaussian noise," *IET Communications*, vol. 6, no. 16, pp. 2513–2524, Nov. 2012.

[17] J. D. Vlok and J. C. Olivier, "Blind sequence-length estimation of low-SNR cyclostationary sequences," *IET Communications*, vol. 8, no. 9, pp. 1578–1588, Jun. 2014.

[18] T. S. Rappaport, *Wireless communications principles and practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.

[19] R. C. Dixon, *Spread spectrum systems with commercial applications*, 3rd ed. New York, NY, USA: John Wiley & Sons, 1994.

[20] *Wireless LAN medium access control (MAC) and physical layer (PHY) specifications*, IEEE 802.11-2007 Std., Accessed: 20 Aug. 2013. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11-2007.pdf

[21] Globalstar, Inc. Accessed: 20 Aug. 2013. [Online]. Available: http://www.globalstar.com

[22] P. Wang, J. Xiao, and L. Ping, "Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems," *IEEE Vehicular Technology Magazine*, vol. 1, no. 3, pp. 4–11, Sep. 2006.

[23] S. M. Kay, *Fundamentals of statistical signal processing: Practical algorithm development.* Upper Saddle River, NJ, USA: Prentice Hall, 2013, vol. 3.

[24] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, and F. Rossi, "GNU scientific library reference manual," no. 1.15, Apr. 2011, Accessed: 30 May 2013. [Online]. Available: http://www.gnu.org/software/gsl/manual/gsl-ref.pdf

[25] J. Bucklew and R. Radeke, "On the Monte Carlo simulation of digital communication systems in Gaussian noise," *IEEE Transactions on Communications*, vol. 51, no. 2, pp. 267–274, Feb. 2003.

[26] J. G. Proakis and M. Salehi, *Digital communications*, 5th ed. Boston, MA, USA: McGraw-Hill, 2007.

[27] J.-M. Muller, *Elementary functions: Algorithms and implementation*, 2nd ed. New York, NY, USA: Birkhäuser, 2006.

[28] W. Stallings, *Wireless communications and networks*. Upper Saddle River, New Jersey, USA: Prentice-Hall, 2002.

[29] B. C. Levy, *Principles of signal detection and parameter estimation*. New York, NY, USA: Springer-Verlag, 2010.

[30] Z. Xiao, D. Jin, L. Su, and L. Zeng, "Simplified direct search method for constrained nonlinear mixed-integer programming of two-dwell serial acquisition schemes," *IEEE Communications Letters*, vol. 15, no. 6, pp. 674–676, 2011.

[31] J. D. Vlok, "Sparse graph codes on a multi-dimensional WCDMA platform," Master's dissertation, Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa, Feb. 2007.

[32] D. Huffman, "The generation of impulse-equivalent pulse trains," *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 10–16, Sep. 1962.

[33] U. Parlitz and S. Ergezinger, "Robust communication based on chaotic spreading sequences," *Elsevier Physics Letters A*, vol. 188, no. 2, pp. 146–150, 1994.

[34] B. M. Popovic, "Generalized chirp-like polyphase sequences with optimum correlation properties," *IEEE Transactions on Information Theory*, vol. 38, no. 4, pp. 1406–1409, Jul. 1992.

[35] L. P. Linde and J. D. Vlok, "Power and spectrally efficient four-dimensional super-orthogonal WCDMA building block for next generation wireless applications," *IEEE Communications Letters*, vol. 10, no. 7, pp. 519–521, Jul. 2006.

[36] B. Rice, "A transceiver design for a high information rate LPI network," in *Proceedings of IEEE Military Communications Conference (MILCOM)*, vol. 3, Monterey, CA, USA, Sep. 1990, pp. 990–993.

[37] R. L. Frank, "Comments on polyphase codes with good periodic correlation properties by Chu, David C," *IEEE Transactions on Information Theory*, vol. 19, no. 2, pp. 244–244, Mar. 1973.

[38] I. Pryra, L. P. Linde, and S. A. Swanepoel, "New family of constant-envelope root-of-unity filtered complex spreading sequences with zero cross-correlation properties," in *Proceedings of IEEE Conference in Africa (AFRICON)*, vol. 1, George, South Africa, Oct. 2002, pp. 299–304.

[39] H.-H. Chen and M. Guizani, *Next generation wireless systems and networks.* Chichester, West Sussex, England: John Wiley & Sons, 2006.

[40] Global positioning system. US government national space-based positioning, navigation, and timing coordination office. Accessed: 20 Aug. 2013. [Online]. Available: http://www.gps.gov

[41] ESA - Navigation - The future - Galileo. Accessed: 20 Aug. 2013. [Online]. Available: http://www.esa.int/esaNA/galileo.html

[42] Conexant systems, Inc. Accessed: 20 Aug. 2013. [Online]. Available: http://www.conexant.com

[43] OPENmeter open public extended network metering. Accessed: 20 Aug. 2013. [Online]. Available: http://www.openmeter.com

[44] R. MacCurdy, R. Gabrielson, E. Spaulding, A. Purgue, K. Cortopassi, and K. Fristrup, "Automatic animal tracking using matched filters and time difference of arrival," *Academy Publisher Journal of Communications*, vol. 4, no. 7, pp. 487–495, Aug. 2009.

[45] J. Gardner, "Spread spectrum systems for EFM and SCADA," in *Proceedings of International School of Hydrocarbon Measurement*, vol. 2, Oklahoma City, OK, USA, May 2010, pp. 381–384.

[46] J. Lehtomäki, "Analysis of energy based signal detection," Ph.D. thesis, Department of Electrical and Information Engineering, Faculty of Technology, University of Oulu, Finland, 2005.

[47] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Elsevier Signal Processing*, vol. 86, no. 4, pp. 639–697, 2006.

[48] R. A. Dillard and G. M. Dillard, *Detectability of spread-spectrum signals.* Norwood, MA, USA: Artech House, 1989.

[49] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523–531, 1967.

[50] S. Davidovici and E. G. Kanterakis, "Radiometric detection of direct-sequence spread-spectrum signals using interference excision," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 4, pp. 576–589, 1989.

[51] J. Lehtomäki, "Detection of spread spectrum signals using a power-law based intercept receiver," in *Proceedings of the 59th IEEE Vehicular Technology Conference (VTC Spring)*, vol. 3, Milan, Italy, May 2004, pp. 1480–1484.

[52] L. Chang, F. Wang, and Z. Wang, "Detection of DSSS signal in non-cooperative communications," in *Proceedings of IEEE International Conference on Communication Technology (ICCT)*, Guilin, Guangxi, China, Nov. 2006, pp. 1–4.

[53] A. W. Houghton and C. D. Reeve, "Detection of spread-spectrum signals using the time-domain filtered cross spectral density," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 142, no. 6, pp. 286–292, Dec. 1995.

[54] G. Burel, "Detection of spread spectrum transmissions using fluctuations of correlation estimators," in *Proceedings of IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Honolulu, Hawaii, USA, Nov. 2000.

[55] G. Burel, C. Bouder, and O. Berder, "Detection of direct sequence spread spectrum transmissions without prior knowledge," in *Proceedings of IEEE Global Communications Conference (GLOBECOM)*, vol. 1, San Antonio, TX, USA, Nov. 2001, pp. 236–239.

[56] C. L. Nikias and M. R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869–891, 1987.

[57] T. T. Soong, *Fundamentals of probability and statistics for engineers*. Buffalo, NY, USA: John Wiley & Sons, 2004.

[58] C. L. Nikias, "Higher-order spectral analysis," in *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, Oct. 1993, pp. 319–319.

[59] P. C. J. Hill, V. E. Comley, and E. R. Adams, "Techniques for detecting and characterising covert communication signals," in *Proceedings of IEEE Military Communications Conference (MILCOM)*, vol. 3, Monterey, CA, USA, Nov. 1997, pp. 1361–1365.

[60] E. R. Adams and P. C. J. Hill, "Detection of direct sequence spread spectrum signals using higher-order statistical processing," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 5, Munich, Germany, Apr. 1997, pp. 3849–3852.

[61] G. E. Carlson, *Signal and linear system analysis*, 2nd ed. New York, NY, USA: John Wiley, 1998.

[62] F. Hlawatsch and F. Auger, *Time-frequency analysis*. Hoboken, NJ, USA: John Wiley & Sons, 2008.

[63] S. G. Mallat, "A theory for multi-resolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[64] N. Ahuja, S. Lertrattanapanich, and N. K. Bose, "Properties determining choice of mother wavelet," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 152, no. 5, pp. 659–664, Oct. 2005.

[65] R. J. R. Landry, P. Mouyon, and D. Lekaïm, "Interference mitigation in spread spectrum systems by wavelet coefficients thresholding," *Wiley European Transactions on Telecommunications*, vol. 9, no. 2, pp. 191–202, Apr. 1998.

[66] Z. Zhao, Z. Sun, and F. Mei, "A threshold detection method of DSSS signal based on STFT," in *Proceedings of IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, vol. 2, Beijing, China, Aug. 2005, pp. 879–882.

[67] M. Medley, G. Saulnier, and P. Das, "Radiometric detection of direct-sequence spread spectrum signals with interference excision using the wavelet transform," in *Proceedings of IEEE International Conference on Communications (ICC)*, vol. 3, New Orleans, LA, USA, May 1994, pp. 1648–1652.

[68] L. Wei and O. Tirkkonen, "Analysis of scaled largest eigenvalue based detection for spectrum sensing," in *Proceedings of IEEE International Conference on Communications (ICC)*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[69] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.

[70] I. T. Jolliffe, *Principal component analysis*, 2nd ed. New York, NY, USA: Springer-Verlag, 2002.

[71] L. I. Smith, "A tutorial on principal component analysis," Feb. 2002, Accessed: 21 Aug. 2013. [Online]. Available: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

[72] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *Springer EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 1–15, 2010.

[73] A. Bejan, "Largest eigenvalues and sample covariance matrices," Master's dissertation, Department of Statistics, University of Warwick, UK, Jun. 2005.

[74] Y. Zeng, C. L. Koh, and Y. C. Liang, "Maximum eigenvalue detection: Theory and application," in *Proceedings of IEEE International Conference on Communications (ICC)*, Beijing, China, May 2008, pp. 4160–4164.

[75] G. H. Golub and C. F. Van Loan, *Matrix computations*, 3rd ed. Baltimore, MD, USA: John Hopkins University Press, 1996.

[76] J. D. Vlok and J. C. Olivier, "Analytic approximation to the largest eigenvalue distribution of a white Wishart matrix," *IET Communications*, vol. 6, no. 12, pp. 1804–1811, Aug. 2012.

[77] Y. Zhan, Z. Cao, and J. Lu, "Spread-spectrum sequence estimation for DSSS signal in non-cooperative communication systems," *IEE Proceedings - Communications*, vol. 152, no. 4, pp. 476–480, Aug. 2005.

[78] E. N. Lorenz, *The essence of chaos*. Seattle, WA, USA: University of Washington Press, 1995.

[79] S. H. Strogatz, *Nonlinear dynamics and chaos*. Cambridge, MA, USA: Westview Press, 2000.

[80] R. Brown, L. Chua, and B. Popp, "Is sensitive dependence on initial conditions nature's sensory device?" *World Scientific International Journal of Bifurcation and Chaos*, vol. 2, no. 1, pp. 193–199, 1992.

[81] E. N. Lorenz, "Predictability: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?" in *The essence of chaos*. Seattle, WA, USA: University of Washington Press, 1995, Appendix 1.

[82] S. Wiggins, *Introduction to applied nonlinear dynamical systems and chaos*, 2nd ed. New York, NY, USA: Springer, 2003.

[83] G. Wang, D. Chen, J. Lin, and X. Chen, "The application of chaotic oscillators to weak signal detection," *IEEE Transactions on Industrial Electronics*, vol. 46, no. 2, pp. 440–444, Apr. 1999.

[84] G. Wang and S. He, "A quantitative study on detection and estimation of weak signals by using chaotic duffing oscillators," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, no. 7, pp. 945–953, Jul. 2003.

[85] H. Jin and K. Wang, "Carrier detection method of BPSK and DSSS signals based on Duffing oscillator," in *Proceedings of the 6th International Conference on ITS Telecommunications*, Chengdu, China, Jun. 2006, pp. 1338–1341.

[86] M. K. Tsatsanis and G. B. Giannakis, "Blind estimation of direct sequence spread spectrum signals in multipath," *IEEE Transactions on Signal Processing*, vol. 45, no. 5, pp. 1241–1252, 1997.

[87] C. Bouder, S. Azou, and G. Burel, "A robust synchronization procedure for blind estimation of the symbol period and the timing offset in spread spectrum transmissions," in *Proceedings of the IEEE 7th International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, Prague, Czech Republic, 2002, pp. 238–241.

[88] M. Nakamura, "Waveform estimation from noisy signals with variable signal delay using bispectrum averaging," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 2, pp. 118–127, Feb. 1993.

[89] A. P. Petropulu and C. L. Nikias, "Signal reconstruction from the phase of the bispectrum," *IEEE Transactions on Signal Processing*, vol. 40, no. 3, pp. 601–610, 1992.

[90] G. Burel and C. Bouder, "Blind estimation of the pseudo-random sequence of a direct sequence spread spectrum signal," in *Proceedings of IEEE Military Communications Conference (MILCOM)*, vol. 2, Los Angeles, CA, USA, Oct. 2000, pp. 967–970.

[91] C. Bouder, S. Azou, and G. Burel, "Performance analysis of a spreading sequence estimator for spread spectrum transmissions," *Elsevier Journal of the Franklin Institute*, vol. 341, no. 7, pp. 595–614, 2004.

[92] S. S. Haykin, *Neural networks and learning machines*. New York, NY, USA: Prentice Hall, 2009, vol. 3.

[93] F. Dominique and J. H. Reed, "Simple PN code sequence estimation and synchronisation technique using the constrained Hebbian rule," *IEE Electronics Letters*, vol. 33, no. 1, pp. 37–38, Jan. 1997.

[94] C. Bouder and G. Burel, "Spread spectrum codes identification by neural networks," in *Systems and Control: Theory and Applications*. WSES, 2000, pp. 257–262.

[95] A. Zanella, M. Chiani, and M. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Transactions on Communications*, vol. 57, no. 4, pp. 1050–1060, 2009.

[96] F. Penna, R. Garello, and M. Spirito, "Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in Wishart matrices," *IEEE Communications Letters*, vol. 13, no. 7, pp. 507–509, 2009.

[97] W. Xu and M. Kaveh, "Analysis of the performance and sensitivity of eigendecomposition-based detectors," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1413–1426, 1995.

[98] A. Edelman and N. R. Rao, "Random matrix theory," *Cambridge University Acta Numerica*, vol. 14, pp. 233–297, May 2005.

[99] I. Telatar, "Capacity of multi-antenna Gaussian channels," *Wiley European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[100] R. Kwan, C. Leung, and P. Ho, "Distribution of ordered eigenvalues of Wishart matrices," *IET Electronics Letters*, vol. 43, no. 5, pp. 31–32, 2007.

[101] I. Johnstone, "Approximate null distribution of the largest root in multivariate analysis," *IMS Annals of Applied Statistics*, vol. 3, no. 4, pp. 1616–1633, 2009.

[102] I. Johnstone, "On the distribution of the largest eigenvalue in principal component analysis," *IMS Annals of Statistics*, vol. 29, no. 2, pp. 295–327, 2001.

[103] C. Tracy and H. Widom, "On orthogonal and symplectic matrix ensembles," *Springer Communications in Mathematical Physics*, vol. 177, no. 3, pp. 727–754, 1996.

[104] C. Tracy and H. Widom, "The distribution of the largest eigenvalue in the Gaussian ensembles," *Calogero-Moser-Sutherland models: CRM series in mathematical physics 4 (Springer-Verlag) (Van Diejen, J.F., Vinet, L. (Eds.))*, pp. 461–472, 2000.

[105] M. Dieng, "Distribution functions for edge eigenvalues in orthogonal and symplectic ensembles: Painlevé representations," Ph.D. thesis, University of California, Davis, USA, Jul. 2005.

[106] M. Prähofer and H. Spohn, "Exact scaling functions for one-dimensional stationary KPZ growth," *Springer Journal of Statistical Physics*, vol. 115, no. 1–2, pp. 255–279, 2004.

[107] F. Bornemann, "On the numerical evaluation of distributions in random matrix theory: A review," *Markov Processes and Related Fields*, vol. 16, no. 4, pp. 803–866, 2010.

[108] M. Prähofer and H. Spohn. (2012) Exact scaling functions for one-dimensional stationary KPZ growth. Accessed: May 2014. [Online]. Available: http://www-m5.ma.tum.de/KPZ

[109] F. James, *Statistical methods in experimental physics*, 2nd ed. Singapore: World Scientific, 2010.

[110] A. Papoulis and S. U. Pillai, *Probability, random variables and stochastic processes*, 4th ed. Singapore: McGraw-Hill, 2002.

[111] E. R. Adams, M. Gouda, and P. C. J. Hill, "Statistical techniques for blind detection & discrimination of m-sequence codes in DS/SS systems," in *Proceedings of the IEEE 5th International Symposium on Spread Spectrum Techniques and Applications*, vol. 3, Sun City, South Africa, Sep. 1998, pp. 853–857.

[112] K. Olivier, J. E. Cilliers, and M. Du Plessis, "Design and performance of wideband DRFM for radar test and evaluation," *IET Electronics Letters*, vol. 47, no. 14, pp. 824–825, Jul. 2011.

[113] R. Everson and S. Roberts, "Inferring the eigenvalues of covariance matrices from limited, noisy data," *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 2083–2091, Jul. 2000.

[114] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge, UK: Cambridge University Press, 2012.

[115] T. Tao, *Topics in random matrix theory (Graduate studies in mathematics)*. Rhode Island, USA: American Mathematical Society, 2012.

[116] C. Li and Y. Poon, "Sum of Hermitian matrices with given eigenvalues: inertia, rank, and multiple eigenvalues," *Canadian Journal of Mathematics*, vol. 62, no. 1, pp. 109–132, 2010.

[117] X. Zhan, "Extremal eigenvalues of real symmetric matrices with entries in an interval," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 3, pp. 851–860, 2006.

[118] J. Wu, "Upper (lower) bounds of the eigenvalues, spread and the open problems for the real symmetric interval matrices," *Wiley Mathematical Methods in the Applied Sciences*, vol. 36, no. 4, pp. 413–421, Mar. 2013.

[119] E. Wigner, "On the distribution of the roots of certain symmetric matrices," *Princeton University Annals of Mathematics*, vol. 67, no. 2, pp. 325–327, Mar. 1958.

[120] D. S. Dean and S. N. Majumdar, "Extreme value statistics of eigenvalues of Gaussian random matrices," *APS Physical Review E*, vol. 77, no. 4, pp. 041 108–1–041 108–12, 2008.

[121] J. W. Silverstein, "Eigenvalues and eigenvectors of large dimensional sample covariance matrices," *AMS Contemporary Mathematics*, vol. 50, pp. 153–159, 1986.

[122] J. Muller, *Elementary functions: Algorithms and implementation*, 2nd ed. Boston, USA: Birkhäuser, 2006.

[123] R. L. Burden and J. D. Faires, *Numerical analysis*, 9th ed. MA, USA: Cengage Learning, 2011.

[124] S. Smale, "Complexity theory and numerical analysis," *Cambridge University Acta Numerica*, vol. 6, pp. 523–551, 1997.

[125] V. Strassen, "Gaussian elimination is not optimal," *Springer Numerische Mathematik*, vol. 13, no. 4, pp. 354–356, 1969.

[126] E. A. Gonzalez, "Determination of the dominant eigenvalue using the trace method," *IEEE Multidisciplinary Engineering Education Magazine*, vol. 1, no. 1, pp. 1–2, 2006.

[127] BLAS (Basic Linear Algebra Subprograms). Accessed: May 2014. [Online]. Available: http://www.netlib.org/blas

[128] GSL - GNU Scientific Library. Accessed: May 2014. [Online]. Available: http://www.gnu.org/software/gsl

[129] D. Middleton, *Non-Gaussian statistical communication theory*. Hoboken, NJ, USA: Wiley-IEEE Press, 2012.

[130] I. Akyildiz, D. Levine, and I. Joe, "A slotted CDMA protocol with BER scheduling for wireless multimedia networks," *IEEE/ACM Transactions on Networking*, vol. 7, no. 2, pp. 146–158, Apr. 1999.

[131] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proceedings of the IEEE*, vol. 68, no. 5, pp. 593–619, May 1980.

[132] E. W. Weisstein, *CRC concise encyclopedia of mathematics*, 2nd ed. Boca Raton, Florida, USA: CRC Press, 2003.

[133] R. Roy and T. Kailath, "ESPRIT–Estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[134] Y. Hua and T. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 5, pp. 814–824, 1990.

[135] Y. Hua and T. Sarkar, "On SVD for estimating generalized eigenvalues of singular matrix pencil in noise," *IEEE Transactions on Signal Processing*, vol. 39, no. 4, pp. 892–900, Apr. 1991.

[136] G. V. Moustakides, G. H. Jajamovich, A. Tajer, and X. Wang, "Joint detection and estimation: Optimum tests and applications," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4215–4229, Jul. 2012.

[137] V. Horniak, "Privacy of communication - Ethics and technology," Master's dissertation, Department of Computer Science and Engineering, Mälardalen University, Sweden, 2004.

[138] Oxford pro bono publico. (2006, Jan.) The justice project: Legal opinion on intercept communication. University of Oxford. Accessed: 22 Aug. 2013. [Online]. Available: http://www2.law.ox.ac.uk/opbp/OPBP%20Intercept%20Evidence%20Report.pdf

[139] European telecommunications standards institute. Lawful interception for mobile networks. Accessed: 22 Aug. 2013. [Online]. Available: http://www.etsi.org/technologies-clusters/technologies/mobile/li-over-mobile-networks

[140] Jammer enforcement. Federal communications commission. Accessed: 22 Aug. 2013. [Online]. Available: http://www.fcc.gov/encyclopedia/jammer-enforcement

[141] International telecommunication union radiocommunication sector (ITU-R). (2012) Radio Regulations Article 5, Vol. 1. Accessed: 22 Aug. 2013. [Online]. Available: http://www.itu.int/ITU-R/terrestrial

[142] D. Kedia, M. Duhan, and S. L. Maskara, "Evaluation of correlation properties of orthogonal spreading codes for CDMA wireless mobile communication," in *Proceedings of IEEE 2nd International Advance Computing Conference (IACC)*, Patiala, India, Feb. 2010, pp. 325–330.

[143] L. Staphorst, "Viterbi decoded linear block codes for narrowband and wideband wireless communication over mobile fading channels," Master's dissertation, Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa, Jul. 2005.

[144] D. J. C. MacKay, *Information theory, inference and learning algorithms.* Cambridge, UK: Cambridge University Press, 2005.

[145] *Maxima: A computer algebra system*, http://maxima.sourceforge.net/, Accessed: 14 Oct. 2013.

APPENDICES

APPENDIX A

SIMULATION OF WIDEBAND SIGNALS IN AWGN

This appendix considers the simulation of spread spectrum signals in AWGN conditions which was used in this thesis to perform Monte-Carlo analysis. Further details on the simulation platform discussed in Section 4.2.2 are also provided.

A.1 Oversampling of signals

Figs. A.1 and A.2 show the temporal and spectral representations of a DSSS signal $x(t)$ spread with the Barker-11 code (10110111000) using 10 samples per chip. The AWGN sequence $w(t) \sim \mathcal{N}(0,1)$ and the sum of the signal and noise (all sampled at the same sampling frequency) are also shown. The signal and noise powers are $P_s = P_n = 1\ W$, such that the SNR is 0 dB[1]. The top graph of Fig. A.1 shows ten spread bits (1001101010) of the signal and the top graph of Fig. A.2 shows the associated power spectrum. Table A.1 contains the parameter values used to plot the graphs in Figs. A.1 and A.2.

Although the SNR is 0 dB for $P_s = P_n$, Fig. A.2 indicates that a significant portion of the noise lies outside the main lobe (or first null-to-null bandwidth) of the signal. The SNR can therefore easily be increased by filtering out the high-frequency noise, although filtering will introduce other effects (such as changes to the signal amplitude and phase, and the introduction of correlation) and will complicate the computer simulation model. Oversampling the digital signal therefore results in a simulation model with a relatively narrow-band information signal in wideband noise. Without filtering, the SNR will therefore be unrealistically low in the simulation, and performance predictions will be overly optimistic if the approach is not amended.

The reasons for sampling at a frequency greater than or equal to the Nyquist rate include avoiding aliasing to allow reconstruction of an analogue signal from its discrete-time equivalent, increasing the spectral distance between the main and mirror spectral lobes to ease filter requirements, and for visual display of signals. For information transmission purposes, it is however possible to sample at frequencies lower than the Nyquist rate.

[1]This SNR measurement does not take sampling bandwidth into account.

Table A.1: Parameter values used in Figs. A.1 and A.2.

Parameter	**Value**	**Parameter**	**Value**
Data rate f_b	1 Mbps	Bit period T_b	1 μs
Chip rate f_{chip}	11 Mcps	Chip period T_{chip}	90.91 ns
Sampling rate f_s	110 MHz	Sampling period T_s	9.091 ns
Chips per bit N	11	Samples per chip N_{spc}	10
Samples per bit $N \times N_{\text{spc}}$	110	Total bits N_T (Fig. A.2)	100

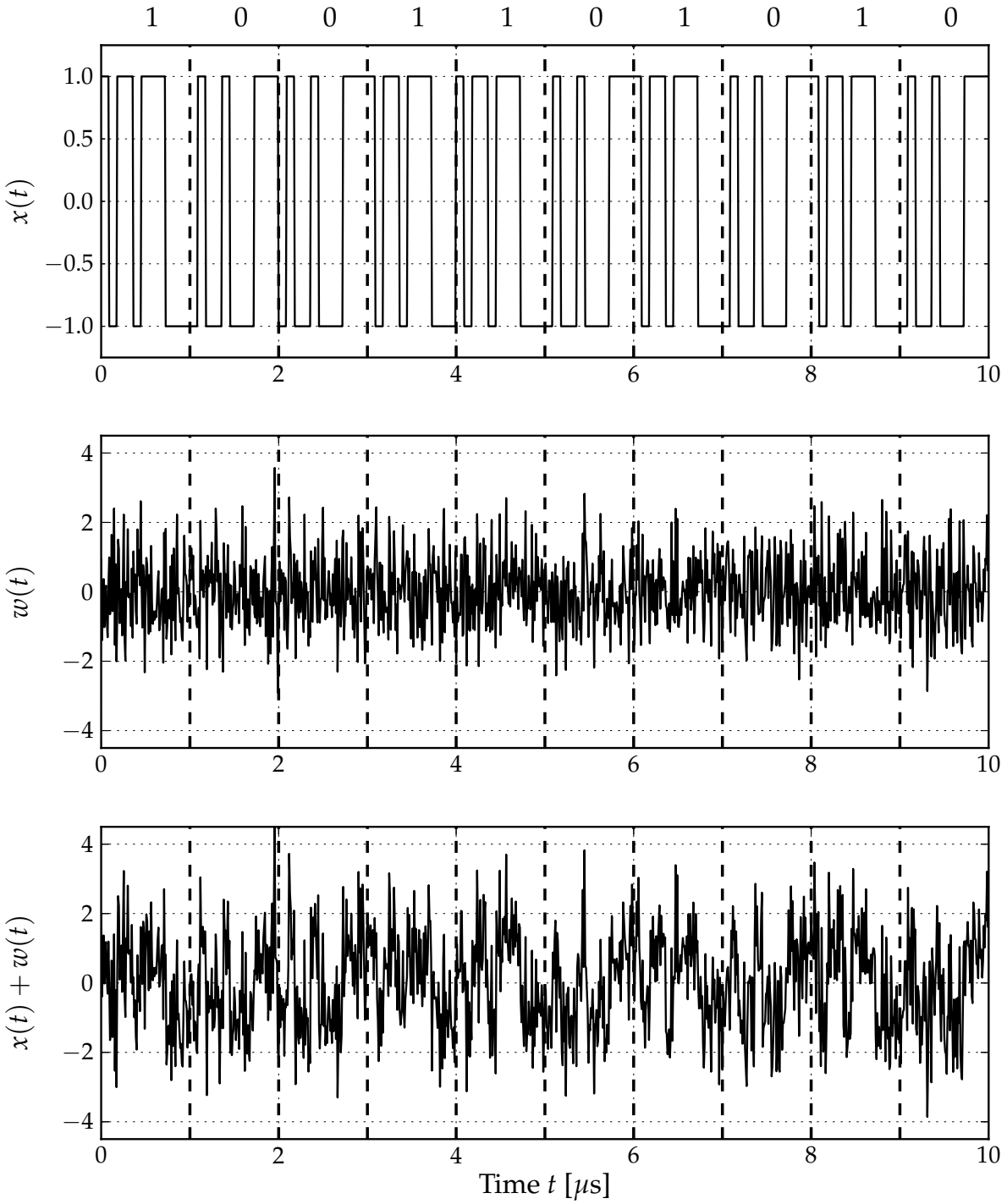


Figure A.1: Temporal representations of Barker-11 sequence and AWGN ($N_{\text{spc}} = 10$).

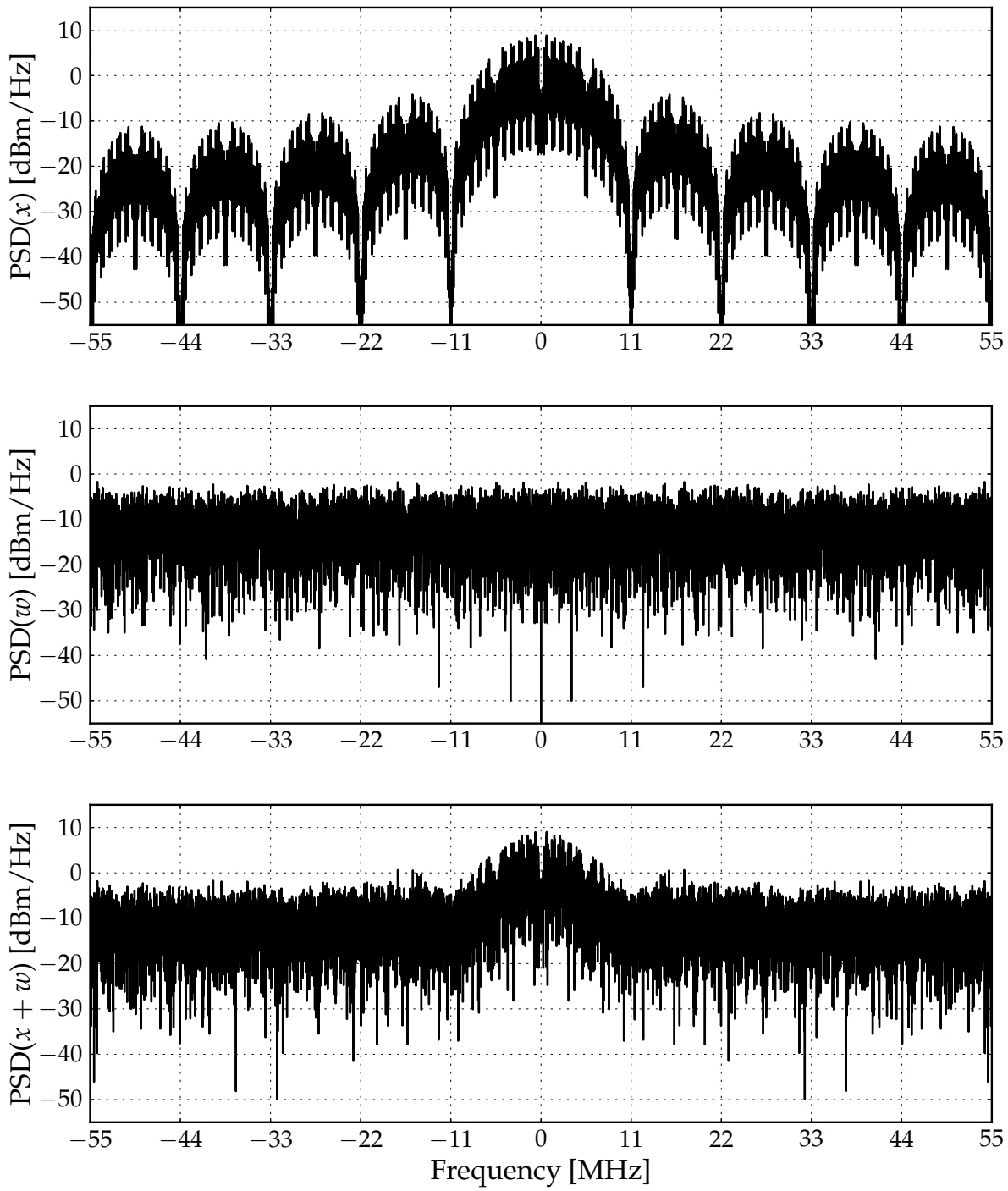


Figure A.2: Spectral representations of Barker-11 sequence and AWGN ($N_{\text{spc}} = 10$).

A.2 Nyquist sampling and bandwidth

According to the Nyquist sampling theorem [61], the sampling rate f_s must exceed the highest frequency component in the continuous-time signal by at least a factor of two, to allow reconstruction of the signal. Assuming the DSSS signal is bandlimited to f_{chip}, the

sampling rate should then be

$$f_s \geq 2f_{\text{chip}} \tag{A.1}$$

$$\therefore \frac{f_s}{f_{\text{chip}}} \geq 2 \tag{A.2}$$

where the factor can be shown to be

$$\frac{f_s}{f_{\text{chip}}} = \frac{T_{\text{chip}}}{T_s} = N_{\text{spc}} \tag{A.3}$$

which is the number of samples used to represent a single chip of the spreading sequence. At least two samples per chip are therefore required to satisfy the Nyquist sampling requirement. However, one sample per chip is sufficient to carry the digital information, as the Nyquist bandwidth [28] (in Hz) is half the chip rate (in chips per second) using an antipodal DSSS sequence. The aim is not to reconstruct a continuous-time signal, and therefore sampling at the chip rate $f_s = f_{\text{chip}}$ or $N_{\text{spc}} = 1$ can be allowed. It is however assumed that perfect symbol or chip synchronisation can be achieved in the receiver, such that sampling occurs at the right instant to recover the digital information.

Figs. A.3 and A.4 show the temporal and spectral representations of the Barker-11 spread signal $x(t)$ using 1 sample per chip. Due to undersampling, the noise $w(t)$ is limited to the Nyquist bandwidth of the digital signal, and no noise filtering is therefore required. As no correlation is introduced due to filtering, white noise scenarios can thus be simulated and comparisons can easily be made with theoretical derivations based on AWGN assumptions (which are readily available in several text books such as [5]) also discussed in Section A.3. Furthermore, the simulation is simplified using $N_{\text{spc}} = 1$, as the minimum number of samples is used, which limits computer memory usage and minimises simulation time.

It is clear that the noise power is higher in Fig. A.4 than in A.2, although the SNR is 0 dB in both figures. The lower sampling frequency limits the noise power to the main spectral lobe of the signal, and the SNR level is therefore more realistic (although the SNR is still only a function of signal and noise powers and not of the sampling bandwidth).

Nothing will be gained by using more than 1 sample per chip in the Monte-Carlo computer simulations, as was done in this thesis to evaluate the performance of detection algorithms. In a practical receiver, it would however make sense to sample at a higher frequency. Advantages include the simplification of symbol synchronisation and the reasons mentioned in Section A.1.

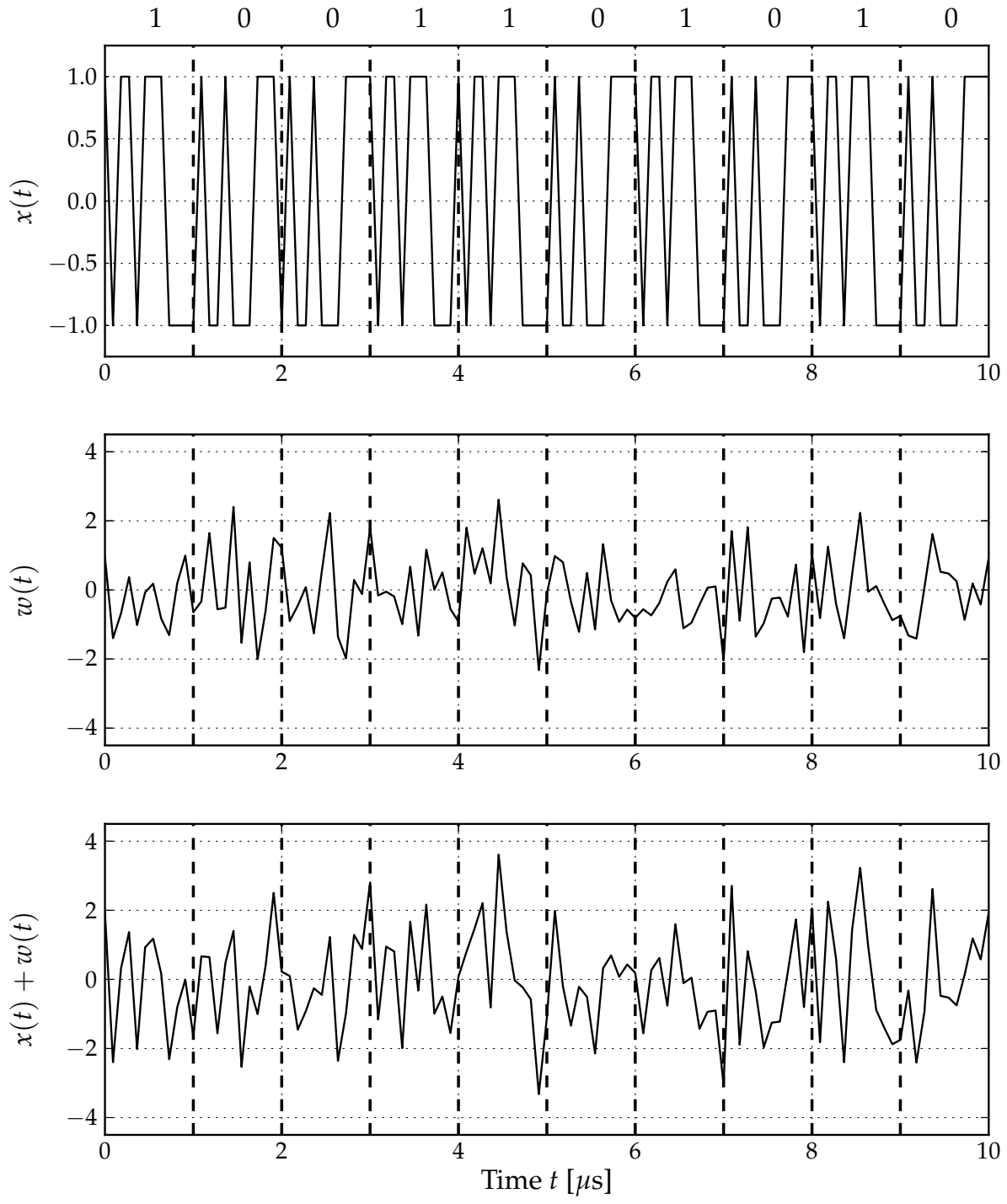


Figure A.3: Temporal representations of Barker-11 sequence and AWGN ($N_{\mathrm{spc}} = 1$).

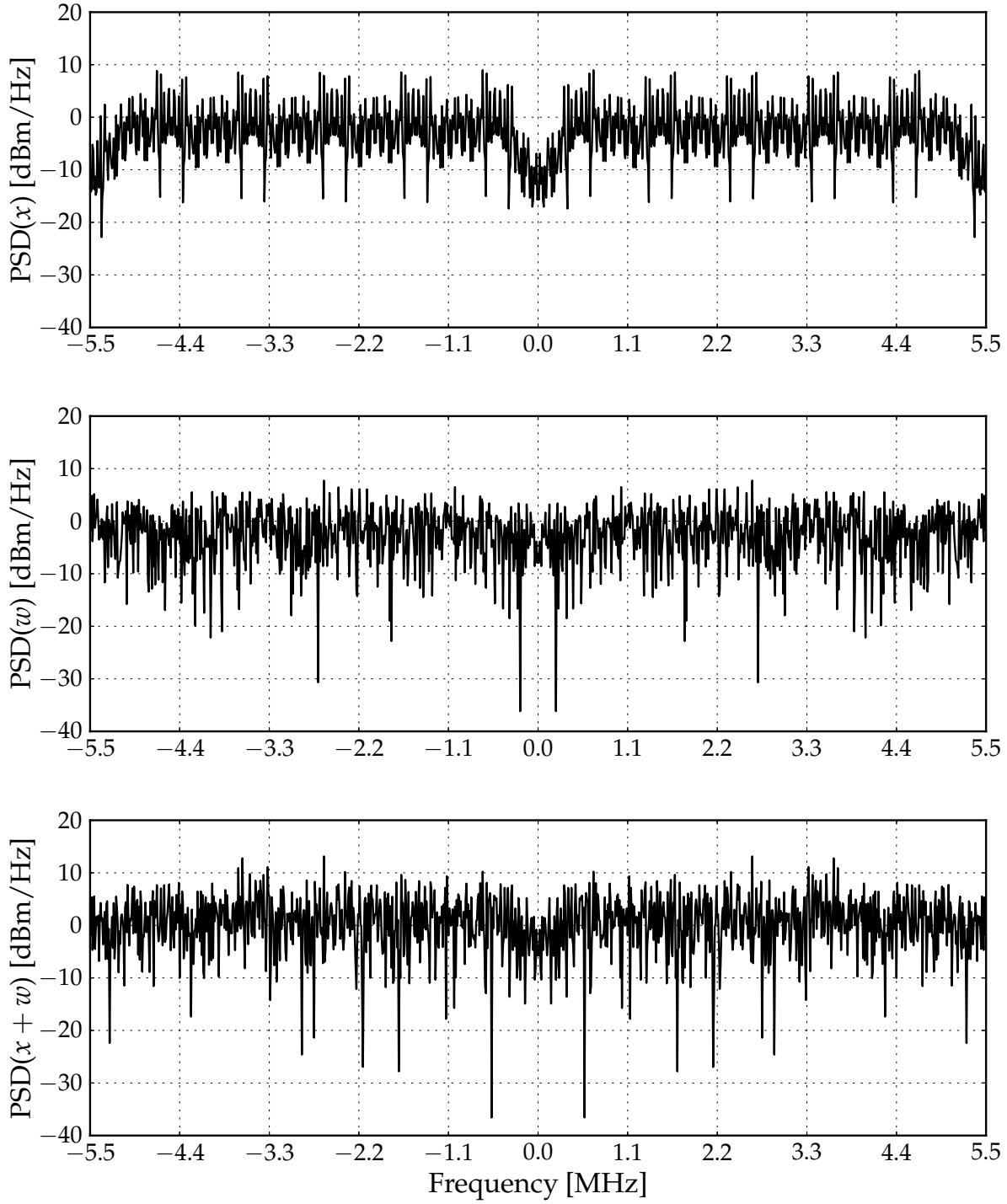


Figure A.4: Spectral representations of Barker-11 sequence and AWGN ($N_{\text{spc}} = 1$).

A.3 Bit error rate performance

The method used in this thesis to validate the DSSS simulation platform shown in Fig. 1.5 and also discussed in Sections 4.2.1 and 5.2 is considered here. By measuring the BER in simulation and comparing it with theoretical derivations, the simulation model and parameters (including the data and noise powers, and the noise distribution) can be confirmed to be correct.

The error probability of BPSK in AWGN is well known as [26]

$$P_e = Q\left(\sqrt{\frac{2\ \varepsilon_b}{N_0}}\right) \tag{A.4}$$

with $Q(\cdot)$ the right-tail probability of the standard normal distribution, ε_b the energy per bit and N_0 the single-sided noise PSD. The quantity ε_b/N_0 is the SNR normalised with the noise-equivalent bandwidth of a digital receiver [143], which provides a convenient metric to compare the performances of different communication systems where the parameter values of the communication systems are available.

In the non-cooperative context, using the SNR metric makes more sense as the intercept receiver typically has no knowledge regarding the information signal (such as the bandwidth, modulation type, bit energy, etc.). Equation (A.4) can be rewritten in terms of SNR by noting that

$$P_s = \frac{\varepsilon_b}{T_b} \tag{A.5}$$

$$\therefore \varepsilon_b = P_s T_b. \tag{A.6}$$

Furthermore, the single-sided PSD can be written as

$$N_0 = \frac{P_N}{f_s/2} = 2P_N T_s \tag{A.7}$$

as the single-sided simulation bandwidth is $f_s/2$. The normalised SNR can therefore be written by combining (A.6) and (A.7) as

$$\frac{\varepsilon_b}{N_0} = \frac{P_s T_b}{2P_N T_s} = \frac{\text{SNR} \times N_{\text{spc}}}{2} \tag{A.8}$$

such that (A.4) can be written as

$$P_e = Q\left(\sqrt{\text{SNR} \times N_{\text{spc}}}\right) \tag{A.9}$$

which was used to plot the graphs shown in Figs. 4.1 and 5.1.

APPENDIX B

ESTIMATION OF COMMUNICATION SIGNAL PARAMETERS

The focus of this thesis is on the detection of DSSS signals. However, parameter estimation is closely related and some detection approaches also rely on estimation principles. This appendix briefly considers some concepts of parameter estimation theory [15, 29].

B.1 Estimation theory

Parameter estimation is concerned with extracting the value of an unknown parameter θ from a received signal. The estimated value or estimator $\hat{\theta}$ is determined by applying a function to the received signal (similar to performing detection). Estimation accuracy can be defined in terms of the average estimation error or mean square error (MSE) measure [15]

$$\begin{aligned} \text{MSE}\left(\hat{\theta}\right) &= E\left[\left(\hat{\theta}-\theta\right)^2\right] && \text{(B.1)} \\ &= \text{var}\left(\hat{\theta}\right) + \left[E\left(\hat{\theta}\right)-\theta\right]^2 && \text{(B.2)} \end{aligned}$$

with the first term the variance, and the second term the bias. If the parameter value is correctly estimated on average, such that $E(\hat{\theta}) = \theta$, the estimator is said to be unbiased. The unbiased estimator with the minimum variance (which can theoretically be derived using the Cramer-Rao lower bound [15]) is then the minimum-variance unbiased estimator. Generally, the MSE cannot be used to realise an estimation formula, as (B.2) depends on θ which is the actual parameter to be estimated (which is unknown). However, several other methods and theorems have been developed which may be used to develop an estimation strategy [15].

Parameter estimation can be performed mainly using two approaches, which depend on assumptions made regarding the unknown parameter. The unknown parameter can be assumed to be either a deterministic constant (classical estimation) or a realisation of a random process, where a prior PDF is assigned to the parameter (Bayesian estimation).

B.2 Joint detection and estimation

Signal detection and parameter estimation are traditionally performed separately and independently in communication problems [129]. Detection is performed first to determine whether the signal of interest is present within noise or not. When a decision is made that the signal of interest is present, the parameter values may be estimated. Although this approach may make sense in several applications, it cannot always be ascertained that the signal of interest is surely present before attempting to perform estimation.

In certain applications detection and estimation can be formulated as a joint problem resulting in improved performance [136]. Other applications require parameter values (or their estimates) in order to perform detection (e.g. the estimator-correlator [5]). Referring to the subject area as either "detection and estimation" or "estimation and detection" is therefore equally valid in many applications.

The separation of signal detection and parameter estimation as distinct operations is being replaced by simultaneous detection and estimation techniques, mainly due to the rapidly changing environment in which modern communication systems operate. It is therefore desirable to constantly update parameter estimates while performing detection on the received signals [29]. If joint detection and estimation becomes a requirement in cooperative communication system design, it will become even more important for COMINT receivers.

B.3 Estimation of DSSS sequence length

Chapter 5 considered the estimation of the sequence length (an integer value) of spreading sequences. The estimation accuracy was expressed in terms of the probability of correct estimation P_{ce}, which in this case is more appropriate than for example the MSE given in (B.2). In many estimation problems, the closer the estimate is to the true parameter value, the better the performance of the estimation technique. The variance around the true value is therefore an indication of the accuracy of an estimator. However, for the application of detecting DSSS signals as considered in Chapters 4 and 5, estimation performance is not a function of how close the estimated value is to the true value. If the actual sequence length is N, an estimated value of $N-1$ instead of $N-2$ does not improve the performance of the DSSS detection algorithm, which depends on the sequence length estimation algorithm. The detection algorithm presented in Chapter 4 requires the correct value N, and any incorrect value will result in the algorithm not functioning (see also Figs. 5.5 and 5.6).

APPENDIX C

PROBABILITY AND LIKELIHOOD

The terms "probability" and "likelihood" are not synonyms in statistical inference [144]. Probability refers to the chance (or odds) of occurrence of an event given the parameters, whereas likelihood refers to the chance of a parameter being or taking on a certain value given the data or outcomes of an experiment. Correct phrases are therefore, for example "probability of an event" and "likelihood of the parameter".

C.1 Probability

Probability is defined in terms of the PDF $p(y|\theta)$ where y represents the data value (or event) and θ the parameter(s). For example, the Gaussian PDF

$$p(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \tag{C.1}$$

expresses the probability of a given data sample value being y, given that the mean and variance have fixed values μ and σ^2. The probability that $y = 0$ (with $\mu = 0$ and $\sigma^2 = 1$) is therefore $1/\sqrt{2\pi}$.

Some example questions to further explain probability are given below.

- Given that the mean is zero and the variance one, what is the probability that the outcome will be between 1.1 and 1.2?
- If a fair coin is flipped 100 times, what is the probability of it landing heads-up every time?
- Given that hypothesis $\mathcal{H}_0$ is true, what is the probability that the received amplitude is greater than 3?

Furthermore, it is also important to distinguish between probability and probability density (or PDF). From Kolmogorov's axioms [110], a probability of an event must always lie in the range $[0, 1]$. A probability density can however grow arbitrarily large, although the integral of the PDF will be one (accounting for all events in the probability space). Thus, (C.1) specifies a density, but the three bulleted examples above are real probabilities.

C.2 Likelihood

Likelihood is determined from the data PDF when the data values are known. The function $p(y|\theta)$ is the PDF of the data y if the parameter value θ is known, and the same function $p(y|\theta)$ is the likelihood function of the parameter θ if the data value y is known. For example, if a statistical experiment (of which the outcome is known to be normally distributed with unity variance) is undertaken and the sample value $y = 0$ is observed, the likelihood function of the mean can be derived from (C.1) as

$$p(y = 0|\mu, \sigma^2 = 1) = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{\mu^2}{2}\right]. \tag{C.2}$$

For this example, the likelihood that $\mu = 0$ is therefore $1/\sqrt{2\pi}$.

Some example questions to further explain likelihood are given below.

- A value of 1.3 has been obtained from the distribution, what is the likelihood of the mean being zero?
- Given that a coin is flipped 100 times and it landed on heads every time, what is the likelihood of the coin being fair?
- Given that the received amplitude is 3.5, what is the likelihood that hypothesis $\mathcal{H}_0$ is true?

APPENDIX D

Q-FUNCTION

D.1 Standard normal distribution

The Q-function is the right-tail probability of the standard normal distribution $\mathcal{N}(0,1)$, which can be expressed as [7]

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

which can be evaluated numerically. However, since the Q-function is often used, it is well tabulated and implemented in several statistical software libraries. The Q-function can also be expressed in terms of the error function [7]

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt \tag{D.2}$$

or the complementary error function [7]

$$\begin{aligned} \operatorname{erfc}(x) &= 1 - \operatorname{erf}(x) && \text{(D.3)} \\ &= \frac{2}{\sqrt{\pi}} \int_x^\infty \exp\left(-t^2\right) dt && \text{(D.4)} \end{aligned}$$

as

$$\begin{aligned} Q(x) &= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) && \text{(D.5)} \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right). && \text{(D.6)} \end{aligned}$$

D.2 General normal distribution

The right-tail probability of the general normal distribution $\mathcal{N}(\mu, \sigma^2)$ can be written as

$$Q_{\mathcal{N}(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^\infty \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] dt \tag{D.7}$$

which should be rewritten in terms of the Q-function given in (D.1) to make use of the software libraries.

By using the substitution

$$u = \frac{t-\mu}{\sigma} \tag{D.8}$$

$$\therefore \frac{du}{dt} = \frac{1}{\sigma} \tag{D.9}$$

$$\therefore dt = \sigma\, du \tag{D.10}$$

in (D.7), the integration limits can be changed as follows

$$t = x \rightarrow u = \frac{x-\mu}{\sigma} \tag{D.11}$$

$$t = \infty \rightarrow u = \infty \tag{D.12}$$

such that (D.7) can be written as

$$Q_{\mathcal{N}(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^\infty \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt \tag{D.13}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\frac{x-\mu}{\sigma}}^\infty \exp\left[-\frac{1}{2}u^2\right] \sigma\, du \tag{D.14}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{x-\mu}{\sigma}}^\infty \exp\left[-\frac{u^2}{2}\right] du \tag{D.15}$$

$$= Q\left(\frac{x-\mu}{\sigma}\right). \tag{D.16}$$

Therefore, the right-tail probability of the RV $X \sim \mathcal{N}(\mu, \sigma^2)$ can be expressed as

$$p(X > \gamma) = Q_{\mathcal{N}(\mu,\sigma^2)}(\gamma) \tag{D.17}$$

$$= p\left(\frac{X-\mu}{\sigma} > \frac{\gamma-\mu}{\sigma}\right) \tag{D.18}$$

$$= Q\left(\frac{\gamma-\mu}{\sigma}\right). \tag{D.19}$$

APPENDIX E

DETECTION OF UNKNOWN DETERMINISTIC SIGNALS

An unknown DSSS signal can be viewed as a random or noise-like signal, especially if the interceptor has no knowledge of the pseudo-random spreading code. In this appendix it is shown that optimal detection (derived using the LRT) of an AWGN signal hidden within AWGN can be achieved by using the energy of the received signal as the test statistic.

E.1 Hypothesis problem

The hypothesis problem can be stated in terms of the received signal $y[n]$ as

$$\mathcal{H}_0 : y[n] = w[n] \tag{E.1}$$

$$\mathcal{H}_1 : y[n] = x[n] + w[n] \tag{E.2}$$

with the sample index $n = 1, 2, \ldots, N$. The background noise is $w[n] \sim \mathcal{N}(0, \sigma_w^2)$, and the signal of interest is $x[n] \sim \mathcal{N}(0, \sigma_x^2)$. The only feature differentiating the two hypotheses is clearly the variance or power content, which can be written as

$$\mathcal{H}_0 : \sigma_y^2 = \sigma_0^2 = \sigma_w^2 \tag{E.3}$$

$$\mathcal{H}_1 : \sigma_y^2 = \sigma_1^2 = \sigma_x^2 + \sigma_w^2 \tag{E.4}$$

since the signal and noise are assumed to be independent.

E.2 Likelihood functions and likelihood ratio

Each of the joint likelihoods $p(\mathbf{y}; \mathcal{H}_0)$ and $p(\mathbf{y}; \mathcal{H}_1)$ can respectively be written as the product of the individual likelihoods $p(y[n]; \mathcal{H}_0)$ and $p(y[n]; \mathcal{H}_1)$ as the received samples (RVs) are i.i.d. [5]. The likelihood functions are therefore given by

$$p(\mathbf{y};\mathcal{H}_0) = \frac{1}{[2\pi\sigma_0^2]^{\frac{N}{2}}} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{n=1}^{N} y^2[n]\right\} \tag{E.5}$$

$$p(\mathbf{y};\mathcal{H}_1) = \frac{1}{[2\pi\sigma_1^2]^{\frac{N}{2}}} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{n=1}^{N} y^2[n]\right\} \tag{E.6}$$

and the log-likelihood ratio (LLR) can then be expressed as

$$\ln \Lambda(\mathbf{y}) = \ln \frac{p(\mathbf{y};\mathcal{H}_1)}{p(\mathbf{y};\mathcal{H}_0)} \tag{E.7}$$

$$= N \ln\left(\frac{\sigma_0}{\sigma_1}\right) + \frac{\sigma_1^2-\sigma_0^2}{2\sigma_0^2\sigma_1^2}\sum_{n=1}^{N} y^2[n]. \tag{E.8}$$

E.3 Test statistic

According to the NP theorem [5] a detection should be declared if the likelihood ratio exceeds a threshold γ, which is determined from the set false alarm rate. Isolating the data-dependent term from (E.8), the test statistic is obtained as

$$T(\mathbf{y}) = \sum_{n=1}^{N} y^2[n] \tag{E.9}$$

which is the energy contained in the received signal y. A detection should be declared if

$$T(\mathbf{y}) > \gamma' \tag{E.10}$$

with γ' the new threshold value, absorbing all the non-data-dependent terms in (E.8), which do not affect the decision region. The detection is therefore based on the energy content of the signal, finding its origin in the multivariate Gaussian distribution.

E.4 Detection performance

In order to calculate the detection performance, the distribution of the test statistic under both hypotheses needs to be determined. The sum of the squares of N independent standard normal RVs $\mathcal{N}(0,1)$ is distributed according to the central Chi-squared distribution with N degrees of freedom [5]. By normalising the variance of each received sample, the

test statistic given in (E.9) can be written as

$$T_{\text{norm}}(\mathbf{y}) = \sum_{n=1}^{N} \left(\frac{y[n]}{\sigma_y} \right)^2 \tag{E.11}$$

$$= \frac{1}{\sigma_y^2} T(\mathbf{y}) \tag{E.12}$$

where $T_{\text{norm}}(\mathbf{y})$ has a central Chi-squared distribution with N degrees of freedom. It can therefore be stated that

$$\frac{T(\mathbf{y})}{\sigma_y^2} \sim \mathcal{X}_N^2. \tag{E.13}$$

The detection performance can therefore be determined from the right-tail probability of the central Chi-squared distribution. The false alarm rate is given as

$$P_{FA} = p(T(\mathbf{y}) > \gamma'; \mathcal{H}_0) \tag{E.14}$$

$$= p\left(\frac{T(\mathbf{y})}{\sigma_0^2} > \frac{\gamma'}{\sigma_0^2} \right) \tag{E.15}$$

$$= Q_{\mathcal{X}_N^2}\left(\frac{\gamma'}{\sigma_0^2} \right) \tag{E.16}$$

from which the optimal threshold level can be calculated as

$$\gamma' = \sigma_0^2 \, Q_{\mathcal{X}_N^2}^{-1}\left(P_{FA}\right). \tag{E.17}$$

Similarly, the detection probability can be shown to be

$$P_D = p(T(\mathbf{y}) > \gamma'; \mathcal{H}_1) \tag{E.18}$$

$$= Q_{\mathcal{X}_N^2}\left(\frac{\gamma'}{\sigma_1^2} \right) \tag{E.19}$$

$$= Q_{\mathcal{X}_N^2}\left(\frac{\sigma_0^2}{\sigma_1^2} \, Q_{\mathcal{X}_N^2}^{-1}\left(P_{FA}\right) \right). \tag{E.20}$$

APPENDIX F

EIGENVALUE MATHEMATICS

This appendix contains further explanation of the eigenvalue decomposition of the SCM of aligned and shifted data matrices, presented in Section 4.3.

F.1 Aligned data matrix

Assume the code sequence $\mathbf{c} = [c_1, c_2, c_3]$ is used to spread the data sequence $\mathbf{d} = [d_1, d_2, d_3]$ and is intercepted and arranged to form the detection matrix (assuming no channel effects)

$$\mathbf{X} = \mathbf{d}^T\mathbf{c} = \begin{bmatrix} d_1c_1 & d_1c_2 & d_1c_3 \\ d_2c_1 & d_2c_2 & d_2c_3 \\ d_3c_1 & d_3c_2 & d_3c_3 \end{bmatrix}. \tag{F.1}$$

The SCM of $\mathbf{X}$ can be expressed as

$$\mathbf{R} = \frac{1}{N}\mathbf{X}^T\mathbf{X} = \alpha_1 \begin{bmatrix} c_1c_1 & c_1c_2 & c_1c_3 \\ c_2c_1 & c_2c_2 & c_2c_3 \\ c_3c_1 & c_3c_2 & c_3c_3 \end{bmatrix} \tag{F.2}$$

with scale factor

$$\alpha_1 = \frac{1}{N}\left(d_1^2 + d_2^2 + d_3^2\right) \tag{F.3}$$

and square matrix dimension $N = 3$.

F.1.1 Eigenvalue calculation

This section shows that aligned matrices of the form given in (F.2) has only one eigenvalue, which is equal to the trace of the matrix. The matrix can be written in the form

$$\mathbf{A} = \begin{bmatrix} k_1c_1 & k_1c_2 & k_1c_3 \\ k_2c_1 & k_2c_2 & k_2c_3 \\ k_3c_1 & k_3c_2 & k_3c_3 \end{bmatrix} \tag{F.4}$$

which is a matrix with all rows (and columns) linearly dependent. Calculating the eigenvalues of $\mathbf{A}$ is equivalent to finding the roots of its characteristic polynomial, defined as [123]

$$p(\lambda) = \det\left(\mathbf{A} - \lambda\mathbf{I}\right). \tag{F.5}$$

The first step in finding the roots is setting the determinant to zero:

$$p(\lambda) = \begin{vmatrix} k_1c_1 - \lambda & k_1c_2 & k_1c_3 \\ k_2c_1 & k_2c_2 - \lambda & k_2c_3 \\ k_3c_1 & k_3c_2 & k_3c_3 - \lambda \end{vmatrix} = 0. \tag{F.6}$$

The values of λ can then be calculated from

$$p(\lambda) = p_1(\lambda) - p_2(\lambda) = 0 \tag{F.7}$$

with

$$\begin{aligned} p_1(\lambda) &= (k_1c_1 - \lambda)(k_2c_2 - \lambda)(k_3c_3 - \lambda) + (k_1c_2)(k_2c_3)(k_3c_1) + (k_2c_1)(k_3c_2)(k_1c_3) \\ &= -\lambda^3 + [k_1c_1 + k_2c_2 + k_3c_3]\lambda^2 - [k_1c_1k_2c_2 + k_1c_1k_3c_3 + k_2c_2k_3c_3]\lambda + 3k_1c_1k_2c_2k_3c_3 \end{aligned} \tag{F.8}$$

and

$$\begin{aligned} p_2(\lambda) &= (k_1c_3)(k_2c_2 - \lambda)(k_3c_1) + (k_1c_2)(k_2c_1)(k_3c_3 - \lambda) + (k_2c_3)(k_3c_2)(k_1c_1 - \lambda) \\ &= -[k_1c_1k_2c_2 + k_1c_1k_3c_3 + k_2c_2k_3c_3]\lambda + 3k_1c_1k_2c_2k_3c_3. \end{aligned} \tag{F.9}$$

By combining (F.7), (F.8) and (F.9), the characteristic polynomial can be written as

$$p(\lambda) = -\lambda^3 + [k_1c_1 + k_2c_2 + k_3c_3]\lambda^2 \tag{F.10}$$

since the lower-order terms sum to zero. By setting (F.10) equal to zero, the only non-zero eigenvalue of $\mathbf{A}$ can directly be found as

$$\lambda = k_1c_1 + k_2c_2 + k_3c_3 \tag{F.11}$$

which is the trace (sum of diagonal elements) of $\mathbf{A}$. From (F.10), it is clear that the other two eigenvalues are zero.

F.1.2 General form

In general, if a code sequence $\mathbf{c} = [c_1, c_2, c_3, \ldots, c_N]$ is used to spread the data sequence $\mathbf{d} = [d_1, d_2, d_3, \ldots, d_N]$ and is intercepted and arranged to form the detection matrix

$$\mathbf{X} = \mathbf{d}^T\mathbf{c} = \begin{bmatrix} d_1c_1 & d_1c_2 & d_1c_3 & \cdots & d_1c_N \\ d_2c_1 & d_2c_2 & d_2c_3 & \cdots & d_2c_N \\ d_3c_1 & d_3c_2 & d_3c_3 & \cdots & d_3c_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_Nc_1 & d_Nc_2 & d_Nc_3 & \cdots & d_Nc_N \end{bmatrix} \tag{F.12}$$

the entry in the i^{th} row and j^{th} column ($i, j \in [1, N]$) of the SCM of $\mathbf{X}$ can be shown to be (following the same approach as in Section F.1)

$$R_{i,j} = \left(\frac{1}{N}\sum_{l=1}^{N} d_l^2\right) c_i c_j. \tag{F.13}$$

The eigenvalues of $\mathbf{R}$ can be found as follows. From linear algebra it is known that the trace of a square matrix $\mathbf{A}$ equals the sum of its eigenvalues [75]

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{N} A_{i,i} = \sum_{i=1}^{N} \lambda_i. \tag{F.14}$$

Since $\mathbf{R}$ is aligned (and all the rows are linearly dependent) it follows that $\mathbf{R}$ will only have one non-zero eigenvalue equal to the trace, which can be expressed using (F.13) and (F.14) as

$$\lambda = \text{tr}(\mathbf{R}) = \sum_{i=1}^{N} R_{i,i} \tag{F.15}$$

$$= \left(\frac{1}{N}\sum_{l=1}^{N} d_l^2\right)\sum_{i=1}^{N} c_i^2. \tag{F.16}$$

F.2 Eigen analysis of non-aligned data matrix

This section will only consider the 3×3 detection matrix for illustration purposes. If the original detection matrix given in (F.1) is cyclically shifted by one chip to the left, the following matrix is formed

$$\mathbf{X}_{-1} = \begin{bmatrix} d_1c_2 & d_1c_3 & d_2c_1 \\ d_2c_2 & d_2c_3 & d_3c_1 \\ d_3c_2 & d_3c_3 & d_1c_1 \end{bmatrix} \tag{F.17}$$

of which the SCM can be shown to be

$$\mathbf{R}_{-1} = \begin{bmatrix} \alpha_1 c_2 c_2 & \alpha_1 c_2 c_3 & \alpha_2 c_1 c_2 \\ \alpha_1 c_2 c_3 & \alpha_1 c_3 c_3 & \alpha_2 c_1 c_3 \\ \alpha_2 c_1 c_2 & \alpha_2 c_1 c_3 & \alpha_1 c_1 c_1 \end{bmatrix} \tag{F.18}$$

with scale factors

$$\alpha_1 = \frac{1}{N}\left(d_1^2 + d_2^2 + d_3^2\right) \tag{F.19}$$

$$\alpha_2 = \frac{1}{N}\left(d_1 d_2 + d_1 d_3 + d_2 d_3\right). \tag{F.20}$$

By following the eigenvalue calculation method of Section F.1.1, the characteristic polynomial of $\mathbf{R}_{-1}$ given in (F.18) can be shown to be

$$p(\lambda) = -\lambda^3 + \alpha_1\left(c_1^2 + c_2^2 + c_3^2\right)\lambda^2 + c_1^2\left(\alpha_2^2 c_3^2 + \alpha_2^2 c_2 c_3 - \alpha_1^2 c_3^2 - \alpha_1^2 c_2^2\right)\lambda. \tag{F.21}$$

Since (F.21) has no constant term (like (F.10)), one root will be $\lambda = 0$. The other two roots can be calculated by solving the quadratic equation

$$-\lambda^2 + \alpha_1\left(c_1^2 + c_2^2 + c_3^2\right)\lambda + c_1^2\left(\alpha_2^2 c_3^2 + \alpha_2^2 c_2 c_3 - \alpha_1^2 c_3^2 - \alpha_1^2 c_2^2\right) = 0. \tag{F.22}$$

Since the sum of the two non-zero eigenvalues of $\mathbf{R}_{-1}$ equals the trace, and the trace remains the same whether the detection matrix $\mathbf{X}$ is aligned or not (compare (F.2) and (F.18)), the largest eigenvalue will be less when $\mathbf{X}$ is not aligned (assuming both eigenvalues are positive). In the aligned case, there is only one non-zero eigenvalue, whereas in the non-aligned case there are two non-zero eigenvalues (see (F.22)).

By cyclically shifting the detection matrix given in (F.17) again by one chip to the left (similar to shifting the original detection matrix given in (F.1) by one chip to the right), the resulting SCM will still have two non-zero eigenvalues. By repeatedly performing cyclic shifting on the detection matrix $\mathbf{X}$, the largest eigenvalue of the SCM will reach its maximum (equal to the matrix trace) every time $\mathbf{X}$ is aligned. The resulting eigenvalue sequence will therefore be a periodic sequence, reaching its peak value once every N cyclic shifts.

It can be assumed that all eigenvalues of $\mathbf{R}$ (for any number of cyclic shifts of $\mathbf{X}$) are real ($\mathbf{R}$ is symmetric) and non-negative ($\mathbf{R}$ is positive semi-definite) [5]. Since $\mathbf{R}$ is of the form $k\mathbf{X}^T\mathbf{X}$, it is symmetric ($\mathbf{R}^T = \mathbf{R}$) since $\left[k\mathbf{X}^T\mathbf{X}\right]^T = k\mathbf{X}^T\mathbf{X}$. Furthermore, $\mathbf{R}$ is positive semi-definite since it can be shown that for any non-zero column vector $\mathbf{z}$ [114]

$$\mathbf{z}^T\mathbf{X}^T\mathbf{X}\mathbf{z} \geq 0. \tag{F.23}$$

APPENDIX G

EIGENVALUE BOUNDS

This appendix contains further explanation of the eigenvalue bounds presented in Sections 4.3.4.1 and 5.4.3.1. The eigenvalues of the matrix $\mathbf{H} = N\mathbf{R} = \mathbf{X}^T\mathbf{X}$ presented here were obtained using the *Maxima* numerical analysis software package [145]. The $N \times N$ DSSS data matrix $\mathbf{X}$ is given by

$$\mathbf{X} = \mathbf{d}^T\mathbf{c} = \begin{bmatrix} d_1c_1 & d_1c_2 & d_1c_3 & \cdots & d_1c_N \\ d_2c_1 & d_2c_2 & d_2c_3 & \cdots & d_2c_N \\ d_3c_1 & d_3c_2 & d_3c_3 & \cdots & d_3c_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_Nc_1 & d_Nc_2 & d_Nc_3 & \cdots & d_Nc_N \end{bmatrix} \tag{G.1}$$

with the spreading code $\mathbf{c} = [c_1, c_2, \ldots, c_N]$ and data sequence $\mathbf{d} = [d_1, d_2, \ldots, d_N]$, each having antipodal elements $c, d \in \pm 1$. The data matrix $\mathbf{X}$ is cyclically shifted continually by moving all row elements to the left, with the first element in each row moving to the last position of the row above. The upper left matrix element is removed, and the lower right element is fed with a new spread bit $d_{N+1}\mathbf{c}$. The eigenvalues of $\mathbf{H} = \mathbf{X}^T\mathbf{X}$ is then calculated for each cyclic shift of $\mathbf{X}$.

Subsequently, the eigenvalues of $\mathbf{H}$ for $N = 3$ to 7 are considered for cyclic shifts of $\mathbf{X}$ from one shift to $N - 1$ shifts. From N shifts the cycle repeats such that N shifts correspond to no shift, $N + 1$ shifts to 1 shift, etc. The underlying pattern is then analysed to obtain a general expression for the eigenvalues in order to find the limits of the largest two eigenvalues λ_1 and λ_2 over the range of cyclic shifts of $\mathbf{X}$ considered.

G.1 3 by 3 matrix

G.1.1 No shift or shifts of integer multiples of N

Using (F.16) the only non-zero eigenvalue of $\mathbf{R}$ can be calculated as

$$\lambda(\mathbf{R}) = \left(\frac{1}{N}\sum_{l=1}^{N} d_l^2\right)\sum_{i=1}^{N} c_i^2 \tag{G.2}$$

$$= \left(\frac{1}{3}\sum_{l=1}^{3} 1\right)\sum_{i=1}^{3} 1 \tag{G.3}$$

$$= 3. \tag{G.4}$$

The corresponding eigenvalue of $\mathbf{H} = N\mathbf{R}$ is therefore $\lambda_1 = 9$. The eigenvalues obtained using *Maxima* are: $\lambda_1 = 9$ and $\lambda_2 = \lambda_3 = 0$.

G.1.2 One or two shifts

The eigenvalues of $\mathbf{H}$ for both one or two shifts of $\mathbf{X}$ are of the form

$$\lambda_1 = \frac{9 + \sqrt{\Delta}}{2} \tag{G.5}$$

$$\lambda_2 = \frac{9 - \sqrt{\Delta}}{2} \tag{G.6}$$

$$\lambda_3 = 0 \tag{G.7}$$

with the discriminant

$$\Delta = (16\, d_1\, d_2\, d_3 + 16\, d_2)\; d_4 + 16\, d_1\, d_3 + 33 \tag{G.8}$$

obtained through simplification using the fact that all data bits $d = \pm 1$ and all code chips $c = \pm 1$. (The simplification follows by replacing all occurrences of c^k and d^k with k even in Δ with 1.) By considering all possible values and combinations of the data bits, the discriminant is found to be within the range $\Delta \in [17, 81]$.

G.2 4 by 4 matrix

G.2.1 No shift

Similar as in Section G.1.1, the eigenvalues of $\mathbf{H}$ obtained using *Maxima* for $N = 4$ are: $\lambda_1 = 16$ and $\lambda_2 = \lambda_3 = \lambda_4 = 0$.

G.2.2 One or three shifts

The non-zero eigenvalues of **H** are of the form

$$\lambda_1 = \frac{16 + \sqrt{\Delta}}{2} \tag{G.9}$$

$$\lambda_2 = \frac{16 - \sqrt{\Delta}}{2} \tag{G.10}$$

with the discriminant

$$\Delta = ((24\,d_2\,d_3 + 24\,d_1\,d_2)\;d_4 + 24\,d_3)\;d_5 + (24\,d_1\,d_2\,d_3 + 24\,d_2)\;d_4 + 24\,d_1\,d_3 + 112 \tag{G.11}$$

with range $\Delta \in [64;\ 256]$ obtained as described in Section G.1.2.

G.2.3 Two shifts

The form of the non-zero eigenvalues of **H** is identical to (G.9) and (G.10), though the discriminant is

$$\Delta = ((32\,d_2\,d_3 + 32\,d_1\,d_2)\;d_4 + 32\,d_3)\;d_5 + (32\,d_1\,d_2\,d_3 + 32\,d_2)\;d_4 + 32\,d_1\,d_3 + 64 \tag{G.12}$$

with range $\Delta \in [0;\ 256]$ obtained as described in Section G.1.2.

G.3 5 by 5 matrix

G.3.1 No shift

Similar as in Section G.1.1, the largest eigenvalue of **H** obtained using *Maxima* for $N = 5$ is $\lambda_1 = 25$, while the remaining four (λ_2 to λ_5) are zero.

G.3.2 One or four shifts

The non-zero eigenvalues of **H** are of the form

$$\lambda_1 = \frac{25 + \sqrt{\Delta}}{2} \tag{G.13}$$

$$\lambda_2 = \frac{25 - \sqrt{\Delta}}{2} \tag{G.14}$$

with the discriminant

$$\Delta = ((32\,d_3\,d_4 + 32\,d_2\,d_3 + 32\,d_1\,d_2)\;d_5 + 32\,d_4)\,d_6 + ((32\,d_2\,d_3 + 32\,d_1\,d_2)\;d_4 + 32\,d_3)\,d_5 + (32\,d_1\,d_2\,d_3 + 32\,d_2)\;d_4 + 32\,d_1\,d_3 + 305 \tag{G.15}$$

with range $\Delta \in [241;\ 625]$ obtained as described in Section G.1.2.

G.3.3 Two or three shifts

The form of the non-zero eigenvalues of $\mathbf{H}$ is identical to (G.13) and (G.14), though the discriminant is

$$\Delta = ((48\,d_3\,d_4 + 48\,d_2\,d_3 + 48\,d_1\,d_2)\;d_5 + 48\,d_4)\,d_6 + ((48\,d_2\,d_3 + 48\,d_1\,d_2)\;d_4 + 48\,d_3)\,d_5 + (48\,d_1\,d_2\,d_3 + 48\,d_2)\;d_4 + 48\,d_1\,d_3 + 145$$

with range $\Delta \in [49;\ 625]$ obtained as described in Section G.1.2.

G.4 6 by 6 matrix

G.4.1 No shift

Similar as in Section G.1.1, the largest eigenvalue of $\mathbf{H}$ obtained using *Maxima* for $N = 6$ is $\lambda_1 = 36$, while the remaining five (λ_2 to λ_6) are zero.

G.4.2 One or five shifts

The non-zero eigenvalues of $\mathbf{H}$ are of the form

$$\lambda_1 = \frac{36 + \sqrt{\Delta}}{2} \tag{G.16}$$

$$\lambda_2 = \frac{36 - \sqrt{\Delta}}{2} \tag{G.17}$$

with the discriminant

$$\begin{aligned}\Delta = {} & ((40\,d_4\,d_5 + 40\,d_3\,d_4 + 40\,d_2\,d_3 + 40\,d_1\,d_2)\;d_6 + 40\,d_5)\;d_7 \\ & + ((40\,d_3\,d_4 + 40\,d_2\,d_3 + 40\,d_1\,d_2)\;d_5 + 40\,d_4)\;d_6 \\ & + ((40\,d_2\,d_3 + 40\,d_1\,d_2)\;d_4 + 40\,d_3)\;d_5 \\ & + (40\,d_1\,d_2\,d_3 + 40\,d_2)\;d_4 + 40\,d_1\,d_3 + 696\end{aligned}$$

with range $\Delta \in [576;\ 1296]$ obtained as described in Section G.1.2.

G.4.3 Two or four shifts

The form of the non-zero eigenvalues of **H** is identical to (G.16) and (G.17), though the discriminant is

$$\begin{aligned}\Delta = &((64\,d_4\,d_5 + 64\,d_3\,d_4 + 64\,d_2\,d_3 + 64\,d_1\,d_2)\;d_6 + 64\,d_5)\;d_7\\ &+ ((64\,d_3\,d_4 + 64\,d_2\,d_3 + 64\,d_1\,d_2)\;d_5 + 64\,d_4)\;d_6\\ &+ ((64\,d_2\,d_3 + 64\,d_1\,d_2)\;d_4 + 64\,d_3)\;d_5\\ &+ (64\,d_1\,d_2\,d_3 + 64\,d_2)\;d_4 + 64\,d_1\,d_3 + 336\end{aligned}$$

with range $\Delta \in [144;\ 1296]$ obtained as described in Section G.1.2.

G.4.4 Three shifts

The form of the non-zero eigenvalues of **H** is identical to (G.16) and (G.17), though the discriminant is

$$\begin{aligned}\Delta = &((72\,d_4\,d_5 + 72\,d_3\,d_4 + 72\,d_2\,d_3 + 72\,d_1\,d_2)\;d_6 + 72\,d_5)\;d_7\\ &+ ((72\,d_3\,d_4 + 72\,d_2\,d_3 + 72\,d_1\,d_2)\;d_5 + 72\,d_4)\;d_6\\ &+ ((72\,d_2\,d_3 + 72\,d_1\,d_2)\;d_4 + 72\,d_3)\;d_5\\ &+ (72\,d_1\,d_2\,d_3 + 72\,d_2)\;d_4 + 72\,d_1\,d_3 + 216\end{aligned}$$

with range $\Delta \in [0;\ 1296]$ obtained as described in Section G.1.2.

G.5 7 by 7 matrix

G.5.1 No shift

Similar as in Section G.1.1, the largest eigenvalue of **H** obtained using *Maxima* for $N = 7$ is $\lambda_1 = 49$, while the remaining six (λ_2 to λ_7) are zero.

G.5.2 One or six shifts

The non-zero eigenvalues of **H** are of the form

$$\lambda_1 = \frac{49 + \sqrt{\Delta}}{2} \tag{G.18}$$

$$\lambda_2 = \frac{49 - \sqrt{\Delta}}{2} \tag{G.19}$$

with the discriminant

$$\begin{aligned}\Delta = &((48\,d_5\,d_6 + 48\,d_4\,d_5 + 48\,d_3\,d_4 + 48\,d_2\,d_3 + 48\,d_1\,d_2)\;d_7 + 48\,d_6)\;d_8\\ &+ ((48\,d_4\,d_5 + 48\,d_3\,d_4 + 48\,d_2\,d_3 + 48\,d_1\,d_2)\;d_6 + 48\,d_5)\,d_7\\ &+ ((48\,d_3\,d_4 + 48\,d_2\,d_3 + 48\,d_1\,d_2)\;d_5 + 48\,d_4)\;d_6\\ &+ ((48\,d_2\,d_3 + 48\,d_1\,d_2)\;d_4 + 48\,d_3)\;d_5\\ &+ (48\,d_1\,d_2\,d_3 + 48\,d_2)\;d_4 + 48\,d_1\,d_3 + 1393\end{aligned}$$

with range $\Delta \in [1249;\ 2401]$ obtained as described in Section G.1.2.

G.5.3 Two or five shifts

The form of the non-zero eigenvalues of $\mathbf{H}$ is identical to (G.18) and (G.19), though the discriminant is

$$\begin{aligned}\Delta = &((80\,d_5\,d_6 + 80\,d_4\,d_5 + 80\,d_3\,d_4 + 80\,d_2\,d_3 + 80\,d_1\,d_2)\;d_7 + 80\,d_6)\;d_8\\ &+ ((80\,d_4\,d_5 + 80\,d_3\,d_4 + 80\,d_2\,d_3 + 80\,d_1\,d_2)\;d_6 + 80\,d_5)\,d_7\\ &+ ((80\,d_3\,d_4 + 80\,d_2\,d_3 + 80\,d_1\,d_2)\;d_5 + 80\,d_4)\;d_6\\ &+ ((80\,d_2\,d_3 + 80\,d_1\,d_2)\;d_4 + 80\,d_3)\;d_5\\ &+ (80\,d_1\,d_2\,d_3 + 80\,d_2)\;d_4 + 80\,d_1\,d_3 + 721\end{aligned}$$

with range $\Delta \in [481;\ 2401]$ obtained as described in Section G.1.2.

G.5.4 Three or four shifts

The form of the non-zero eigenvalues of $\mathbf{H}$ is identical to (G.18) and (G.19), though the discriminant is

$$\begin{aligned}\Delta = &((96\,d_5\,d_6 + 96\,d_4\,d_5 + 96\,d_3\,d_4 + 96\,d_2\,d_3 + 96\,d_1\,d_2)\;d_7 + 96\,d_6)\;d_8\\ &+ ((96\,d_4\,d_5 + 96\,d_3\,d_4 + 96\,d_2\,d_3 + 96\,d_1\,d_2)\;d_6 + 96\,d_5)\;d_7\\ &+ ((96\,d_3\,d_4 + 96\,d_2\,d_3 + 96\,d_1\,d_2)\;d_5 + 96\,d_4)\;d_6\\ &+ ((96\,d_2\,d_3 + 96\,d_1\,d_2)\;d_4 + 96\,d_3)\;d_5\\ &+ (96\,d_1\,d_2\,d_3 + 96\,d_2)\;d_4 + 96\,d_1\,d_3 + 385\end{aligned}$$

with range $\Delta \in [97;\ 2401]$ obtained as described in Section G.1.2.

G.6 General form and bounds on λ_1 and λ_2

As illustrated in Sections G.1 to G.5 and discussed in Section 4.3.3, the matrix $\mathbf{H}$ and consequently the SCM $\mathbf{R}$ has at most two non-zero eigenvalues, irrespective of the value of N and how many cyclic shifts are applied to the data matrix $\mathbf{X}$. From the expressions derived in Sections G.1 to G.5, the general form of the largest two eigenvalues of $\mathbf{H}$ can be expressed as

$$\lambda = \frac{N^2 \pm \sqrt{\Delta}}{2} \tag{G.20}$$

which can be understood from the fact that the simplified characteristic polynomial of $\mathbf{H}$ has a maximum order of two (see e.g. (F.21) and (F.22)). The roots of the characteristic polynomial can therefore be calculated using the quadratic formula, which resembles (G.20). The ranges for the discriminant value Δ in (G.20), found in Sections G.1 to G.5, are summarised in Table G.1. The different values for the minimum correspond to different numbers of cyclic shifts applied to $\mathbf{X}$.

Table G.1: Ranges of discriminant values.

N	$\Delta_{\min}$	$\Delta_{\max}$
3	17	81
4	min(0, 64)	256
5	min(49, 241)	625
6	min(0, 144, 576)	1296
7	min(97, 481, 1249)	2401

From Table G.1 it is clear that

$$\Delta \in \begin{cases} \left[0;\ N^4\right] & (N \text{ even}) \\ \\ \left[2N^2 - 1;\ N^4\right] & (N \text{ odd}) \end{cases} \tag{G.21}$$

from which the upper and lower limits of λ_1 and λ_2 can be calculated.

G.6.1 Upper and lower limits of λ_1

The upper limit of the largest eigenvalue can be expressed using (G.20) and (G.21) as

$$\begin{aligned} \lambda_{1,\max} &= \frac{N^2 + \sqrt{\Delta_{\max}}}{2} \\ &= \frac{N^2 + \sqrt{N^4}}{2} \\ &= N^2. \end{aligned}$$

The lower limit of the largest eigenvalue can be expressed using (G.20) and (G.21) as

$$\begin{aligned}\lambda_{1,\min} &= \frac{N^2+\sqrt{\Delta_{\min}}}{2} \\ &= \begin{cases} \dfrac{N^2}{2} & (N \text{ even}) \\ \dfrac{N^2+\sqrt{2N^2-1}}{2} & (N \text{ odd}) \end{cases}\end{aligned}$$

G.6.2 Upper and lower limits of λ_2

The upper limit of the second largest eigenvalue can be expressed using (G.20) and (G.21) as

$$\begin{aligned}\lambda_{2,\max} &= \frac{N^2-\sqrt{\Delta_{\min}}}{2} \\ &= \begin{cases} \dfrac{N^2}{2} & (N \text{ even}) \\ \dfrac{N^2-\sqrt{2N^2-1}}{2} & (N \text{ odd}) \end{cases}\end{aligned}$$

The lower limit of the second largest eigenvalue can be expressed using (G.20) and (G.21) as

$$\begin{aligned}\lambda_{2,\min} &= \frac{N^2-\sqrt{\Delta_{\max}}}{2} \\ &= \frac{N^2-\sqrt{N^4}}{2} \\ &= 0.\end{aligned}$$

G.6.3 Summary

Using the upper and lower limits given in Sections G.6.1 and G.6.2, the eigenvalue bounds can be summarised as follows.

$$\lambda_1 \in \begin{cases} \left[\frac{N^2}{2};\ N^2\right] & (N \text{ even}) \\ \left[\frac{N^2+\sqrt{2N^2-1}}{2};\ N^2\right] & (N \text{ odd}) \end{cases} \tag{G.22}$$

$$\lambda_2 \in \begin{cases} \left[0;\ \frac{N^2}{2}\right] & (N \text{ even}) \\ \left[0;\ \frac{N^2-\sqrt{2N^2-1}}{2}\right] & (N \text{ odd}) \end{cases} \tag{G.23}$$

G.6.4 Examples

Figs. G.1 and G.2 show example eigenvalue sequences of $\mathbf{H} = \mathbf{X}^T\mathbf{X}$ when $\mathbf{X}$ (defined in (G.1)) is cyclically shifted continuously. The spreading sequence used in Fig. G.1 is given by

$$\mathbf{c} = [1, -1, 1, -1, -1, -1, -1, 1, -1, -1, -1, 1, 1, 1, 1] \tag{G.24}$$

with $N = 15$ to illustrate the case where N is odd. The spreading sequence used in Fig. G.2 is given by

$$\mathbf{c} = [1, -1, 1, -1, -1, -1, -1, 1, -1, -1, -1, 1, 1, 1, 1, 1] \tag{G.25}$$

with $N = 16$ to illustrate the case where N is even. The cyclic shift parameter τ shown in the two figures indicates the number of cyclic shifts applied to $\mathbf{X}$. The upper and lower bounds given in (G.22) and (G.23) are also shown on the graphs.

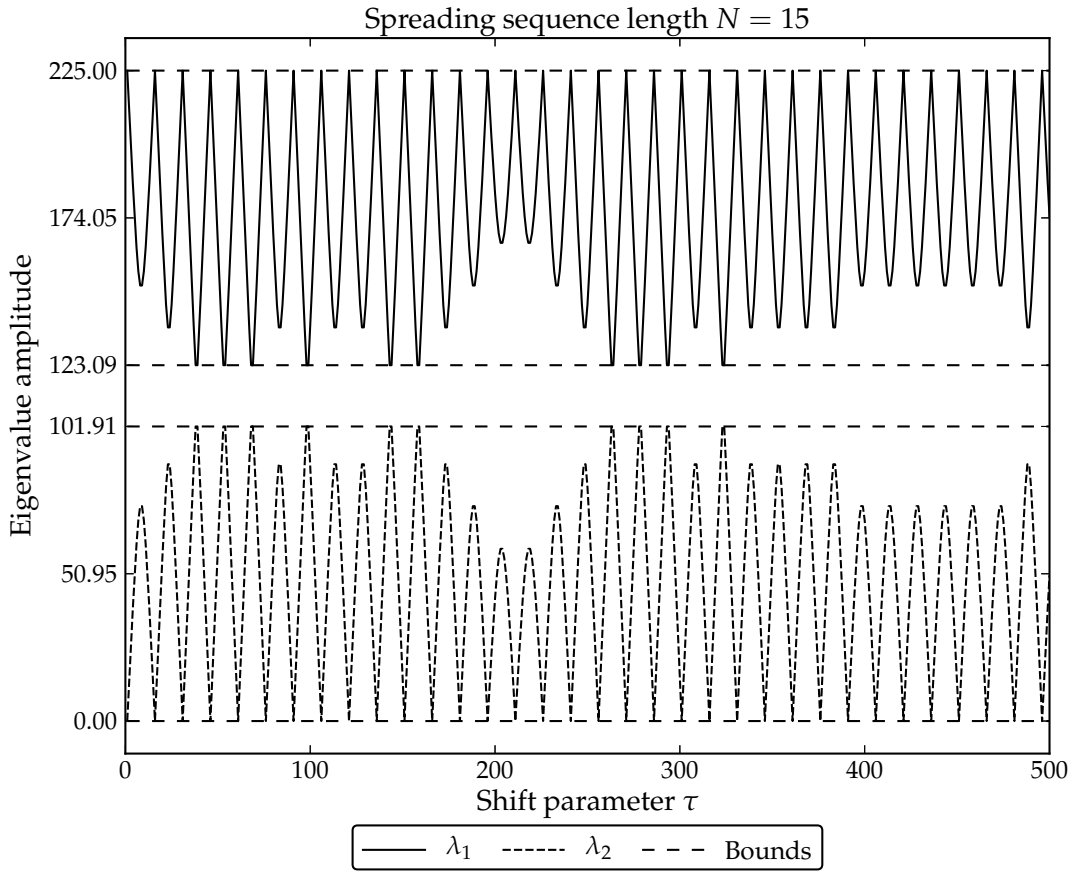


Figure G.1: Eigenvalue sequences for odd sequence length.

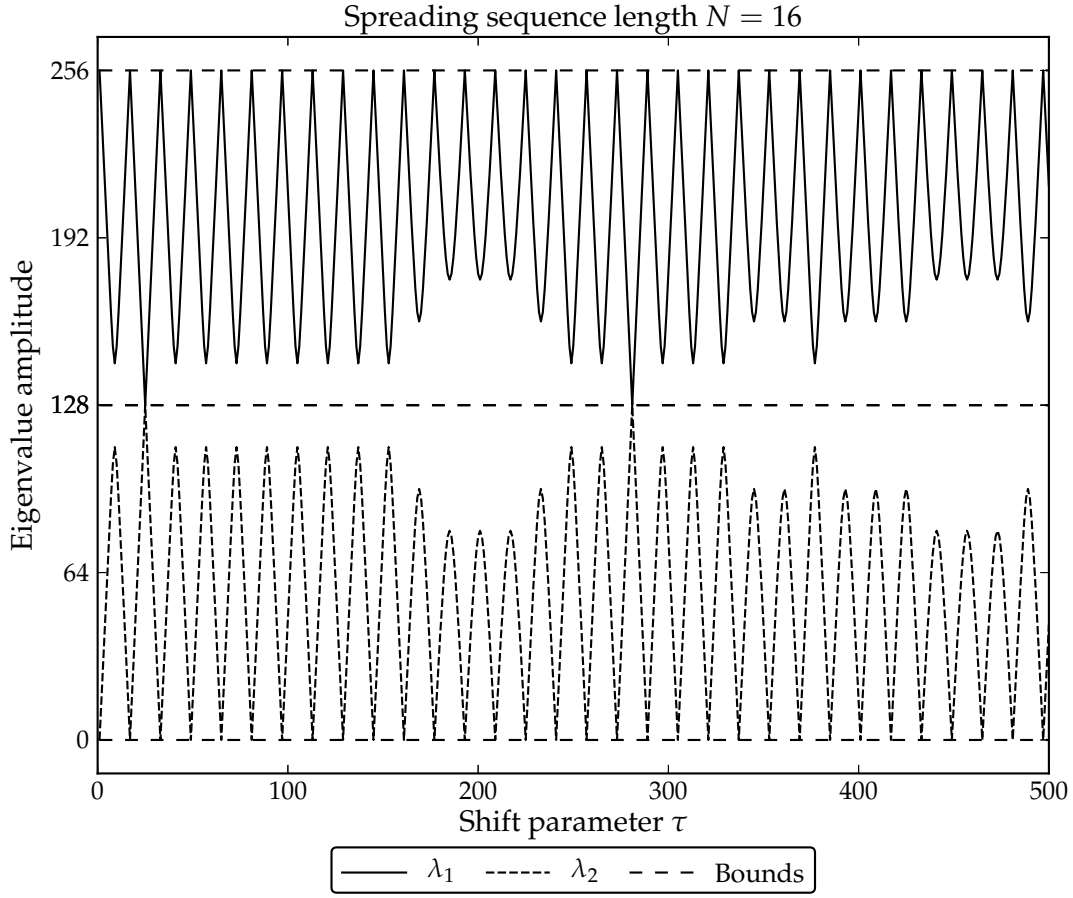


Figure G.2: Eigenvalue sequences for even sequence length.

APPENDIX H

STATISTICAL DERIVATIONS

This appendix considers statistical derivations of importance mostly for Chapter 5.

H.1 Variance

The variance of a RV x can be expressed as [26]

$$\begin{aligned}
\sigma_x^2 = \sigma^2(x) &= E\left[(x - E[x])^2\right] && \text{(H.1)}\\
&= E\left[x^2 - 2xE[x] + (E[x])^2\right] && \text{(H.2)}\\
&= E\left[x^2\right] - 2E\left[xE[x]\right] + (E[x])^2 && \text{(H.3)}\\
&= E\left[x^2\right] - 2E[x]E[x] + (E[x])^2 && \text{(H.4)}\\
&= E\left[x^2\right] - (E[x])^2 && \text{(H.5)}\\
&= E\left[x^2\right] - \mu_x^2. && \text{(H.6)}
\end{aligned}$$

The variance of cx, with c a constant value, can be expressed as

$$\begin{aligned}
\sigma^2(cx) &= E\left[(cx)^2\right] - (E[cx])^2 && \text{(H.7)}\\
&= c^2 E\left[x^2\right] - c^2 (E[x])^2 && \text{(H.8)}\\
&= c^2\sigma^2(x). && \text{(H.9)}
\end{aligned}$$

H.2 Statistics of the product of two independent RVs

This section considers the statistics (mean μ and variance σ^2) of the product of two independent RVs. The product of two i.i.d. (normally distributed) RVs $w_n w_{n+k}$ was considered in Section 5.3.3.2 (see Equation (5.14)).

The expected value of $Z = XY$, with RVs X and Y independent, can be written as

$$E[Z] = E[XY] = E[X]E[Y] \quad \text{(H.10)}$$

$$\therefore \mu_z = \mu_x\mu_y. \quad \text{(H.11)}$$

By combining (H.6) and (H.11), the variance of Z can be expressed as

$$\begin{aligned}
\sigma_z^2 &= E[Z^2] - (E[Z])^2 && \text{(H.12)}\\
&= E[X^2Y^2] - (E[XY])^2 && \text{(H.13)}\\
&= E[X^2]E[Y^2] - (E[X]E[Y])^2 && \text{(H.14)}\\
&= (\sigma_x^2 + \mu_x^2)(\sigma_y^2 + \mu_y^2) - (\mu_x^2\mu_y^2) && \text{(H.15)}\\
&= \sigma_x^2\sigma_y^2 + \sigma_x^2\mu_y^2 + \sigma_y^2\mu_x^2 + \mu_x^2\mu_y^2 - (\mu_x^2\mu_y^2) && \text{(H.16)}\\
&= \sigma_x^2\sigma_y^2 + \sigma_x^2\mu_y^2 + \sigma_y^2\mu_x^2. && \text{(H.17)}
\end{aligned}$$

Equations (H.11) and (H.17) therefore provide the relationships between the means and variances of the product Z and the two RVs X and Y.