A SEQUENTIAL CLASSIFICATION ALGORITHM FOR AUTOREGRESSIVE PROCESSES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÜNEŞ OTLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

SEPTEMBER 2011

Approval of the thesis:

A SEQUENTIAL CLASSIFICATION ALGORITHM FOR AUTOREGRESSIVE PROCESSES

submitted by **GÜNEŞ OTLU** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. İsmet Erkmen
Head of Department, **Electrical and Electronics Engineering** _____

Assoc. Prof. Dr. Çağatay Candan
Supervisor, **Electrical and Electronics Engineering Dept., METU** _____

Assoc. Prof. Dr. Tolga Çiloğlu
Co-supervisor, **Electrical and Electronics Engineering Dept., METU** _____

Examining Committee Members:

Prof. Dr. Mete Severcan
Electrical and Electronics Engineering Dept., METU _____

Prof. Dr. S. Sencer Koç
Electrical and Electronics Engineering Dept., METU _____

Assoc. Prof. Dr. Tolga Çiloğlu
Electrical and Electronics Engineering Dept., METU _____

Assoc. Prof. Dr. Çağatay Candan
Electrical and Electronics Engineering Dept., METU _____

Dr. Ülkü Doyuran
ASELSAN Inc. _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: GÜNEŞ OTLU

Signature :

ABSTRACT

A SEQUENTIAL CLASSIFICATION ALGORITHM FOR AUTOREGRESSIVE PROCESSES

Otlu, Güneş

M.Sc., Department of Electrical and Electronics Engineering

Supervisor : Assoc. Prof. Dr. Çağatay Candan

Co-Supervisor : Assoc. Prof. Dr. Tolga Çiloğlu

September 2011, 70 pages

This study aims to present a sequential method for the classification of the autoregressive processes. Different from the conventional detectors having fixed sample size, the method uses Wald's sequential probability ratio test and has a variable sample size. It is shown that the suggested method produces the classification decisions much earlier than fixed sample size alternative on the average. The proposed method is extended to the case when processes have unknown variance. The effects of the unknown process variance on the algorithm performance are examined. Finally, the suggested algorithm is applied to the classification of fixed and rotary wing targets. The average detection time and its relation with signal to noise ratio are examined.

Keywords: Sequential Detection, Autoregressive Modeling, Target classification, Wald, SPRT

ÖZ

ÖZBAĞLANIMLI SÜREÇLER İCİN DİZİSEL SINIFLANDIRMA ALGORİTMASI

Otlu, Güneş
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Doç. Dr. Çağatay Candan
Ortak Tez Yöneticisi : Doç. Dr. Tolga Çiloğlu
Eylül 2011, 70 sayfa

Bu çalışma, özbağlanımlı süreçler için dizisel bir sınıflandırma algoritması sunmayı amaçlamaktadır. Genellikle kullanılan sabit boyutlu algılayıcıların aksine, bu metot değişken örnek boyutlarına sahip Wald' un dizisel olasılık oran testini kullanmaktadır. Önerilen metodun sınıflandırma kararını ortalama olarak sabit örnek boyutuna sahip alternatifinden daha önce verdiği gösterilmiştir. Önerilen metot, süreçlerin varyansının bilinmediği durum için genişletilmiştir. Bilinmeyen süreç varyansının algoritma performansı üzerindeki etkileri incelenmiştir. Son olarak, önerilen algoritma sabit ve dönen kanatlı hedeflerin sınıflandırmasına uygulanmıştır. Ortalama karar zamanı ve bunun sinyal gürültü oranıyla bağlantısı araştırılmıştır.

Anahtar Kelimeler: Dizisel Tespit, Otoregresif Modelleme, Hedef Sınıflandırma, Wald, SPRT

to my family...

ACKNOWLEDGMENTS

I thank to my supervisor, Çağatay Candan for his guidance and support. He was encouraged me at every level of this work. I am very grateful to him for his patience during our research meetings.

I am also thankful to ASELSAN and METU-EE members who have collected the experimental data used. I would like to express my deep and sincere respect to my co-supervisor Tolga Çiloğlu and the thesis jury members Mete Severcan, Sencer Koç and Ülkü Doyuran.

I would like to thank to my department Hardware Design Engineering in ASELSAN for letting me to involve in this thesis work. Also I would like to thank to TÜBİTAK (The Scientific and Technical Research Council of Turkey) for their support during my M. Sc. studies with their scholarship.

Finally, I owe my thanks to my family, my beloved wife (Duygu Otlu), my loving mother (Muhterem Otlu), my dear father (Fehmi Otlu) and my dear brother (Deniz Otlu) for their undying love, encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AR	:	Autoregressive
ARMA	:	Autoregressive Moving Average
ASN	:	Average Sample Number
CWGN	:	Complex White Gaussian Noise
FSS	:	Fixed Sample Size
MA	:	Moving Average
PDF	:	Probability Density Function
PSD	:	Power Spectrum Density
RCS	:	Radar Cross Section
SNR	:	Signal to Noise Ratio
SPRT	:	Sequential Probability Ratio Test
WGN	:	White Gaussian Noise

CHAPTER 1

INTRODUCTION

1.1 Motivation and Background

Modern sensor systems can detect targets with low latency at all ranges and can work properly under high noise and interference levels. The delay in the detection decision is important in initiating the counter measures to the targets. A reduced delay is also helpful in scheduling of sensor modes. As an example, a typical surveillance system can enlarge its coverage area by making target present or absent decisions in reduced time. The method of detection with low delay is desired to work at noisy environments and in the long ranges. The detection of radar targets has been studied extensively in the literature, but lesser attention is paid to the detection delay.

The target detection is typically performed on multiple observation samples. The detection with fixed sample size (FSS) is the most well known procedure. This detection procedure collects a predefined number of samples first and once the sample collection period is over, processes them to make a decision on the target presence. The sample size is determined by the desired performance requirements, [2]. The sequential probability ratio test (SPRT) is another effective procedure for the multiple-sample detection problem which is described originally by Wald, [3]. The sample size of this procedure is variable. In other words, the termination time of the test changes with the input. For two hypothesis case, this procedure is optimal in the sense that the decision sequence ends with the minimum number of samples (on the average) to achieve a desired probability of error, [4]. It is shown that the SPRT needs much fewer samples than the procedures having fixed sample sizes, [5].

Since the sequential probability ratio test does not require the sample size selection in advance,

it is also very suitable for online processing. The decision to terminate the process depends on the scores generated by the observations collected up to that instant. The average sample number (ASN) is defined as the mean number of sample points to successfully meet the desired probability of errors of first and second types. With respect to ASN measure, it is known that the sequential probability ratio test usually terminates in an about 50 per cent savings in the number of observations over the most efficient test procedure having a fixed sample size, [3].

The detection time is a random variable in this sequential procedure. As expected, the targets that are difficult to discriminate have to be observed for longer periods while the classification of the targets that are easier to discriminate requires less time. The sensor and the other computational resources can be utilized more efficiently because of this behavior of the sequential approach. So that an overall enhancement in the detection performance can be observed.

The SPRT is also extensively used in the problem of change detection. This problem can be defined as the change of the parameter θ of the time series from θ_0 to $\theta_1 \neq \theta_0$ at an unknown change time. The main goal is to detect and estimate the change in one or more parameters of interest. Changes can be grouped as additive and nonadditive (spectral) changes. Additive changes take place in the mean value of the observation. Nonadditive case results in changes in the variance, correlations, spectral characteristics or dynamics of the signal, [6]. Also these changes can be analyzed as changes in a regression model, ARMA model or in a state-space model by using different modeling issues. The online change detection algorithms use the SPRT by processing the online data streams. With the aid of the Wald's SPRT, the cumulative sum (CUSUM) algorithm was proposed in [7]. CUSUM method uses the idea that integration of the probability ratios of the signal with adaptive thresholds and it is frequently used in the online change detection algorithms for detection of a change in a parameter in AR modeling, [8], [9], [10]. In addition CUSUM method can be used for sequential detection of a target in clutter, [11]. Recursive algorithms are also used for the detection of a change in the autoregressive processes, [12], [13].

In a typical detection problem, some target specific parameters are required to implement the classification algorithm. The required parameters can be obtained by fitting the data to available analytical models. For modeling time series, linear regression models are often used. The general form of the linear regression models includes Finite Impulse Response

(FIR), Auto-Regressive (AR) and Auto-Regressive Moving Average (ARMA) models as some special cases. The AR model is useful for modeling time series and often used in radar systems (clutter modeling), human EEG, earthquake analysis, speech segmentation areas, [14].

In many applications, the autoregressive modeling is preferred to another widely used method called as periodogram. AR modeling is utilized in the analysis of spectrum because of its better spectral resolution, [15]. Therefore, in areas requiring high spectral resolution, such as determining the doppler spectrum, AR or ARMA models can be preferred in comparison to simple periodogram, that is the windowed fast Fourier transformation, followed by magnitude squaring and averaging, [16]. In addition, AR modeling has low complexity for spectrum estimation, [17].

The methods of sequential detection and AR models have been proposed for the problems of underwater acoustics. The sequential detector based on Page test is applied to active sonar detection in [18]. In passive acoustic detection, AR modeling is used in the interference cancelation procedure, [19], [20]. In addition, in active sonar detection adaptive prewhiteners can be designed by using the autoregressive methods, [21], [22].

In this thesis, we apply the methods of sequential detection with AR models to the problem of plane - helicopter classification. It is well known that a doppler frequency shift of carrier frequency proportional to target range rate of the target is observed on the radar echo. In addition, the rotating objects such as the rotor blades of the helicopter causes a modulation on the radar return signal. The analysis of moving and rotary target effects is given in [23], [24], [25], [26]. The work on utilizing the effect of the main rotor in detection is given in [27], [1]. The discrimination of helicopters from fixed-wing targets is given in [28], [29]. To the best of our knowledge, the SPRT with AR modeled spectrums has not been applied to the problem of plane - helicopter (fixed wing - rotary wing) classification.

In [30], a general solution to the target discrimination problem using SPRT and AR models is presented. The suggested method, different from the work presented here, uses two mean square prediction filters. The likelihood ratio is composed of the probability density functions of the prediction errors for each hypothesis. It is shown that this structure is optimal for the signatures with Gaussian distribution. It is also stated that the structure can be applied to the signatures with non-Gaussian distributions but its optimality is lost.

Lastly in [31], the binary hypothesis test, as in here, is chosen as two different AR processes. The order of the AR model is chosen as 2 and the process coefficients are complex valued. The simulations for probability of detection as a function of sample size and power ratio are established for constant false alarm rate and an application of this algorithm on the experimental radar data is presented. The presented work contains similar detection results for the SPRT.

1.2 Scope of Thesis

Radar systems are to detect moving targets or the changes in the received signal as quickly as possible with a reasonable algorithm complexity. Proposed method brings a convenient approach for detection in reduced decision time in comparison to fixed sample size systems. This method enables reduced power consumption and better scheduling of the radar coverage area through reduced detection time.

The proposed algorithm is a discrimination algorithm based on the sequential statistical hypothesis testing using AR models. The main idea is to bring the results in literature on the Wald's sequential probability ratio test and AR modeling together and to propose a solution to the target discrimination problem.

The performance of the proposed method is investigated by means of Monte Carlo simulations. The examined performance criterion are ASN (the dependence of ASN on SNR) the probability of type-one (False alarm) and type-two (Miss) classification errors.

1.3 Outline of Thesis

Thesis begins with an introduction of the problem and a summary of the literature research on the related topics. The sequential detection, autoregressive processes and the advantages of sequential tests are qualitatively summarized.

In Chapter 2, the review of the sequential detection procedure is given by considering the two-sided scheme of Wald.

In Chapter 3, the problem definition and an explanation of the proposed algorithm are given

along with some simulation results. The method is illustrated with the numerical examples and the feasibility/limitation of the method is studied with simulations. Brief review of the related methods that are used in the development of the method is also given.

In Chapter 4, the method is applied to the problem of discrimination of the rotary and fixed wing targets. The experimental data used in this chapter has been collected through the joint work ASELSAN and METU-EE members.

The last chapter presents the conclusions on the applicability of the proposed method and outlines some further research directions that can be explored.

CHAPTER 2

BACKGROUND

In order to provide a foundation for the proposed work, we summarize some basic results in detection theory, multivariate distributions, sequential tests and auto-regressive processes in this chapter.

2.1 Detection Theory

2.1.1 Binary Hypothesis Tests

Assume that a set of N observations, $x_1, x_2, \ldots, x_N$; is collected together in the vector $\mathbf{x}$ as shown below:

$$\mathbf{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \tag{2.1}$$

By processing the observation vector, as a classification decision is to be generated. The choices of classification are referred as hypotheses. In general, these hypotheses are $H_0, H_1, \ldots, H_{M-1}$. In the binary hypothesis case, the output is reduced to H_0 and H_1. For binary hypothesis testing, there are four possible assignment cases and their probabilities are denoted as follows:

1) H_0 true and the decision is H_1: Probability of False Alarm (P_{FA})
2) H_0 true and the decision is H_0: $(1 - P_{FA})$

3) H_1 true and the decision is H_1: Probability of Detection (P_D)

4) H_1 true and the decision is H_0: Probability of Miss ($P_M = 1 - P_D$)

2.1.2 Gaussian Probability Density Function

The Gaussian probability density function (pdf) for a scalar random variable **x** having a mean μ_x and a variance σ_x^2 is defined as;

$$p(\mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp[-\frac{1}{2\sigma_x^2}(x-\mu_x)^2] \tag{2.2}$$

It can be denoted by $N(\mu_x, \sigma_x^2)$ and said that $x \sim N(\mu_x, \sigma_x^2)$

If a $nx1$ vector **x** is composed of the scalar gaussian random variables, x, the multivariate Gaussian pdf of **x** becomes

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} \det(\mathbf{C}_x)^{\frac{1}{2}}} \exp[-\frac{1}{2}(\mathbf{x}-\mu_x)^T \mathbf{C}_x^{-1}(\mathbf{x}-\mu_x)] \tag{2.3}$$

and is denoted by $\mathbf{x} \sim N(\mu_x, \mathbf{C}_x)$.

If **x** is a complex Gaussian random vector, then the pdf of **x** is,

$$p(\mathbf{x}) = \frac{1}{\pi^N \det(\mathbf{C}_x)} \exp[-(\mathbf{x}-\mu_x)^T \mathbf{C}_x^{-1}(\mathbf{x}-\mu_x)] \tag{2.4}$$

Likelihood ratio is denoted by $\Lambda(\mathbf{x})$ and defined as

$$\Lambda(\mathbf{x}) \triangleq \frac{p_x(\mathbf{x}|H_1)}{p_x(\mathbf{x}|H_0)} \tag{2.5}$$

2.2 The Sequential Detection

Let $p(x, \Theta)$ denote probability density function (pdf) of the random variable x. Let H_0 be the hypothesis that $\Theta = \Theta_0$ and H_1 the hypothesis that $\Theta = \Theta_1$, Thus , the pdf of x is given by $p(x, \Theta_1)$ when H_1 is true , and by $p(x, \Theta_0)$ when H_0 is true.

For any positive integer m, the pdf that a sample $x_1 \ldots x_k$ is obtained is given by

$$p_{0k} = \prod_{n=1}^{k} p_0(x_n) \tag{2.6}$$

when H_0 is true , and by

$$p_{1k} = \prod_{n=1}^{k} p_1(x_n) \tag{2.7}$$

when H_1 is true.

2.2.1 The Derivation of Upper and Lower Limits for Sequential Probability Ratio Test

The relationship among probability of false alarm (α), probability of detect (β), upper limit (A) and lower limit (B) of the test can be defined as follows: For any given sample $x_1 \ldots x_k$ of type 1, the probability of obtaining such a sample is at least A times as large under hypothesis H_1 as under hypothesis H_0. Thus, the probability measure of the total of all samples of type 1 is also at least A times as large under H_1 as under H_0, [3]. The probability measure of the totality of all samples of type 1 is the same as the probability that the sequential process will terminate with the acceptance of H_1 (rejection of H_0). The latter probability is equal to α when H_0 is true and to $1 - \beta$ when H_1 is true. Thus, we obtain the inequality

$$1 - \beta \leq A\alpha \tag{2.8}$$

This inequality can be written as

$$A \leq \frac{1-\beta}{\alpha} \tag{2.9}$$

Thus, $\frac{1-\beta}{\alpha}$ is an upper limit for A, [3].

A lower limit for B can be derived in a similar way. In fact, for any given sample $x_1 \ldots x_k$ of type 0 the probability of obtaining such a sample under H_1 is at most B times as large as the probability of obtaining such a sample when H_0 is true. Thus, also the probability of accepting H_0 is at most B times as large when H_1 is true as when H_0 is true. Since the probability of accepting H_0 is $1-\alpha$ when H_0 is true and β when H_1 is true, we obtain the inequality

$$\beta \leq (1-\alpha)B \tag{2.10}$$

This inequality can be written as

$$B \geq \frac{\beta}{1-\alpha} \tag{2.11}$$

Thus, $\frac{\beta}{1-\alpha}$ is a lower limit for B.

These relations are called as Wald bounds and they are fundamental for setting the detection thresholds of SPRT, [3]. A practical usage of the Wald sequential probability ratio test is to use the upper and lower bounds discussed above and by taking logarithm they are used as shown below:

$$\ln A = \ln\left(\frac{1-\beta}{\alpha}\right) \tag{2.12}$$

$$\ln B = \ln\left(\frac{\beta}{1-\alpha}\right) \tag{2.13}$$

2.2.2 The Test Procedure of Sequential Probability Ratio Test

The sequential probability ratio test for testing H_0 against H_1 is defined as follows: Two positive constants A and B ($B < A$) are chosen. At each stage of the experiment (at the kth trial), the probability ratio $\frac{p_{1k}}{p_{0k}}$ is computed. If

$$B < \frac{p_{1k}}{p_{0k}} < A \tag{2.14}$$

the experiment is continued by taking an additional observation. If

$$\frac{p_{1k}}{p_{0k}} \leq B \tag{2.15}$$

the process is terminated with the rejection of H_0 (acceptance of H_1)

$$\frac{p_{1k}}{p_{0k}} \geq A \tag{2.16}$$

the process is terminated with the acceptance of H_0.

Take logarithm of the both sides and say $\lambda_k = \ln \frac{p_{1k}}{p_{0k}}$ then Equation 2.17 summarizes the SPRT algorithm.

$$\lambda_k = \begin{cases} \geq \ln A & \rightarrow \quad \text{stop and decide } H_1 \\ \leq \ln B & \rightarrow \quad \text{stop and decide } H_0 \\ otherwise & \rightarrow \quad \text{continue} \end{cases} \tag{2.17}$$

As an example to illustrate the operation of sequential classification, a DC level classification algorithm is formed. The hypotheses for the DC level classification algorithm are:

$$H_0 : x[n] = A_0 + w_0[n] \qquad n = 0, 1, \ldots, N-1 \tag{2.18}$$

$$H_1 : x[n] = A_1 + w_1[n] \qquad n = 0, 1, \ldots, N-1 \tag{2.19}$$

where A_0 and A_1 are mean value of the data of the hypotheses, w_0 and w_1 are WGN with zero mean and variance σ_w^2. Then the LRT for these hypothese becomes:

$$\lambda(x) = \frac{p(x|H_1)}{p(x|H_0)} = \frac{\frac{1}{\sqrt{2\pi\sigma_x^2}} \exp[-\frac{1}{2\sigma_x^2}(x - A_1)^2]}{\frac{1}{\sqrt{2\pi\sigma_x^2}} \exp[-\frac{1}{2\sigma_x^2}(x - A_0)^2]} \tag{2.20}$$

After taking logarithm and making cancelations, we get the loglikelihood ratio:

$$\log \lambda(x) = \frac{1}{2\sigma_x^2}(A_0 - A_1)(A_0 + A_1 - 2x) \tag{2.21}$$

Let k be the number of observations, or length of the sample sequence, then the loglikelihood ratio is equal to

$$\log \Lambda_k = \log \Lambda_{k-1} + \frac{1}{2\sigma_x^2}(A_0 - A_1)(A_0 + A_1 - 2x[k]) \tag{2.22}$$

To illustrate this classification algorithm, the parameters are chosen as $A_0 = 1, A_1 = 2, \sigma_w^2 = 1$ and Wald limits are set such as type 1 and type 2 errors will be under 0.01. The classification procedure is illustrated in Figure 2.1.

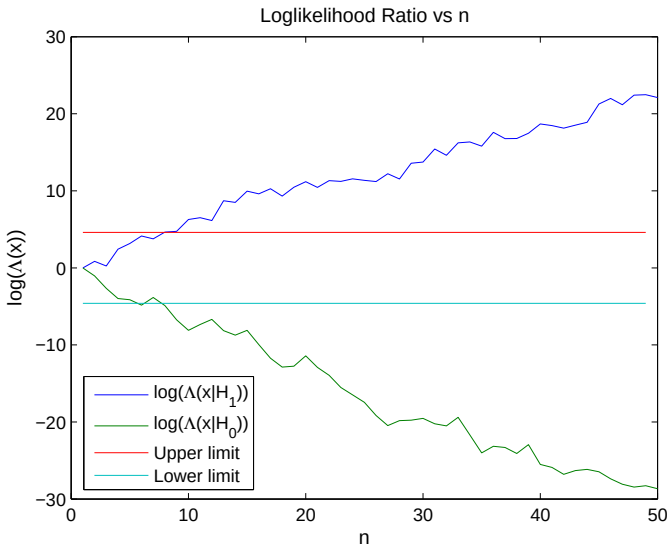


Figure 2.1: Classification Procedure Example of SPRT

2.3 Autoregressive Models

Filtering white noise with a causal linear shift-invariant filter having a rational system function can be used to generate random processes. One of the special types of these random processes, autoregressive processing, is explained in this section. The basic characteristics

and properties of autoregressive processes, their autocorrelation sequences and power spectrum of these processes is given. It is beneficial to start with the generalized form of these type of the random processes, autoregressive moving average processes.

2.3.1 Autoregressive Moving Average Processes

Autoregressive moving average (ARMA) processes are generated by filtering white noise $w(n)$ with a causal linear shift-invariant filter that has a rational system function with p poles and q zeros (Figure 2.23). Therefore, the power spectrum of these processes have, twice of their filter response, $2p$ poles and $2q$ zeros reciprocally.

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q(k) z^{-k}}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \tag{2.23}$$

The power spectrum of white noise is $P_w(z) = \sigma_w^2$ therefore the power spectrum of $x(n)$ is

$$P_x(z) = \sigma_w^2 \frac{B_q(z) B_q^*(1/z^*)}{A_p(z) A_p^*(1/z^*)} \tag{2.24}$$

in terms of the frequency variable w,

$$P_x(e^{jw}) = \sigma_w^2 \frac{|B_q(e^{jw})|^2}{|A_p(e^{jw})|^2} \tag{2.25}$$

Assuming that the filter is stable, the output $x(n)$ is also wide-sense stationary and related with $w(n)$ by

$$x(n) + \sum_{l=1}^{p} a_p(l) x(n-l) = \sum_{l=0}^{q} b_q(l) w(n-l) \tag{2.26}$$

Multiplying both sides by $x^*(n-k)$ and taking expected value the equation becomes

$$E\{x(n)x^*(n-k)\} + \sum_{l=1}^{p} a_p(l) E\{x(n-l)x^*(n-k)\} = \sum_{l=0}^{q} b_q(l) E\{w(n-l)x^*(n-k)\} \tag{2.27}$$

$w(n)$ is WSS and then $x(n)$ and $w(n)$ are jointly WSS therefore,

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \sum_{l=0}^{q} b_q(l) r_{wx}(k-l) \tag{2.28}$$

In order to write $r_{wx}(k)$ in terms of σ_w^2 and $h(k)$, first find $x(n)$ as

$$x(n) = h(n) * w(n) = \sum_{m=-\infty}^{\infty} w(m) h(n-m) \tag{2.29}$$

then the cross-correlation $r_{wx}(k)$ can be written as

$$\begin{aligned} E\{w(n-l)x^*(n-k)\} &= E\{\textstyle\sum_{m=-\infty}^{\infty} w(n-l)w^*(m)h^*(n-k-m)\} \\ &= \textstyle\sum_{m=-\infty}^{\infty} E\{w(n-l)w^*(m)\}h^*(n-k-m) \\ &= \textstyle\sum_{m=-\infty}^{\infty} \sigma_w^2 \delta(n-l-m)h^*(n-k-m) \\ &= \sigma_w^2 h^*(l-k) \end{aligned} \tag{2.30}$$

Substitute Eq.(2.30) into Eq.(2.28) then

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \sigma_w^2 \sum_{l=0}^{q} b_q(l) h^*(l-k) \tag{2.31}$$

Denoting the right side of the equation by $c_q(k)$ and assuming that $h(n)$ is causal, $c_q(k)$ can be written as

$$c_q(k) = \sum_{l=k}^{q} b_q(l) h^*(l-k) = \sum_{l=0}^{q-k} b_q(l+k) h^*(l) \tag{2.32}$$

The Yule-Walker equations becomes

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \begin{cases} \sigma_w^2 c_q(k) & ; \quad 0 \le k \le q \\ 0 & ; \quad k > q \end{cases} \tag{2.33}$$

which, in matrix form become

$$\begin{bmatrix} r_x(0) & r_x(-1) & \dots & r_x(-p) \\ r_x(1) & r_x(0) & \dots & r_x(-p+1) \\ \vdots & \vdots & & \vdots \\ r_x(q) & r_x(q-1) & \dots & r_x(q-p) \\ r_x(q+1) & r_x(q) & \dots & r_x(q-p+1) \\ \vdots & \vdots & & \vdots \\ r_x(q+p) & r_x(q+p-1) & \dots & r_x(q) \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = \sigma_w^2 \begin{bmatrix} c_q(0) \\ c_q(1) \\ \vdots \\ c_q(p) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.34}$$

2.3.2 Autoregressive Processes

Autoregressive process is a type of ARMA(p, q) processes with $q = 0$. This type requires a filter having a rational system function with p poles and no zeros. AR process $x(n)$ is generated at the output by using white noise at the input of this all-pole filter.(Figure 2.2).

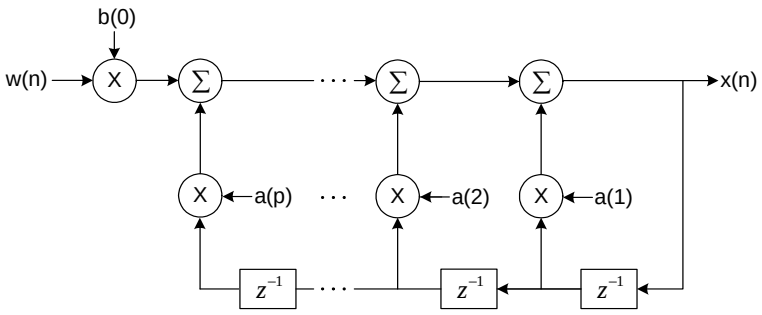


Figure 2.2: The Autoregressive filtering

This filtering process has a frequency response (Eqn. 2.35)

$$H(z) = \frac{b(0)}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \tag{2.35}$$

and the output signal has the power spectrum of

$$P_x(z) = \sigma_w^2 \frac{|b(0)|^2}{A_p(z) A_p^*(1/z^*)} \tag{2.36}$$

in terms of w,

$$P_x(e^{jw}) = \sigma_w^2 \frac{|b(0)|^2}{|A_p(e^{jw})|^2} \tag{2.37}$$

The Yule-Walker equations for an autoregressive process of order of p, AR(p) process, can be found by using Eqn. 2.33 with $q = 0$, $c_0(0) = b(0)h^*(0) = |b(0)|^2$.

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \begin{cases} \sigma_w^2 |b(0)|^2 & ; \quad k = 0 \\ 0 & ; \quad k > 0 \end{cases} \tag{2.38}$$

$$\begin{bmatrix} r_x(0) & r_x(-1) & \dots & r_x(-p) \\ r_x(1) & r_x(0) & \dots & r_x(-p+1) \\ \vdots & \vdots & & \vdots \\ r_x(p) & r_x(p-1) & \dots & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ \vdots \\ a_p(p) \end{bmatrix} = \sigma_w^2 |b(0)|^2 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.39}$$

2.3.3 Moving Average Processes

This type is the other type of the ARMA(p, q) process with $p = 0$. An MA(q) process can be generated by filtering white noise ($w(n)$) having unit variance with an FIR filter

$$H(z) = \sum_{k=0}^{q} b_q(k) z^{-k} \tag{2.40}$$

Therefore, this type has a power spectrum

$$P_x(z) = \sigma_w^2 B_q(z) B_q^*(1/z^*) \tag{2.41}$$

where σ_w^2 represents the power spectrum of the input noise. If it is written in terms of w the equation becomes

$$P_x(e^{jw}) = \sigma_w^2 |A_p(e^{jw})|^2 \tag{2.42}$$

In order to find the Yule-Walker equations for an MA(q) process, Eqn. 2.33 can be used with $a_p(k) = 0$ and $h(n) = b(n)$ and calculated as

$$r_x(k) = \sigma_w^2 \sum_{l=1}^{q-|k|} b_q(l + |k|) b_q^*(l) = \tag{2.43}$$

2.4 Levinson-Durbin Recursion

Levinson-Durbin recursion is used for solving a specially structed matrix equations (The Yule-Walker equations). It is capable of order recursively updating the solution. The Yule-Walker equations for an AR(p) process require to solve a set of linear equations of the form $\mathbf{R}_x \mathbf{a}_p = \mathbf{b}$. If standard method is used it would requires $O(p^3)$ operations. However, if $\mathbf{R}_x$ is a Hermitian Toeplitz matrix which is the condition for the autocorrelation matrix, then by using Levinson-Durbin recursion algorithm these equations can be solved in $O(p^2)$ operations. [32]

To develop recursion, at first, error (ϵ) is needed to be modeled. For the pth recursion, the modeling error is ϵ_p and equivalent to

$$\epsilon_p = r_x(0) + \sum_{l=1}^{p} a_p(l) r_x(l) \tag{2.44}$$

The normal equations for the autocorrelation is

$$\begin{bmatrix} r_x(0) & r_x^*(1) & \dots & r_x^*(p) \\ r_x(1) & r_x(0) & \dots & r_x^*(p-1) \\ \vdots & \vdots & & \vdots \\ r_x(p) & r_x(p-1) & \dots & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ \vdots \\ a_p(p) \end{bmatrix} = \epsilon_p \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.45}$$

In vector notation, it can be written as

$$\mathbf{R}_p \mathbf{a}_p = \epsilon_p \mathbf{u}_1 \tag{2.46}$$

This equation leads to p+1 equations, so that the p+1 unknowns which are $a_p(1), a_p(2), \ldots, a_p(p)$ and ϵ_p can be calculated. The Levinson-Durbin algorithm provides a method to calculate the solutions of these equations recursively. This means that, a_{k+1} can be found by the help of the coefficients calculated before. The recursion is started with the solution for the model of order $k = 0$ as

$$a_0(0) = 1 \tag{2.47}$$

$$\epsilon_0 = r_x(0) \tag{2.48}$$

At the kth-order, the equation becomes $\mathbf{R}_k \mathbf{a}_k = \epsilon_k \mathbf{u}_1$. Assuming $\mathbf{a}_k$ is known, in order to calculate $\mathbf{a}_{k+1}$ the equation becomes $\mathbf{R}_{k+1} \mathbf{a}_{k+1} = \epsilon_{k+1} \mathbf{u}_1$.

Supposing a zero is added to the end of the vector a_k, and write the equation as

$$\begin{bmatrix} r_x(0) & r_x^*(1) & \ldots & r_x^*(k) & r_x^*(k+1) \\ r_x(1) & r_x(0) & \ldots & r_x^*(k-1) & r_x^*(k) \\ \vdots & \vdots & & \vdots & \vdots \\ r_x(k) & r_x(k-1) & \ldots & r_x(0) & r_x^*(1) \\ r_x(k+1) & r_x(k) & \ldots & r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_k(1) \\ \vdots \\ a_k(k) \\ 0 \end{bmatrix} = \begin{bmatrix} \epsilon_k \\ 0 \\ \vdots \\ 0 \\ \gamma_k \end{bmatrix} \tag{2.49}$$

The new parameter γ_k is equal to

$$\gamma_k = r_x(k+1) + + \sum_{i=1}^{k} a_k(i) r_x(k+1-i) \tag{2.50}$$

Because of the Hermitian Toeplitz property of R_{k+1} the Eqn. 2.49 can be written as

$$\begin{bmatrix} r_x(0) & r_x(1) & \dots & r_x(k) & r_x(k+1) \\ r_x^*(1) & r_x(0) & \dots & r_x(k-1) & r_x(k) \\ \vdots & \vdots & & \vdots & \vdots \\ r_x^*(k) & r_x^*(k-1) & \dots & r_x(0) & r_x(1) \\ r_x^*(k+1) & r_x^*(k) & \dots & r_x^*(1) & r_x(0) \end{bmatrix} \begin{bmatrix} 0 \\ a_k(k) \\ \vdots \\ a_k(1) \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma_k \\ 0 \\ \vdots \\ 0 \\ \epsilon_k \end{bmatrix} \tag{2.51}$$

If the complex conjugate of Eqn. 2.51 is taken and combined with the Eqn. 2.49, for any constant Γ_{k+1}, the equality of

$$\mathbf{R}_{k+1}\left(\begin{bmatrix} 1 \\ a_k(1) \\ \vdots \\ a_k(k) \\ 0 \end{bmatrix} + \Gamma_{k+1}\begin{bmatrix} 0 \\ a_k^*(k) \\ \vdots \\ a_k^*(1) \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \epsilon_k \\ 0 \\ \vdots \\ 0 \\ \gamma_k \end{bmatrix} + \begin{bmatrix} \gamma_k^* \\ 0 \\ \vdots \\ 0 \\ \epsilon_k^* \end{bmatrix} \tag{2.52}$$

is provided. If Γ_{k+1} is set as

$$\Gamma_{k+1} = -\frac{\gamma_k}{\epsilon_k^*} \tag{2.53}$$

then Eqn. 2.52 becomes

$$\mathbf{R}_{k+1}\mathbf{a}_{k+1} = \epsilon_{k+1}\mathbf{u}_1 \tag{2.54}$$

where

$$\mathbf{a}_{k+1} = \begin{bmatrix} 1 \\ a_k(1) \\ \vdots \\ a_k(k) \\ 0 \end{bmatrix} + \Gamma_{k+1}\begin{bmatrix} 0 \\ a_k^*(k) \\ \vdots \\ a_k^*(1) \\ 1 \end{bmatrix} \tag{2.55}$$

and

$$\epsilon_{k+1} = \epsilon_k + \Gamma_{k+1}\gamma_k^* = \epsilon_k(1 - |\Gamma_{k+1}|^2) \tag{2.56}$$

Therefore, the Eqn. 2.55 can be written as

$$a_{k+1}(i) = a_k(i) + \Gamma_{k+1}a_k^*(j - i + 1) \tag{2.57}$$

and referred as *Levinson order-update equation*.

CHAPTER 3

CLASSIFICATION OF AUTOREGRESSIVE PROCESSES

The main objective of this work is making the quickest detection of one of the two hypotheses modeled with different Autoregressive (AR) processes through the sequential probability ratio test. In this method, AR coefficients of the processes of two hypothesis can be assumed to be either known or unknown non-random variables. Both cases are studied in this chapter. By using SPRT, the aim is to shorten the average detection time with respect to the fixed sample size methods.

The detection problem can be interpreted as deciding on the synthesis filter generating the AR(p) processes that are in the hypothesis test. This can be done by comparing the PSD (Power Spectrum Density) of the input with the PSD of the AR(p_0) process and AR(p_1) process. Under the null hypothesis H_0, it is assumed that the data is an AR process with p_0 coefficients $a_0(1), a_0(2), \ldots, a_0(p_0)$. Under the alternative hypothesis H_1, it is assumed that the data is an AR process having p_1 coefficients $a_1(1), a_1(2), \ldots, a_1(p_1)$ (Figure 3.1).

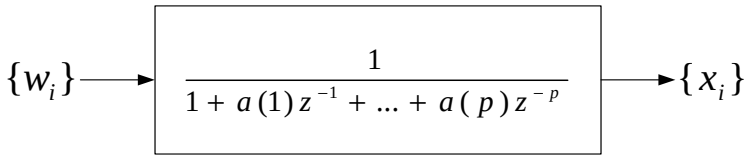


Figure 3.1: The Autoregressive model

If x is a real random variable with $x \sim N(\mu_x, \sigma_x^2)$

For $\mathbf{x} = [x_1 \ldots x_k]$, the multivariate Gaussian pdf is

$$p(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} (\det \mathbf{C}_x)^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\mu_x)^T \mathbf{C}_x^{-1}(\mathbf{x}-\mu_x)} \tag{3.1}$$

Assuming zero mean data the pdf can be written as;

$$p(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} (\det \mathbf{R}_x)^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}^T \mathbf{R}_x^{-1}\mathbf{x})} \tag{3.2}$$

The variance σ_w^2 can be written separately from the autocorrelation function ($\mathbf{R}_x$) as

$$\mathbf{R}_x = \sigma_w^2 \mathbf{R}_a \tag{3.3}$$

where $\mathbf{R}_a$ is the filter autocorrelation function and in the matrix form it is equal to

$$\mathbf{R}_a = \begin{bmatrix} r_a[0] & r_a[-1] & r_a[-2] & \dots & r_a[-(m-1)] \\ r_a[1] & r_a[0] & r_a[-1] & \dots & r_a[-(m-2)] \\ r_a[2] & r_a[1] & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & r_a[0] & r_a[-1] \\ r_a[m-1] & r_a[m-2] & \dots & r_a[1] & r_a[0] \end{bmatrix} \tag{3.4}$$

where each of the matrix element $r_a[k]$ is equal to

$$r_a[k] = \begin{cases} -\sum_{l=1}^{p} a[l] r_a[k-l] & \text{for } k \geq 1 \\ -\sum_{l=1}^{p} a[l] r_a[k-l] + 1 & \text{for } k = 0 \end{cases} \tag{3.5}$$

From the equality it is obvious that $r_a[-k] = r_a^*[k]$ and therefore the filter autocorrelation matrix $\mathbf{R}_a$ is Hermitian and Toeplitz.

In order to calculate autocorrelation matrix $\mathbf{R}_a$, Yule-Walker equations in matrix form can be used as:

$$\begin{bmatrix} r_a[0] & r_a[-1] & r_a[-2] & \dots & r_a[-(p-1)] \\ r_a[1] & r_a[0] & r_a[-1] & \dots & r_a[-(p-2)] \\ r_a[2] & r_a[1] & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & r_a[0] & r_a[-1] \\ r_a[p-1] & r_a[p-2] & \dots & r_a[1] & r_a[0] \end{bmatrix} \begin{bmatrix} a[1] \\ a[2] \\ \vdots \\ \vdots \\ a[p-1] \end{bmatrix} = - \begin{bmatrix} r_a[1] \\ r_a[2] \\ \vdots \\ \vdots \\ r_a[p] \end{bmatrix} \tag{3.6}$$

Substitute (3.3) in (3.2), then

$$p(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} \sigma_w^{-N} (\det \mathbf{R}_a)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_w^2}(\mathbf{x}^T \mathbf{R}_a^{-1} \mathbf{x})} \tag{3.7}$$

If the variance of the time series (σ_x^2) is known, the variance of the white noise (σ_x^2) for each hypothesis can be calculated as

$$\sigma_{w_0}^2 = \frac{\sigma_x^2}{r_{a_{H_0}}(0)} \tag{3.8}$$

$$\sigma_{w_1}^2 = \frac{\sigma_x^2}{r_{a_{H_1}}(0)} \tag{3.9}$$

Therefore the likelihood ratio for known variance case becomes:

$$\Lambda(\mathbf{x}) = \frac{p(\mathbf{x}|H_1, \sigma_{w_1}^2)}{p(\mathbf{x}|H_0, \sigma_{w_0}^2)} = \frac{(2\pi)^{-\frac{N}{2}} \sigma_{w_1}^{-N} (\det \mathbf{R}_{a_1})^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_{w_1}^2}(\mathbf{x}^T \mathbf{R}_{a_1}^{-1} \mathbf{x})}}{(2\pi)^{-\frac{N}{2}} \sigma_{w_0}^{-N} (\det \mathbf{R}_{a_0})^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_{w_0}^2}(\mathbf{x}^T \mathbf{R}_{a_0}^{-1} \mathbf{x})}} \tag{3.10}$$

After making cancelations and taking logarithm of both sides, we get the loglikelihood ratio:

$$\ln \Lambda(\mathbf{x}) = N \ln \frac{\sigma_{w_0}^2}{\sigma_{w_1}^2} + \frac{1}{2} \ln \frac{\det \mathbf{R}_{a_0}}{\det \mathbf{R}_{a_1}} + \frac{1}{2} \mathbf{x}^T \left(\frac{1}{\sigma_{w_0}^2} \mathbf{R}_{a_0}^{-1} - \frac{1}{\sigma_{w_1}^2} \mathbf{R}_{a_1}^{-1} \right) \mathbf{x} \tag{3.11}$$

However, if the variance σ_w^2 is unknown, it has to be estimated with the Maximum Likelihood (ML) method. To maximize $\ln p(\mathbf{x})$ with respect to σ_w^2, we first take the logarithm of the pdf

given in equation 3.7, and then take the derivative with respect to σ_w^2 and equate it to zero as (3.12),

$$\frac{\partial \ln p(\mathbf{x})}{\partial \sigma_w^2} = -\frac{N}{2}\frac{1}{\sigma_w^2} + \frac{1}{2\sigma_w^4}\mathbf{x}^T\mathbf{R}_a^{-1}\mathbf{x} = 0 \tag{3.12}$$

$$\Rightarrow \widehat{\sigma}_w^2 = \frac{1}{N}\mathbf{x}^T\mathbf{R}_a^{-1}\mathbf{x} \tag{3.13}$$

Therefore, the variance for the hypothesis H_0 can be estimated as $\widehat{\sigma}_{w_0}^2 = \frac{1}{N}\mathbf{x}^T\mathbf{R}_{a_0}^{-1}\mathbf{x}$ while the variance estimate for the hypothesis H_1 is equal to $\widehat{\sigma}_{w_1}^2 = \frac{1}{N}\mathbf{x}^T\mathbf{R}_{a_1}^{-1}\mathbf{x}$.

Using (3.13) in (3.7) the equation becomes

$$p(\mathbf{x}) = (2\pi)^{-\frac{N}{2}}\left(\frac{1}{N}\mathbf{x}^T\mathbf{R}_a^{-1}\mathbf{x}\right)^{-\frac{N}{2}}(\det\mathbf{R}_a)^{-\frac{1}{2}}e^{-\frac{N}{2}} \tag{3.14}$$

Therefore, in the unknown power case the likelihood ratio is equal to

$$\Lambda(\mathbf{x}) = \frac{(2\pi)^{-\frac{N}{2}}(\frac{1}{N}\mathbf{x}^T\mathbf{R}_{a_1}^{-1}\mathbf{x})^{-\frac{N}{2}}(\det\mathbf{R}_{a_1})^{-\frac{1}{2}}e^{-\frac{N}{2}}}{(2\pi)^{-\frac{N}{2}}(\frac{1}{N}\mathbf{x}^T\mathbf{R}_{a_0}^{-1}\mathbf{x})^{-\frac{N}{2}}(\det\mathbf{R}_{a_0})^{-\frac{1}{2}}e^{-\frac{N}{2}}} \tag{3.15}$$

Taking logarithm and making cancelations yields the final loglikelihood ratio:

$$\ln\Lambda(\mathbf{x}) = \frac{1}{2}(\ln\det\mathbf{R}_{a_0} - \ln\det\mathbf{R}_{a_1}) + \frac{N}{2}\left(\ln(\mathbf{x}^T\mathbf{R}_{a_0}^{-1}\mathbf{x}) - \ln(\mathbf{x}^T\mathbf{R}_{a_1}^{-1}\mathbf{x})\right) \tag{3.16}$$

In order to calculate $\mathbf{R}_a^{-1}$ and $\det\mathbf{R}_a$ in the loglikelihood ratio, the proposed method uses inverse Levinson-Durbin Recursion, [33], as shown in Algorithm 1.

Algorithm 1 The Recursion method to update $\mathbf{R}_a^{-1}$ and $\det \mathbf{R}_a$

Initialize recursion

$\rho_0 = 1$

$\epsilon_0 = \mathbf{r}_a(0)$

$\mathbf{R}_a^{-1} = \frac{1}{\mathbf{r}_a(0)}$

$\det \mathbf{R}_{a_0} = \mathbf{r}_a(0)$

for all $n = 0, 1, \ldots, m-1$ **do**

$\Gamma_{n+1} = -\frac{\mathbf{r}_a(0:n)\rho_n^R}{\epsilon_n}$

$\rho_{n+1} = [\rho_n \; 0] + \Gamma_{n+1}[0 \; (\rho_n^R)^*]$

$\epsilon_{n+1} = \epsilon_n[1 - |\Gamma_{n+1}|]^2$

$\det \mathbf{R}_{a_{n+1}} = \epsilon_{n+1} \det \mathbf{R}_{a_n}$

$\mathbf{R}_{a_{n+1}}^{-1} = \begin{bmatrix} 0 \\ \mathbf{R}_{a_n}^{-1} \end{bmatrix} + \frac{1}{\epsilon_{n+1}} \rho_{n+1} \rho_{n+1}^H$

end for

In order to find the inverse autocorrelation matrix of the AR processes having pth order, Eqn. 3.17 also can be used, [34].

$$\mathbf{R}_a^{-1} = \frac{1}{\sigma_w^2}(\mathbf{A}_1\mathbf{A}_1^H - \mathbf{A}_2\mathbf{A}_2^H) \tag{3.17}$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are lower triangular Toeplitz matrices and for $N \geq p$ can be calculated as

$$(\mathbf{A}_1)_{i,j} = \begin{cases} 1, & i = j \\ a_{i-j}, & i > j \\ 0, & i > j \end{cases} \tag{3.18}$$

$$(\mathbf{A}_2)_{i,j} = \begin{cases} a_{N-i+j}^*, & i \geq j \\ 0, & i < j \end{cases} \tag{3.19}$$

and $a_k = 0$ for $k < 0$ and $k > p$.

The outline of the proposed algorithm is given in Algorithm 2

Algorithm 2 Summary of the Proposed Algorithm

Initialize $x[0], \mathbf{a}_{H_0}, \mathbf{a}_{H_1}$

For a desired P_{FA} (α) and P_{MISS} (β) compute test bounds from (2.12) and (2.13)

for all $n > 0$ **do**

Update $\mathbf{R}_0^{-1}, \mathbf{R}_1^{-1}$ by using Algorithm 1

Update $\det \mathbf{R}_0^{-1}, \det \mathbf{R}_1^{-1}$ by using Algorithm 1

Update loglikelihood $\eta = \ln \Lambda(\mathbf{x})$ from (3.16)

if $\eta \geq \ln A$ **then**

Decide H_1, break;

else if $\eta \leq \ln B$ **then**

Decide H_0, break;

else

Continue;

end if

end for

In the next section, the described algorithm is examined using Monte Carlo simulations. At the beginning, the variance of the observation data is assumed to be known, i.e. no need to estimate the variance. In this section, the effects of closeness and the value of the AR coefficients are investigated. Then, the accuracy of Wald limits, the upper and lower thresholds, are examined by changing false alarm and miss rates. In the second part, the algorithm with unknown variance is studied and the performance of this part is compared with the first part (known variance) in order to observe the effect of the ML estimation. Finally, the proposed algorithm is compared with a similar classification algorithm having a fixed decision period. In other words, the results of the proposed algorithm are compared with the results of a similar algorithm based on the fixed observation length.

3.1 Sequential Classification of AR(p) Processes : Known Variance Case

In this section, the variance of the processes is assumed to be known as well as the filter coefficients generating processes. In order to produce the test data, white noise is filtered with causal linear shift-invariant filters in order to generate the AR processes. The hypotheses for the classification of AR(1) processes are equal to

$$H_0 : x[n] = a_{H_0} x_0[n-1] + w_0[n] n = 0, 1, \ldots, N-1 \tag{3.20}$$

$$H_1 : x[n] = a_{H_1} x_1[n-1] + w_1[n] n = 0, 1, \ldots, N-1 \tag{3.21}$$

$$\tag{3.22}$$

where a_{H_0} and a_{H_1} are AR(1) coefficients of the processes, $w_0[n]$ and $w_1[n]$ are WGN samples.

The performance of the algorithm is investigated by means of numerical comparisons through Monte Carlo simulations. The examined metrics are the probability of false alarm, miss and average sampling number (ASN) or average detection time. Simulations are made by setting the upper and lower thresholds to the Wald bounds. The parameters are set such that the false alarm rate $\alpha = 0.01$ and the miss rate $\beta = 0.01$. Also Monte Carlo simulations are made for 2000 trials and each data vector have a length of 500.

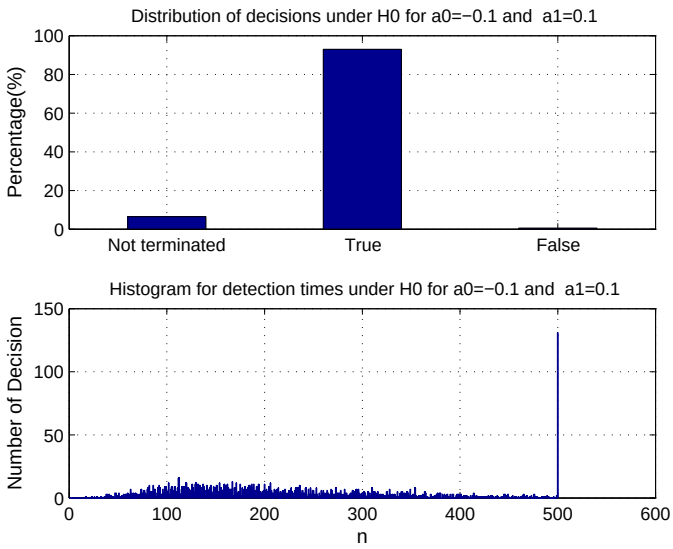
3.1.1 Effect of AR Coefficient Closeness on the Performance of the Algorithm

When the AR coefficients of the models to H_0 and H_1 hypotheses, $\mathbf{a}_{H_0}$ and $\mathbf{a}_{H_1}$ gets closer to each other and the false alarm rate and miss rate are fixed, we may expect the tests to take a larger number of samples until termination.

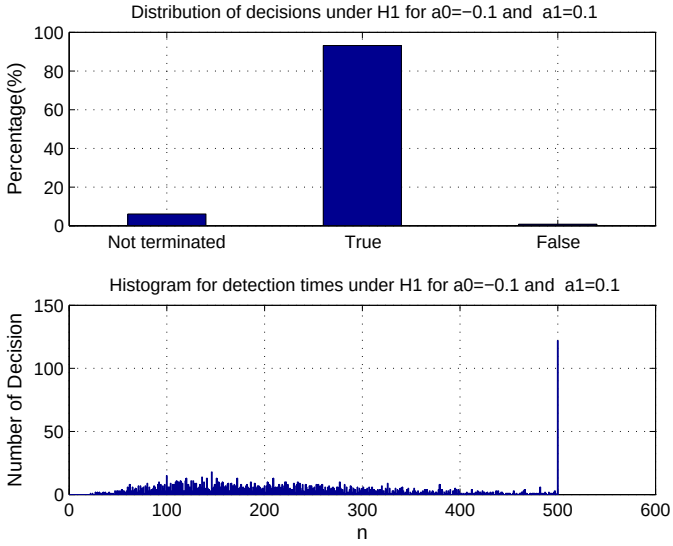
To investigate the effect of the closeness of AR coefficients, we first set AR(1) coefficients to the AR models as $a_{H_0} = -0.1$ and $a_{H_1} = 0.1$ and then increase the difference in between these coefficients and compare the results. The input to the classification algorithm is the $\mathbf{x}_0$ vector generated under hypothesis H_0. When the input is fixed to process related to the H_0 hypothesis, false alarm rate of the system is examined. (Figure 3.2(a)) Then, the input is fixed to H_1 hypothesis and miss rate is studied. (Figure 3.2(b))

Bar graph represents the percentage of the number of decisions after the 2000 trials. The SPRT algorithm makes 3 types of decisions. The test can terminate with a decision H_0 or H_1. If the algorithm can not be able to make any decision after 500 samples, the test is said to be unterminated. The histogram plot represents the distribution of the termination of total 2000 trials. The mean of the distribution is equal to the ASN of the trial.

Secondly, we repeat the same procedure for $a_{H_0} = -0.125$ and $a_{H_1} = 0.125$. (Figure 3.3(a),



(a) $a_{H_0} = -0.1$, $a_{H_1} = 0.1$, FA=0.0059, ASN=227.3160, Unterminated Percent 6.45%

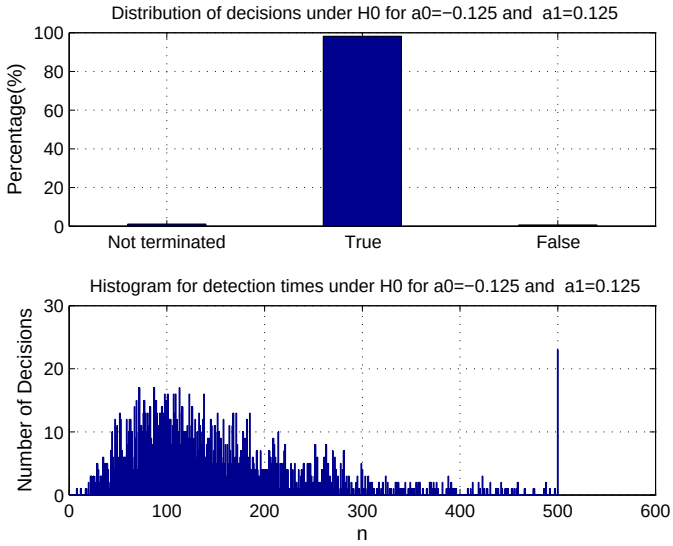


(b) $a_{H_0} = -0.1$, $a_{H_1} = 0.1$, MISS=0.0085, ASN=230.1190, Unterminated Percent 6.05%

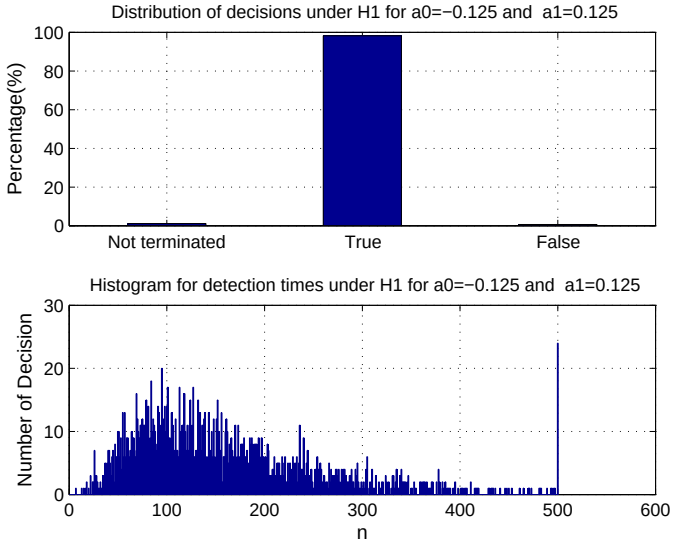
Figure 3.2: $a_{H_0} = -0.1$, $a_{H_1} = 0.1$ False Alarm and Miss statistics for known variance case

Figure 3.3(b)) And finally, the algorithm is tested by setting the AR(1) coefficients as $a_{H_0} = -0.15$ and $a_{H_1} = 0.15$. (Figure 3.4(a), Figure 3.4(b))

As expected, as the distance between the AR coefficients increases, unterminated trials and the average detection number decreases at fixed false alarm and miss rate values. (Figure 3.1) Because the rates are decided by the Wald thresholds and they are adjusted to the same value



(a) $a_{H_0} = -0.125$, $a_{H_1} = 0.125$, FA=0.0066, ASN=153.3720, Unterminated Percent 1.15%



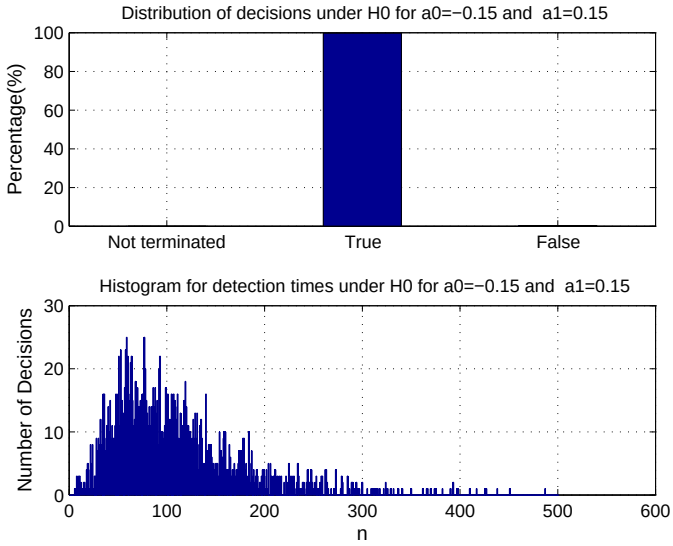
(b) $a_{H_0} = -0.125$, $a_{H_1} = 0.125$, MISS=0.0066, ASN=160.8205, Unterminated Percent 1.15%

Figure 3.3: $a_{H_0} = -0.125$, $a_{H_1} = 0.125$ False Alarm and Miss statistics for known variance case

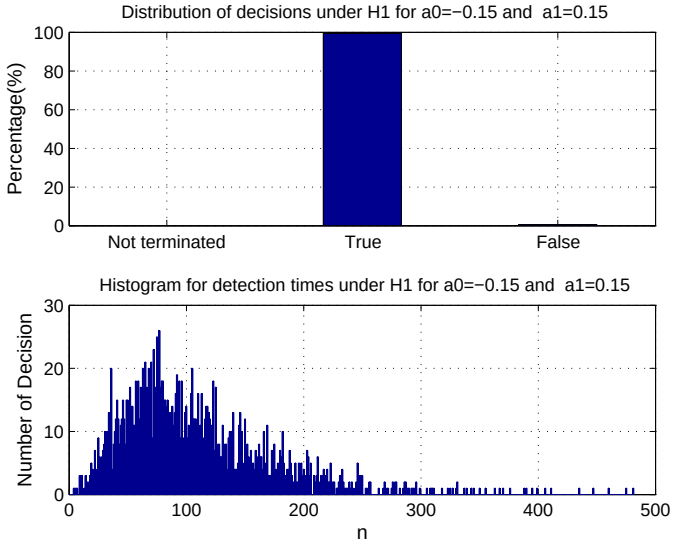
(1%) for all the trials.

3.1.2 Accuracy of Wald Thresholds and Achieved False Alarm and Miss Rates

Finding the correct threshold value is the critical task and for many tests more difficult than constructing the optimal detector. However, the Wald limits bypass this problem effortlessly



(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA=0.0030, ASN=110.3460, Unterminated Percent 0%



(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, MISS=0.0055, ASN=109.2190, Unterminated Percent 0%

Figure 3.4: $a_{H_0} = -0.15$, $a_{H_1} = 0.15$ False Alarm and Miss statistics for known variance case

Table 3.1: ASN with respect to the closeness of AR(1) coefficients for known variance case

a_{H_0}	a_{H_1}	ASN Under H_0	ASN Under H_1
-0.1	0.1	227.3160	230.1190
-0.125	0.125	153.3720	160.8205
-0.15	0.15	110.3460	109.2190

and they are extreme easy to calculate and valid for all likelihood ratio tests, [3]. The sequential probability ratio test with Wald limits says that,

$$\lambda_k = \begin{cases} \geq \ln A & \rightarrow \quad \text{stop and decide } H_1 \\ \leq \ln B & \rightarrow \quad \text{stop and decide } H_0 \\ otherwise & \rightarrow \quad \text{continue} \end{cases} \tag{3.23}$$

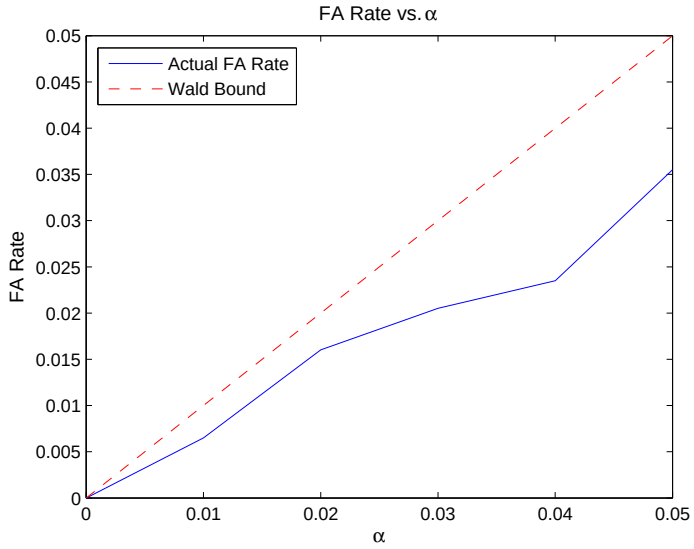
where $\frac{1-\beta}{\alpha}$ is an upper limit for A and $\frac{\beta}{1-\alpha}$ is a lower limit for B. And using this limit values as upper and lower thresholds guarantees to provide desired false alarm and miss rates.

α is the parameter that sets the false alarm rate in Wald sequential probability ratio test. By setting this parameter according to Wald limits, the actual false alarm probability ($p(H_1|H_0)$) is guaranteed to be below the desired rate. Also, for the miss rate the parameter β is used and also setting this parameter guarantees that actual miss rate ($p(H_0|H_1)$ is below the desired value. In this section, the accuracy of these limits is examined by setting the error rates initially to 0.01 and then increasing these parameter to 0.05. To examine the actual false alarm rate and α relation, as an input under H_0 hypothesis is applied to the algorithm. (Figure 3.5(a)) An input under H_1 hypothesis is applied to the algorithm so that the relation of actual miss rate and β is checked. (Figure 3.5(b))

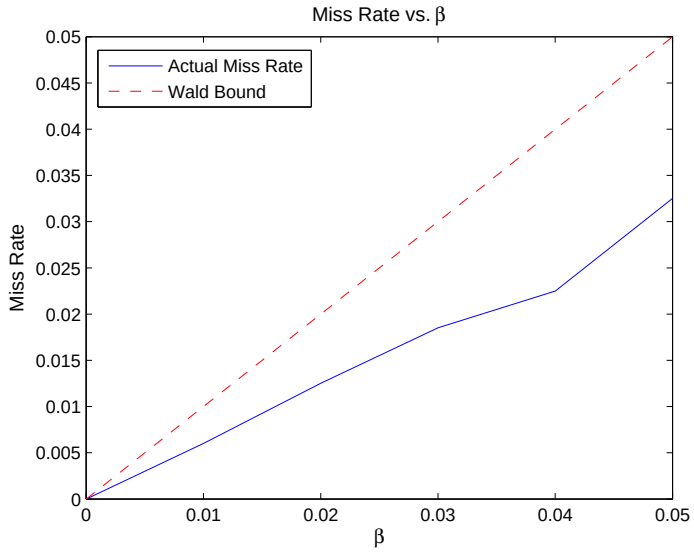
3.1.3 Effects of the Value of the AR Coefficients to the Performance of the Algorithm

In a typical signal modeling problem, the outputs of stable filters are used to model a given random process. With AR modeling, the filter response of the rational system function has poles according to the AR coefficients of the process. When the AR coefficients are chosen, special attention should be paid on the stability of the filter. For an AR(1) process AR coefficients can take the values between 1 and -1 in order to make the filter stable.

The effect of chosen AR coefficients on the false alarm and miss rates, ASN and the number of unterminated trials are observed by running the algorithm for different AR coefficients. Firstly, the effect of the AR coefficients to the performance of the algorithm with known power is examined and the simulation results are in Appendix A.1 and the ASN' s are listed in Table 3.2.



(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA rate vs α



(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, Miss rate vs β

Figure 3.5: $a_{H_0} = -0.15$, $a_{H_1} = 0.15$ FA vs α and MISS vs β for known variance case

Table 3.2: ASN with respect to the value of AR(1) coefficients for known power case

a_{H_0}	a_{H_1}	ASN Under H_0	ASN Under H_1
0	0.3	110.1410	111.3980
0.3	0.6	104.8395	117.5130
0.6	0.9	95.5640	140.9580

Secondly, the effect of the AR coefficients to the performance of the algorithm with unknown power is examined and the simulation results can be reached in Appendix A.2. The resultant ASN' s are listed in Table 3.3.

Table 3.3: ASN with respect to the value of AR(1) coefficients for unknown power case

a_{H_0}	a_{H_1}	ASN Under H_0	ASN Under H_1
0	0.3	112.8660	107.6575
0.3	0.6	97.5770	81.4765
0.6	0.9	66.5160	42.7900

Against the expectations, the ASN values for the unknown variance case are smaller than the ASN values for the known variance case. However, the results of the known power case and the unknown power case have to be evaluated independently because assumptions are different in the derivation of the loglikelihood ratios. For the unknown power case, as the pole of the filter approaches the unit circle in the complex plane, the resultant ASN decreases.

3.2 Sequential Classification of AR(p) Processes : Unknown Variance Case

If the variance of the random processes is not known, the maximum likelihood (ML) estimation of the time series has to be calculated according to Equation 3.24 and inserted in to likelihood ratio.

$$\widehat{\sigma}_w^2 = \frac{1}{N}\mathbf{x}^T\mathbf{R}_a^{-1}\mathbf{x} \tag{3.24}$$

The algorithm estimates the variance of a white Gaussian random noise with zero mean and a power σ_w^2. (Figure 3.6)

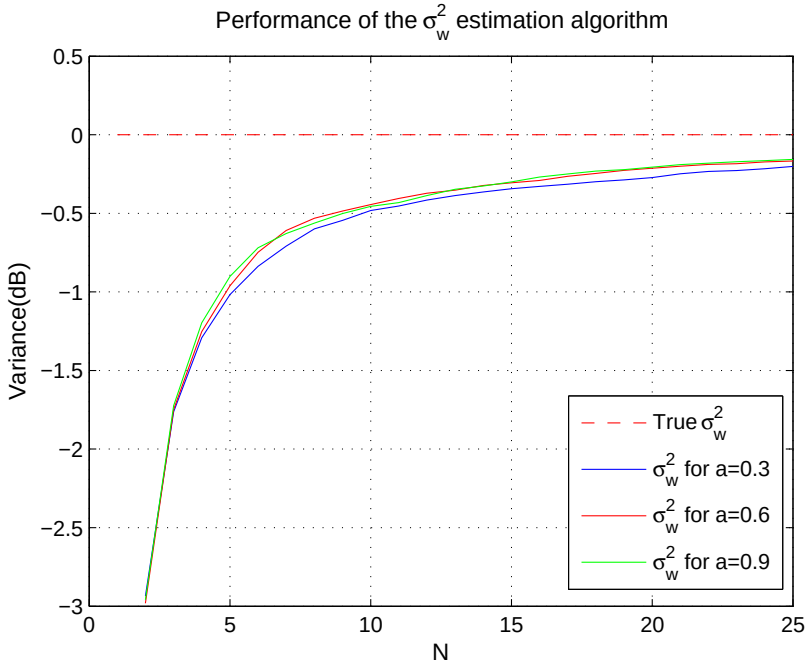
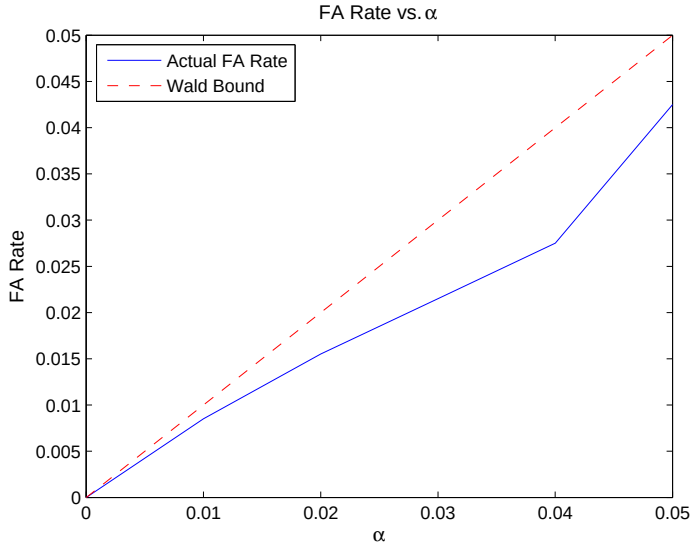


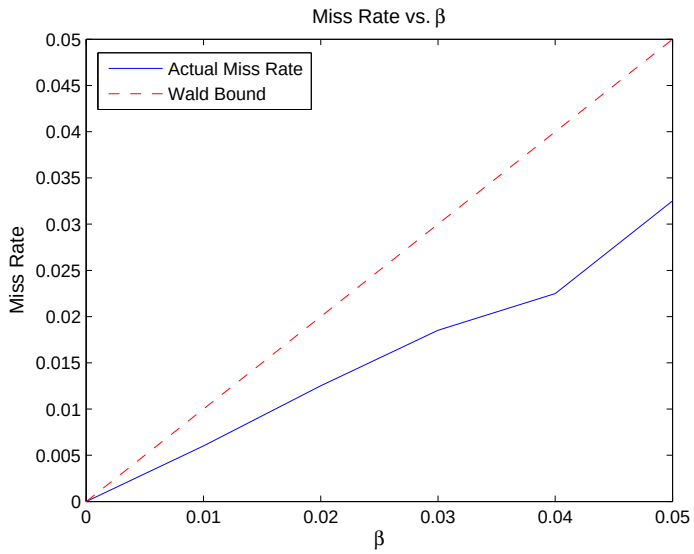
Figure 3.6: Performance of the estimation algorithm, true $\sigma_w^2 = 0\ dB$

Because of using an estimate instead of the true value for the variance, the algorithm may require some additional samples for the initial variance calculation to be reliable. To reduce the negative effects of this delay in decision, this number is chosen by considering the results of the previous simulation. The decision start sample number (the first output of the SPRT test) is chosen as 10. In this way, the likelihood ratio is calculated using first 10 samples and then SPRT test proceeds by updating both the hypotheses scores and the variance estimate.

Due to the fact that the Wald limits are pessimistic bounds for the false alarm and miss rates, the simulation results show that the bounds are still valid for the unknown variance case despite the inaccuracies of the initial estimations. (Figure 3.7)



(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA rate vs α

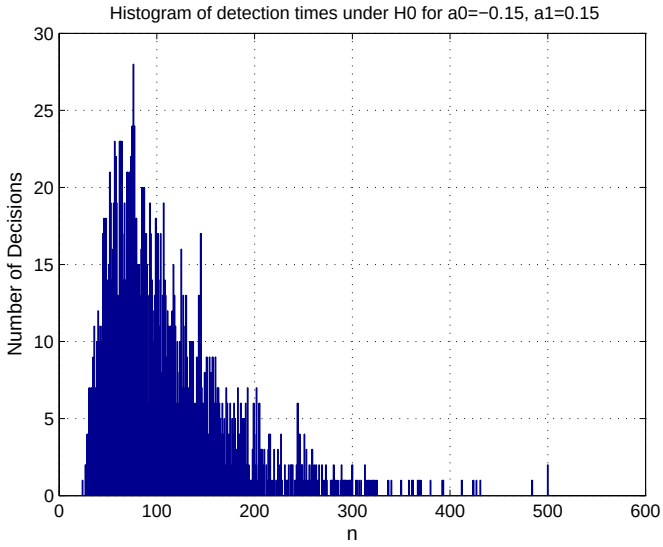


(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, Miss rate vs β

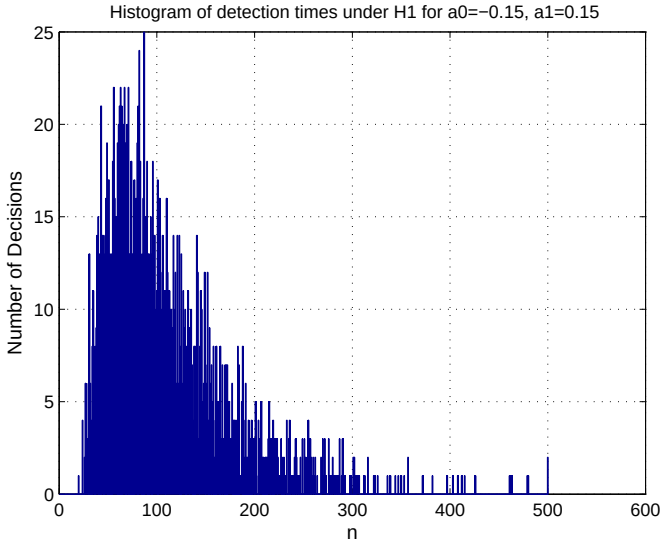
Figure 3.7: $a_{H_0} = -0.15$, $a_{H_1} = 0.15$ FA vs α and MISS vs β Unknown Variance

For the unknown variance case, histogram of the required sample numbers, resultant error rates, ASN and unterminated percentage of the trials are illustrated in Figure 3.8 by choosing AR coefficients of H_0 as $a_{H_0} = -0.15$ and H_1 as $a_{H_1} = 0.15$. Figures are prepared by using simulated data under H_0 (Figure 3.8(a)) and H_1 (Figure 3.8(b)) hypotheses. The Wald limits

are set to $\alpha = \beta = 0.01$, i.e. the expected false alarm and miss rates have to be under 1%. The ASN (Average Sample Number) which is the mean number of sample points demanded to make a decision is 111.9560 under H_0 and 111.2125 under H_1. Also resulted false alarm rate is 0.0075 and miss rate is 0.0085 which are below the Wald limits as expected.



(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA=0.0075, ASN=111.9560, Unterminated Percent 0.1000%



(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, MISS=0.0085, ASN=111.2125, Unterminated Percent 0.1000%

Figure 3.8: The Histograms for the Detection Times with $a_{H_0} = -0.15$, a_{H_1}=0.15

3.2.1 Unknown Variance Sequential Detection of AR(p) Processes for Complex Time Series

If the time series vector is complex Gaussian distributed random process with zero mean, the associated multivariate complex Gaussian pdf is

$$p(\mathbf{x}) = \frac{1}{\pi^N \det(\mathbf{C}_x)} \exp[-(\mathbf{x})^H \mathbf{C}_x^{-1}(\mathbf{x})] \tag{3.25}$$

and is denoted by $\mathbf{x} \sim CN(0, \mathbf{C}_x)$.

Loglikelihood ratio is denoted by $\ln \Lambda(\mathbf{x})$ and it is

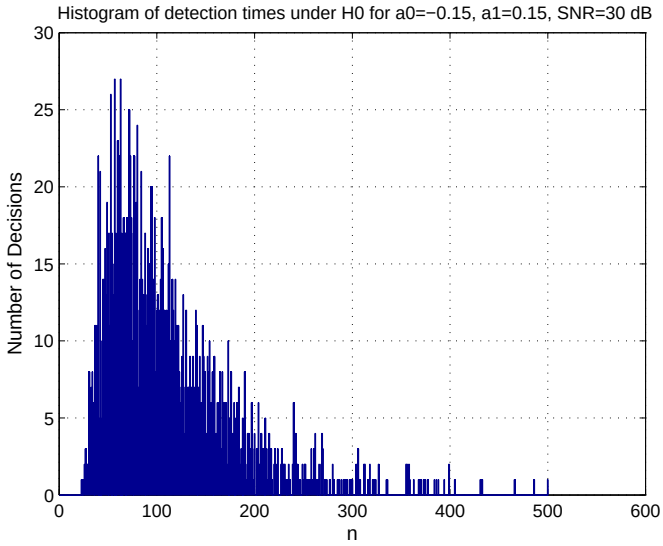
$$\ln \Lambda(\mathbf{x}) = (\ln \det \mathbf{R}_{a_0} - \ln \det \mathbf{R}_{a_1}) + N(\ln(\mathbf{x}^H \mathbf{R}_{a_0}^{-1}\mathbf{x}) - \ln(\mathbf{x}^H \mathbf{R}_{a_1}^{-1}\mathbf{x})) \tag{3.26}$$

3.3 The Effect of SNR on the Detection of AR(p) Processes

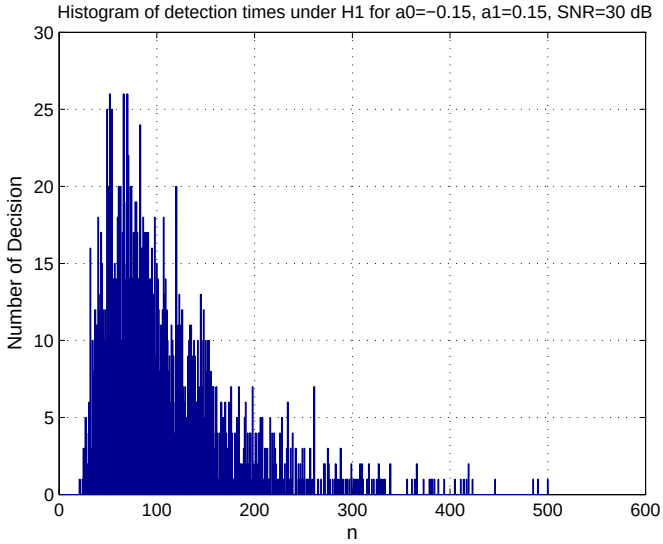
The aim of this section is to examine the effect of the additive white noise on the detection of AR signals . AR signals are produced by filtering the white noise by p AR coefficients for the pth order AR processes. As the white noise is added to these AR processes the structure of these signals are distorted. (This distortion due to the noise is not accounted in the hypothesis definitions in this work.) In this section, the algorithm is tried for different SNR values by fixing other parameters. Firstly, by adding white noise to the AR signal, the SNR of this signal is adjusted as 30 dB. (Figure 3.9)

The simulations are repeated for the SNR values of 20 dB and 10 dB and the results are illustrated in Figure 3.10 and Figure 3.11 respectively. The ASN under H_0 is equal to 112.5365 at 30 dB SNR but it increases to 113.3295 and 120.5155 as the value of the SNR decreases to 20 dB and 10 dB respectively. The results are similar for the hypothesis H_1. Finally at 10 dB SNR the error rates exceed the adjusted limits ($\alpha = \beta = 0.01$) by the Wald thresholds.

It is observed that as the SNR decreases the ASN is increased and also the probability of errors are increased (Figure 3.9, Figure 3.10, Figure 3.11) Therefore, the probability of errors vs. SNR is examined in the next set of simulations.



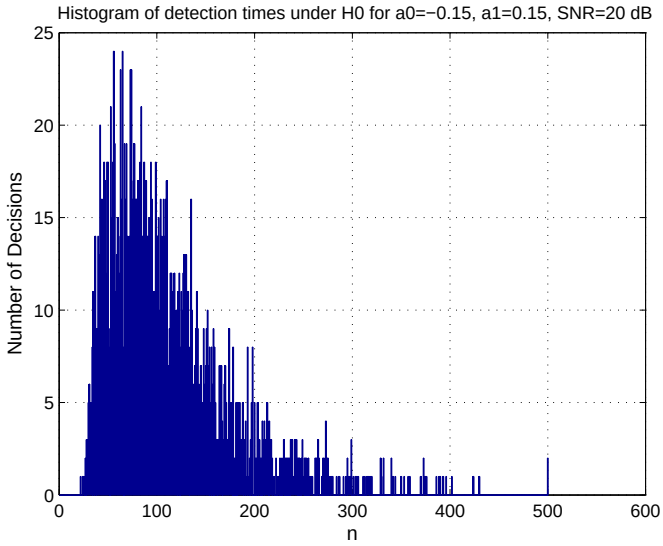
(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA=0.0070, ASN=112.5365, Unterminated Percent 0.0500%



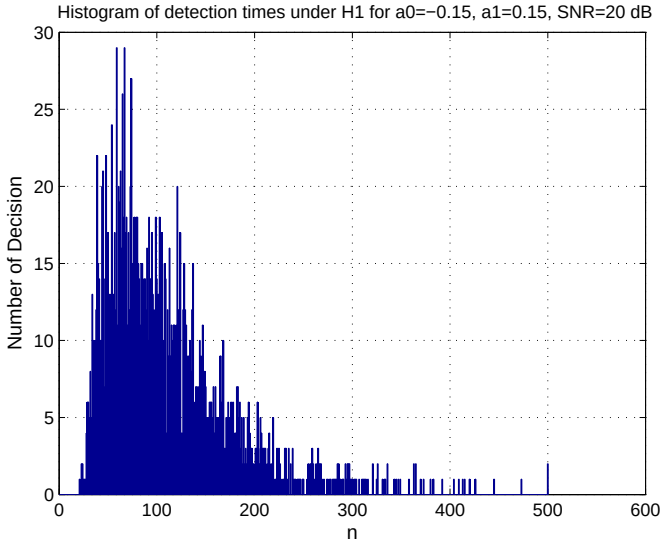
(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, MISS=0.0055, ASN=111.7505, Unterminated Percent 0.0500%

Figure 3.9: The Histograms for the Detection Times with SNR=30 dB, $a_{H_0} = -0.15$, a_{H_1}=0.15

In order to study the effect of additive CWGN, after the production of AR process by filtering the white noise with complex AR coefficients, extra complex Gaussian white noise is added to the output process. In these simulations, the value used for the AR coefficients of the hypotheses H_0 and H_1 are listed in the Table 3.4 and the hypotheses represented the helicopter echo and CWGN respectively. The value of the AR(5) coefficients for the helicopter hypothesis (H_0) is derived at the Section 4.1.2 by fitting the AR model to the experimental data. The



(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA=0.0075, ASN=113.3295, Unterminated Percent 0.1000%

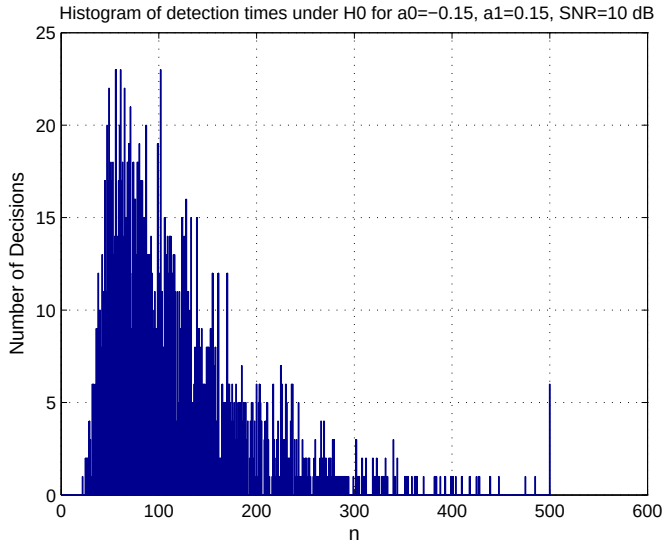


(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, MISS=0.0090, ASN=112.5795, Unterminated Percent 0.1000%

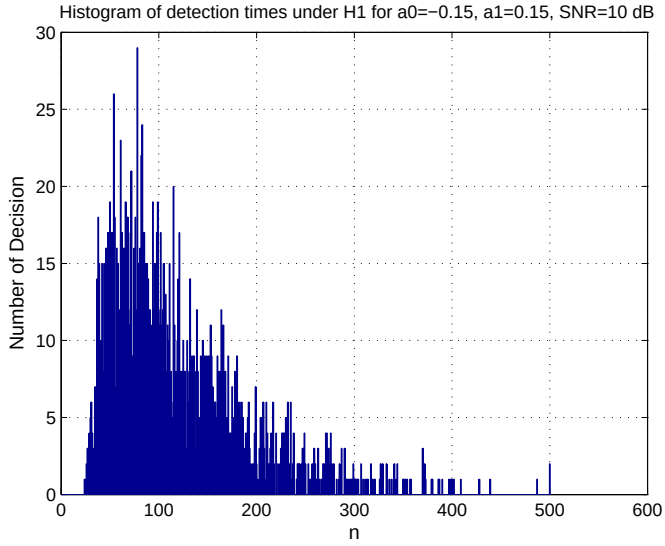
Figure 3.10: The Histograms for the Detection Times with SNR=20dB, $a_{H_0} = -0.15$, a_{H_1}=0.15

alternative hypothesis represents the CWGN. (By this configuration, it is aimed to examine the detection performance of the sequential algorithm when one hypothesis is related to the helicopter echo and the other one is the white noise.)

The Wald limits are also valid for the detection of signals in CWGN. However, as the decision start number (N_{start}) gets larger the algorithm is able to work for the signals having low SNR



(a) $a_{H_0} = -0.15$, a_{H_1}=0.15, FA=0.0115, ASN=120.5155, Unterminated Percent 0.3000%



(b) $a_{H_0} = -0.15$, a_{H_1}=0.15, MISS=0.0105, ASN=120.9145, Unterminated Percent 0.1000%

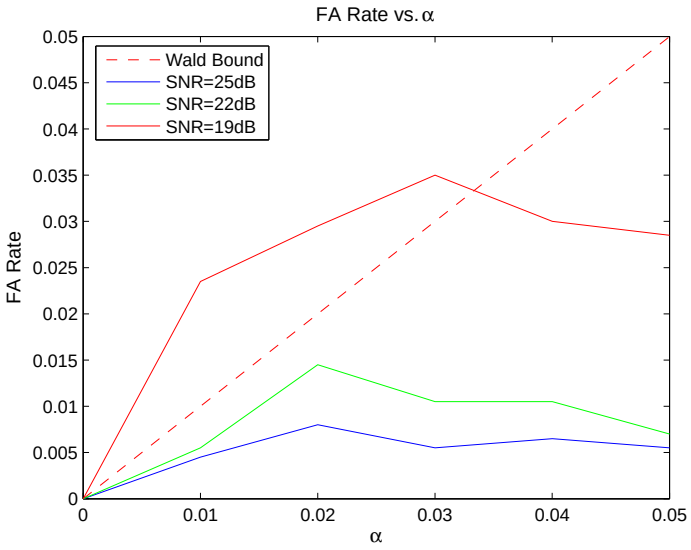
Figure 3.11: The Histograms for the Detection Times with SNR=10dB, $a_{H_0} = -0.15$, a_{H_1}=0.15

values. If the decision start number is chosen as 20, the algorithm starts to take an action after 20th sample. For the first simulation N_{start} is chosen as 20 and it is observed that the algorithm works under the adjusted error rates above 19 dB SNR by increasing the value of the α parameter from 0.01 to 0.05 (Figure 3.12(a)). The power spectrum estimates of the the data vectors under helicopter hypothesis (H_0), under CWGN hypothesis (H_1) and under helicopter hypothesis (H_0) having 19 dB SNR by adding CWGN are illustrated at Figure 3.12(b).

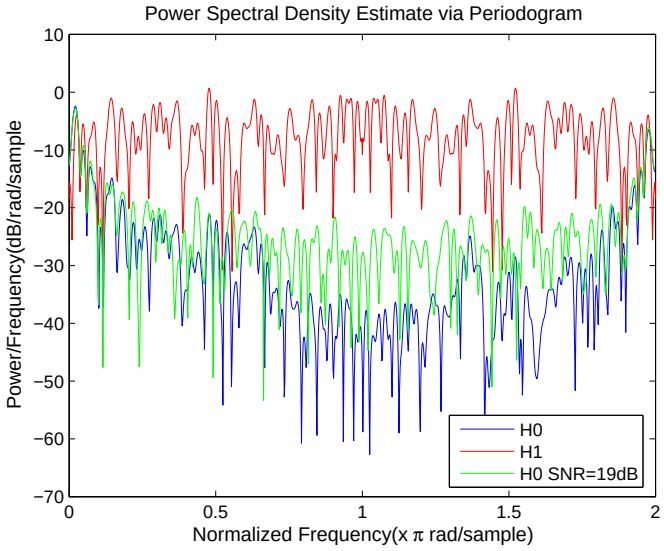
Table 3.4: AR coefficients for the helicopter (H_0) and CWGN (H_1) hypothesis

Hypothesis	H_0	H_1
$a(1)$	0	-1.2542-0.4250i
$a(2)$	0	0.5051+0.3334i
$a(3)$	0	-0.7718+0.0352i
$a(4)$	0	0.6383+0.2075i
$a(5)$	0	-0.1109-0.1649i

The simulation is repeated for the case of $N_{start} = 50$. The method operates adequately at 14 dB SNR (Figure 3.13(a)). And it is observed that by increasing the parameter N_{start}, the algorithm is able to work under lower SNR's.

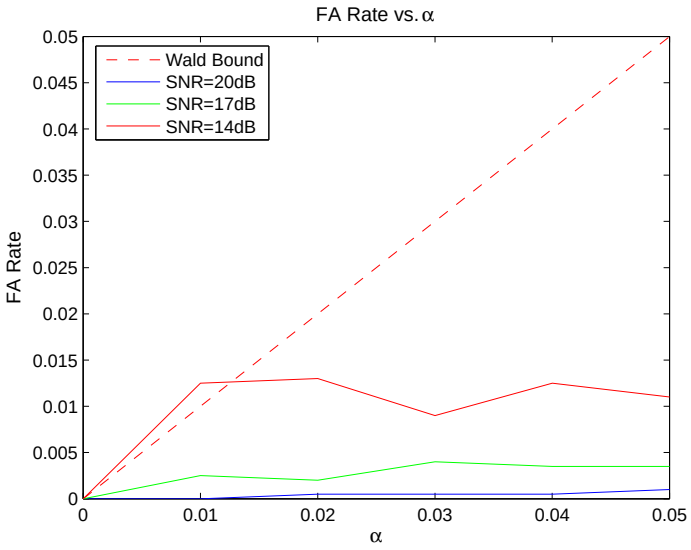


(a) $p(H_1|H_0)$ vs α wrt SNR, $N_{start} = 20$

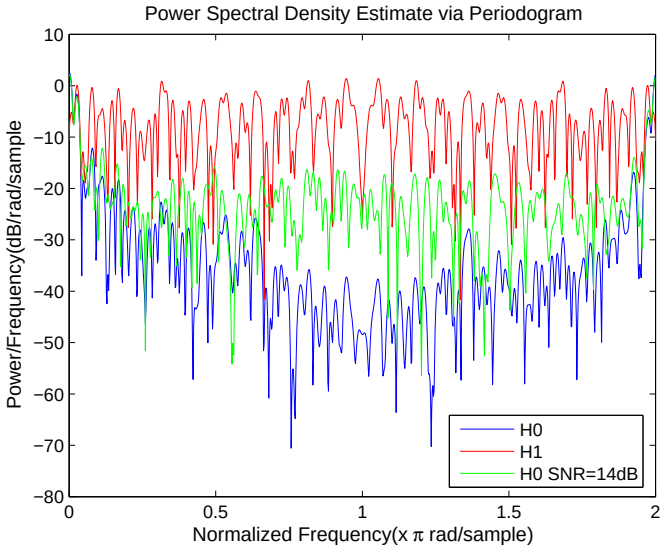


(b) Power Spectrum Estimates of H_0, H_1 and H_0 SNR=19 dB, $N_{start} = 20$

Figure 3.12: Power Spectrum Estimates for $N_{start} = 20$



(a) $p(H_1|H_0)$ vs α wrt SNR, $N_{start} = 50$



(b) Power Spectrum Estimates of H_0, H_1 and H_0 SNR=14 dB, $N_{start} = 50$

Figure 3.13: Power Spectrum Estimates for $N_{start} = 50$

3.4 Comparison of Sequential and Fixed Length Detection of AR(p) Processes

In this section, we compare the proposed method with a fixed sample size classifier. We first outline an algorithm that makes detection with a fixed sample size. The performance of this method is compared with the performance of SPRT test given in the earlier sections.

In the figure below, N (Fixed sample size) is increased to 250 and $P_{FA} = P_{MISS}$ is chosen so that P_D vs N is analyzed.

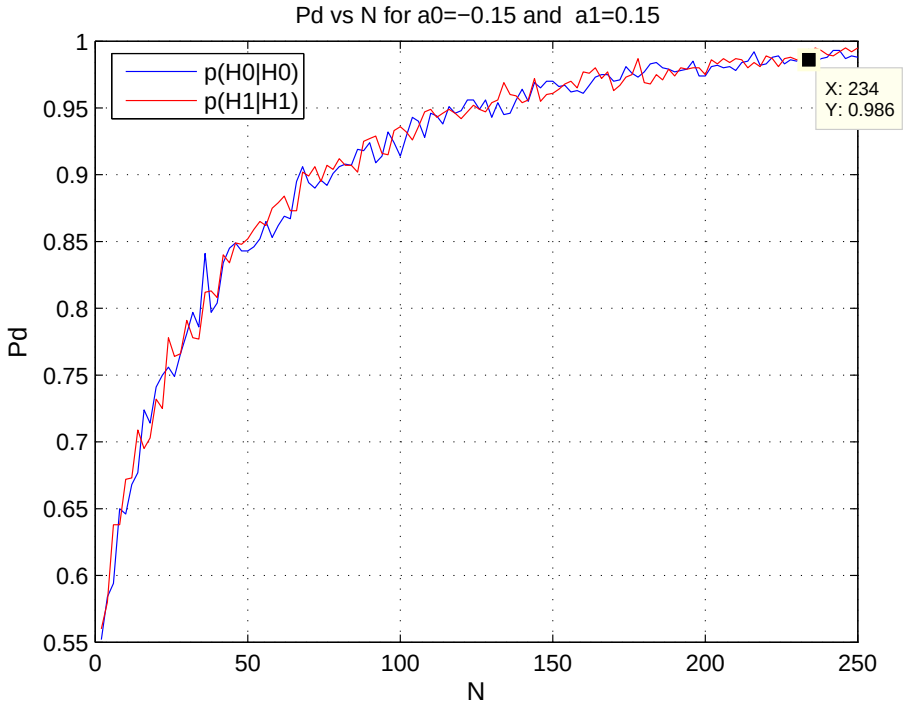
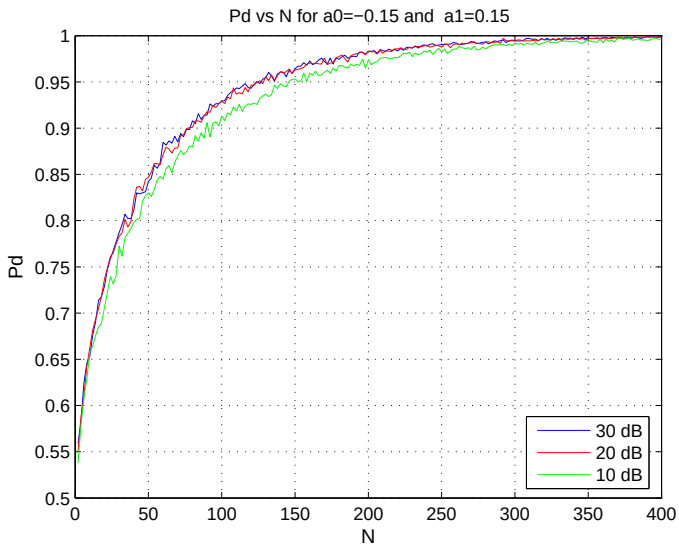


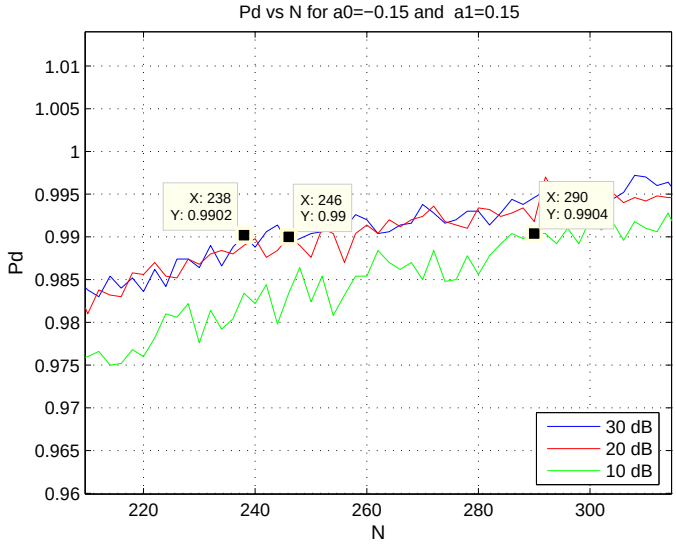
Figure 3.14: P_D vs Sample Number for $a_{H_0} = -0.15$, $a_{H_1} = 0.15$

For different SNR values the previous experiment is repeated in the next figure. In order to achieve same error rates, the requirement for increasing FSS can be observed.

To achieve the same error rates, the sample size for an FSS algorithm has to increase as the SNR of the time series decreases. This simulations (Figure 3.15(b)) can be compared with the results of the sequential algorithm (Figure 3.9, Figure 3.10, Figure 3.11). In order to achieve same error rates ($P_{FA} = P_{MISS} = 0.01$) in equal detection conditions ($a_{H_0} = -0.15$, $a_{H_1} = 0.15$) the requirement of average sampling numbers (ASN) are shown in Table 3.5:



(a) P_D vs Sample Number for $P_{FA} = P_{MISS}$, $a_{H_0} = -0.15$, $a_{H_1} = 0.15$



(b) The Sample Numbers for different SNR's to achieve $P_{FA} = P_{MISS} = 0.01$, $a_{H_0} = -0.15$, $a_{H_1} = 0.15$

Figure 3.15: P_D vs Fixed Sample Size(N) for different SNR's, $a_{H_0} = -0.15$, $a_{H_1} = 0.15$

Table 3.5: The ASN for Sequential and Fixed Sample Size algorithms

SNR	$30dB$	$20dB$	$10dB$
Wald' s SPRT	112	113	121
Fixed Sample Size	238	246	290

In order to examine the proposed method at lower than 10 dB SNR, we increase the N_{start} number and make some simulations. The results of these simulations are listed in Table 3.6:

Table 3.6: The performance results of proposed method at low SNR's

$Nstart$	SNR	Error Rate	ASN	Unterminated
20	8 dB	0.0175	122.0	% 0.20
25	8 dB	0.0130	122.3	% 0.15
30	8 dB	0.0125	124.9	% 0.15
30	6 dB	0.0221	134.8	% 0.55
30	4 dB	0.0273	145.7	% 1.05

Below 10 dB SNR, an increase in the $Nstart$ number is not enough to provide the preset error probability by Wald limits. This problem can be arouse from the addition of WGN disrupt the AR structure of the hypotheses. As a future work, how to reflect the effect of the noise on the hypotheses will be analyzed.

CHAPTER 4

AN APPLICATION EXAMPLE : ROTARY - FIXED WING CLASSIFICATION

In the previous chapter, a sequential algorithm to make a fast decision with an acceptable false alarm and miss rates, is given. This method uses the sequential probability ratio test with Wald's thresholds. Also in order to decrease the complexity of the algorithm, the Levinson-Durbin recursions is implemented to calculate required matrix inverses. The algorithm can also be used if the process has unknown variance with some initial delay to reliable estimate the unknown variance. The other parameters of the processes (the order of the process and the pole locations) are assumed to be known.

The advantages of variable sample size detection can be utilized in rapid detection of targets having different spectral characteristics such as rotary and fixed wing targets. The only pre-processing step requirement to use the suggested algorithm is to model the hypotheses with autoregressive models.

4.1 Target Models

In the proposed algorithm requires AR models for each hypothesis. To calculate the model coefficients, the algorithm need to have some knowledge of PSD of the each process.

4.1.1 Modeling of Backscattering from Hovering Helicopter

When the radar pulse impinges on the hovering helicopter, the backscattered signal is affected by several components of the helicopter. Basically, the components of the helicopter that

affects the echo signal can be divided into 4 parts. These parts are the body of the helicopter, the hub, the main rotor blade and the tail rotor blade. (Figure 4.1)

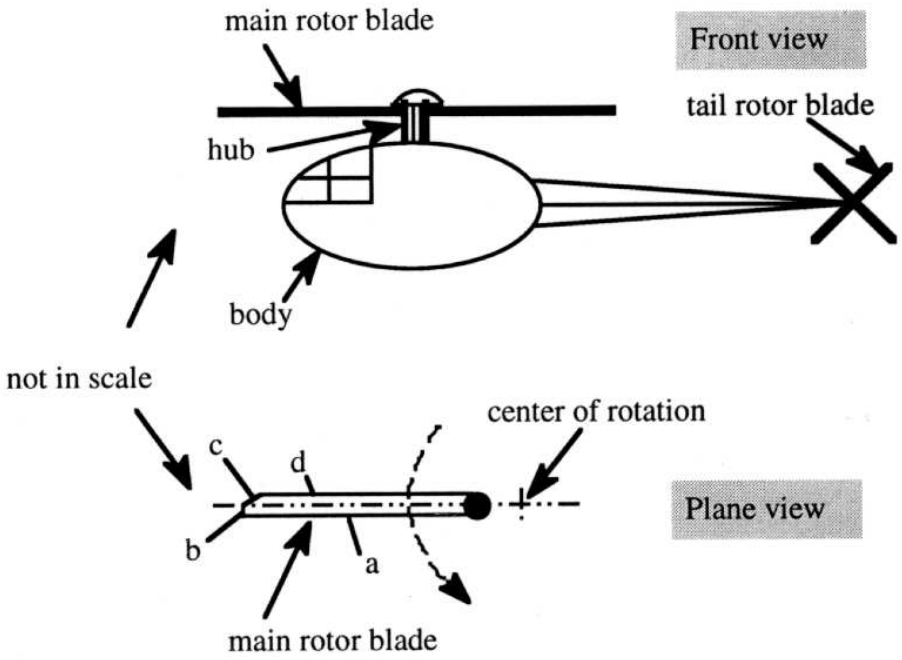


Figure 4.1: Contributions to EM backscattering from the parts of the Hovering Helicopter, [1]

The contribution of the body to the PSD is at only zero frequency. The hub has a triangle shaped PSD centered at $f = 0\ Hz$. The value of f_{hub} is dependent to the radar cross section (RCS) and the radial velocity of the hub. The approaching and receding parts of the main rotor blade impacts the PSD differently because of the shape of the blade. The part that approaches have a greater RCS with respect to the receding part. Therefore, the approaching blade has a higher level of the PSD than receding blade. f_{max} has a relationship with the tip speed. The final part is the tail rotor blades. The rotation plane of the tail rotor blades is perpendicular to the radar generally. Therefore, the contribution of the scattering from the tail rotor blades is little with respect to the other parts, [1]. Figure 4.2.

Following simulations and comparisons are prepared with the experimental data which is gathered from an experiment for the analysis of the doppler and flash effects of a helicopter. If the experimental data from the helicopter returns is analyzed, the constructed PSD estimate (Periodogram) given in Figure 4.3 is closely related to the one given in [1].

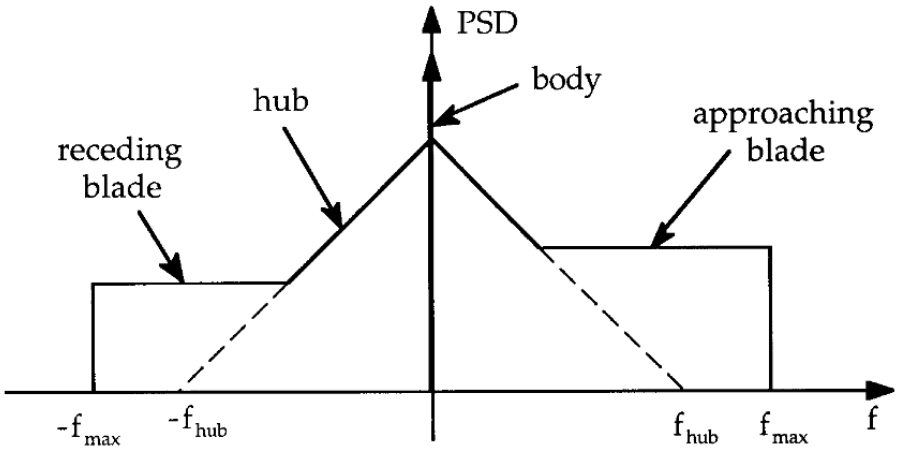


Figure 4.2: PSD of signal backscattered from the parts of the Hovering Helicopter, [1]

4.1.2 Autoregressive Modeling

A parametric model is selected to model a given time series The only requirement, after the selection of the model, is to estimate the parameters of the model. The number of parameters and their values should be properly chosen to accurately model the time series. It is also desired to represent the time series with as few parameters as necessary.

Parametric modeling have 3 types: Autoregressive (AR) model, Moving Average (MA) model and the Autoregressive-Moving Average (ARMA) model. Each of them is appropriate for different PSD forms. The AR Model can be used for PSD's having sharp peaks, and not having deep valleys. The MA models is suitable for the opposite type of PSD's, i.e., having deep valleys and not sharp peaks. Because the ARMA model comprises the characteristics of this two models, it is appropriate for the spectra containing both of these contrast characteristics. [32]

In our detection procedure, each of the two hypothesis has a different power spectrum density (PSD) with sharp peaks. Therefore, they have to be modeled by AR processes. AR coefficients for each of the two hypotheses have to be known or estimated from the collected data. (The data used in this modeling is collected through the collaboration of ASELSAN and METU-EE members.) From the time series data, the autocorrelation matrix are estimated. Then, these autocorrelation estimations are used to find the AR coefficients of the autoregressive processes.

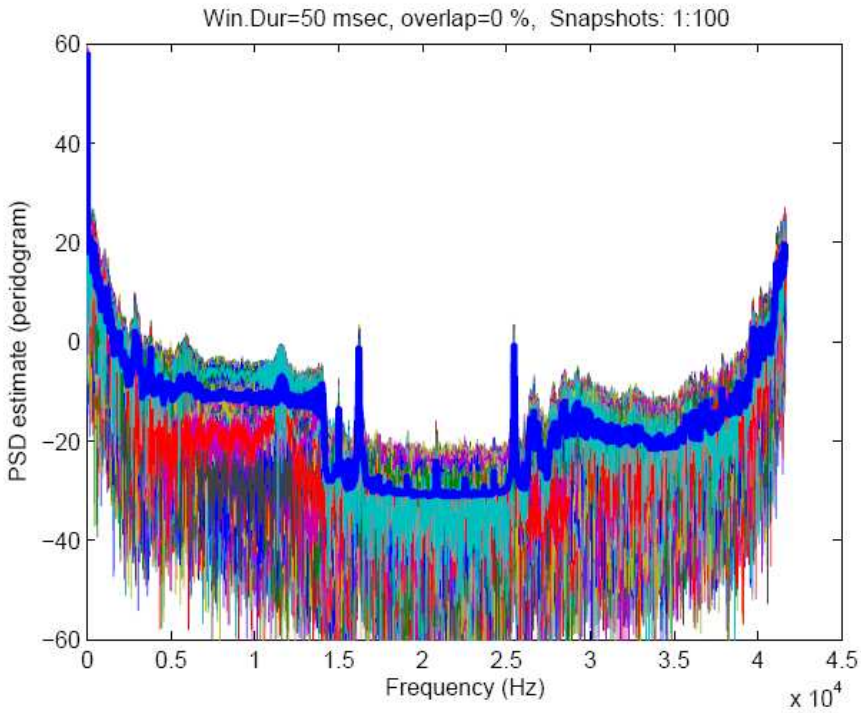


Figure 4.3: PSD estimate of the signal backscattered from the Helicopter

In order to model a process as an autoregressive process of the pth order AR(p), the parameter set of $(\sigma_x^2, a_p(1), a_p(2), \ldots, a_p(p))$ has to be calculated. If it is assumed that the AR process is generated by filtering white noise having unit variance,$w(n)$, then the all pole filter equation has the form of

$$H(z) = \frac{b(0)}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \tag{4.1}$$

Therefore, the Yule-Walker equations are satisfied by this autocorrelation sequence of AR process (Eqn. 4.2).

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \begin{cases} \sigma_w^2 |b(0)|^2 & ; \quad k = 0 \\ 0 & ; \quad k > 0 \end{cases} \tag{4.2}$$

Using the conjugate symmetry of $r_x(k)$, for $k = 1, 2, \ldots, p$, these equations can be written in matrix form as

$$\begin{bmatrix} r_x(0) & r_x^*(1) & \dots & r_x^*(p-1) \\ r_x(1) & r_x(0) & \dots & r_x^*(p-2) \\ \vdots & \vdots & & \vdots \\ r_x(p-1) & r_x(p-2) & \dots & r_x(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = \sigma_w^2 |b(0)|^2 \begin{bmatrix} r_x(1) \\ r_x(2) \\ \vdots \\ r_x(p) \end{bmatrix} \tag{4.3}$$

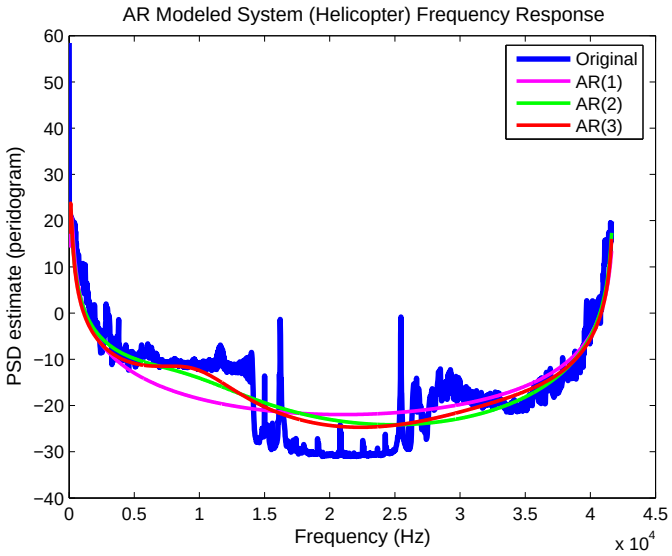
Yule-Walker method says that if the autocorrelations $r_x(k)$ are known or estimated, a linear equation system is to be solved for the AR coefficients. Therefore, using this method the AR spectral estimation can be done and the results are illustrated for different number of AR coefficient as

4.1.3 Modeling of Backscattering from a Fixed Wing Target

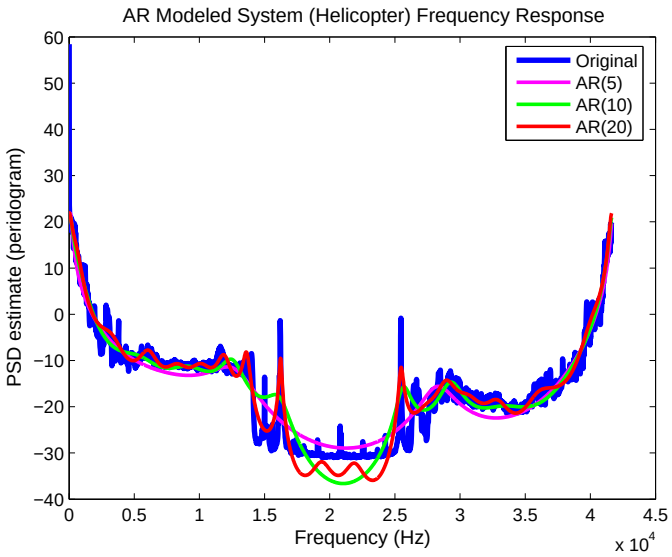
The backscattering of a radar pulse from a constant object contributes no Doppler shift, or a DC component after down conversion. If the object is moving, then it has a spread with respect to its unknown velocity which can be modeled as a random variable spread around the mean velocity. However, the spread of velocity for a fixed wing target is very low during coherent processing interval in comparison with the rotating parts of the helicopter. Therefore, the spectrum spread of the fixed wing target is narrower than the spread of the helicopter.

In addition, if the radar antenna is also rotating, the spread originating from this rotation can be wider than the spread due to velocity differences of the the fixed wing target. Therefore, the spread of PSD around the center frequency corresponding to mean velocity can be solely due to antenna rotation.

In this thesis, the fixed wing target is modeled as an AR(1) process with the correlation coefficient of -0.999. (This model is similar to the exponential clutter models.) In other words the model parameter is chosen such as the pole of the filter is close to z=1 in the z-plane. The final AR model coefficients for the hypotheses are listed in the Table 4.1. And the PSD estimates of the AR modeled hypotheses are illustrated in the Figure 4.5.



(a) PSD estimates of the AR modeled Helicopter signal AR(1), AR(2), AR(3)



(b) PSD estimates of the AR modeled Helicopter signal AR(5), AR(10), AR(20)

Figure 4.4: PSD estimates of the AR modeled Helicopter signal

4.2 Simulation Results and Analysis of the Algorithm with the Estimated Models

In this section, the performance of the classification algorithm is tested. For this aim, the complex white Gaussian noise is added to the input signals. CWGN has a negative effect on the performance. In order to see this negative effect, the SNR is decreased by adding CWGN until the algorithm can not operate at the desired false alarm and miss rates. False alarm and

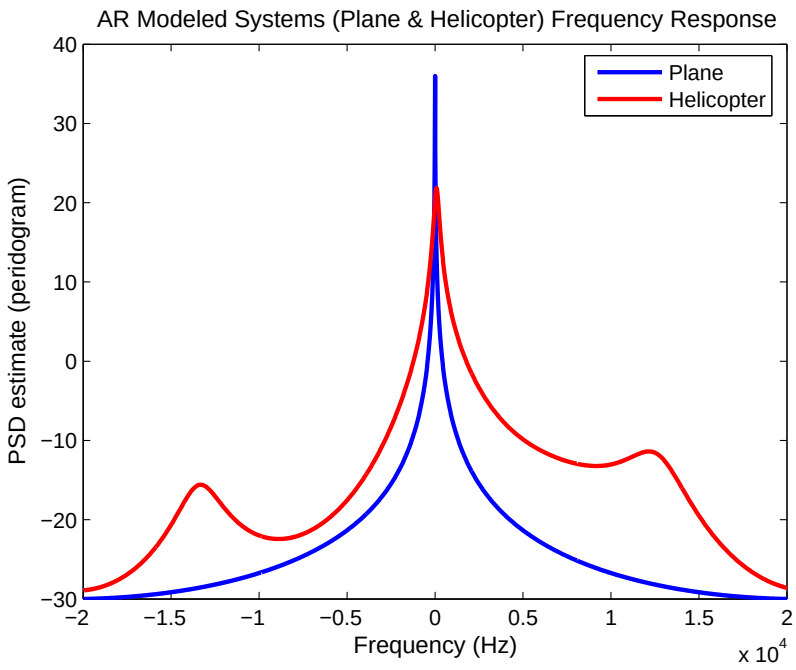


Figure 4.5: PSD estimates of the AR modeled Plane and Helicopter signals

miss rates are adjusted to ($\alpha = \beta = 10^{-6}$) for all of these simulations in this section.

By analyzing the PSD estimates of the echo returned from the helicopter (Figure 4.3), the echo signal (H_0 hypothesis) can be assumed to have a noise floor at -35 dB and therefore, the SNR of the experiment data under H_0 hypothesis can be calculated as 25 dB. At this SNR, the classiication algorithm with AR(5) helicopter model decides the target as helicopter at the samples n =[43 87] as illustrated in Figure 4.6.

To examine the effect of the used AR model order on the performance of the classification of helicopter, Monte-Carlo simulations are made for 2000 trials by using different AR helicopter

Table 4.1: AR coefficients for the fixed wing target or plane (H_0) and helicopter (H_1) hypotheses

Hypothesis	Plane (H_0)	Helicopter (H_1)
$a(1)$	-0.999	-1.2542-0.4250i
$a(2)$	0	0.5051+0.3334i
$a(3)$	0	-0.7718+0.0352i
$a(4)$	0	0.6383+0.2075i
$a(5)$	0	-0.1109-0.1649i

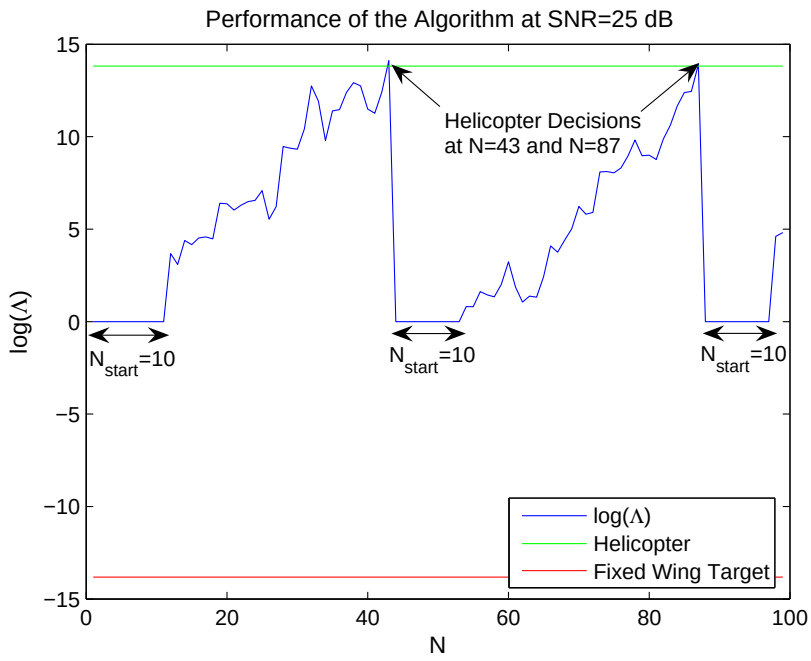


Figure 4.6: Performance of the Detection of Helicopter at 25 *dB* SNR

models having different number of coefficients. To obtain the histogram of the termination times, we run the algorithm for the same helicopter data by using AR(3), AR(5) and AR(10) helicopter models at 23 *dB* SNR.

AR(3) model results in the all trials are unterminated in other words no decision is made during 100 samples in 2000 trials. (Figure 4.7) By using AR(5) model, the %9.55 of the trials are unterminated but the remaining trials end with helicopter decision and no wrong decision is made. (Figure 4.8) In AR(10) model, %30.15 of the trials remain unterminated after 100 sample and more than half of the remaining trials end with plane decision so the error rate is 0.6822. (Figure 4.9)

Adding CWGN to the helicopter time series to get an SNR of 23 *dB* results in a delay in the detection time of the helicopter. While making the first decision at n =43 with 25 *dB* SNR, the algorithm makes helicopter decision with an ASN of 75.6795 with 23 *dB* SNR. (Figure 4.8)

To examine the performance of the algorithm with AR models having different number of coefficients, the test is repeated for 21 *dB* SNR. AR(3) model remains mostly unterminated, however this time, it makes 2 wrong decisions in 2000 trials. (Figure 4.10) In AR(5) model, the unterminated percent increases to %85.15 however the error rate does not increased in

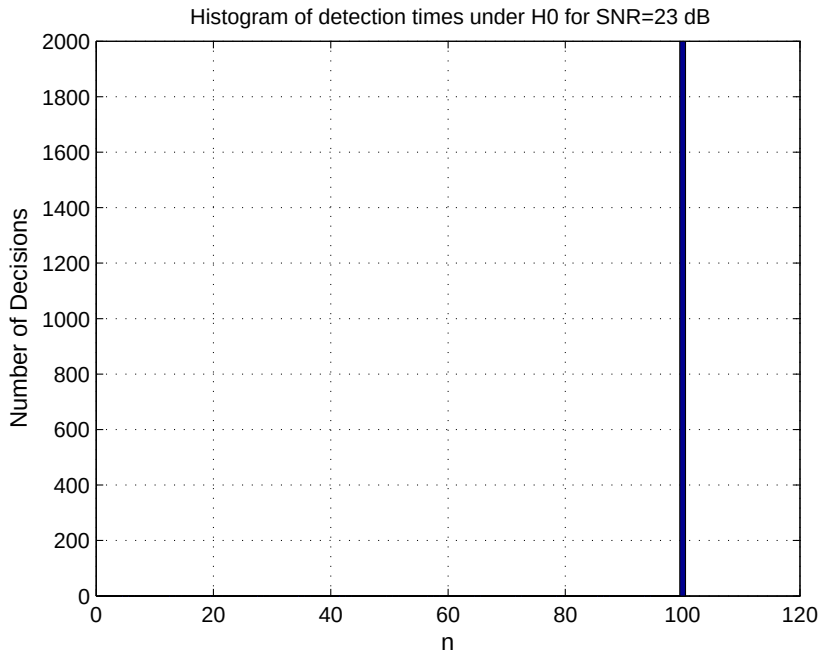


Figure 4.7: Helicopter classification(AR(3)) at 23 *dB* SNR, ASN=100, FA=0, Unterminated Percent=%100

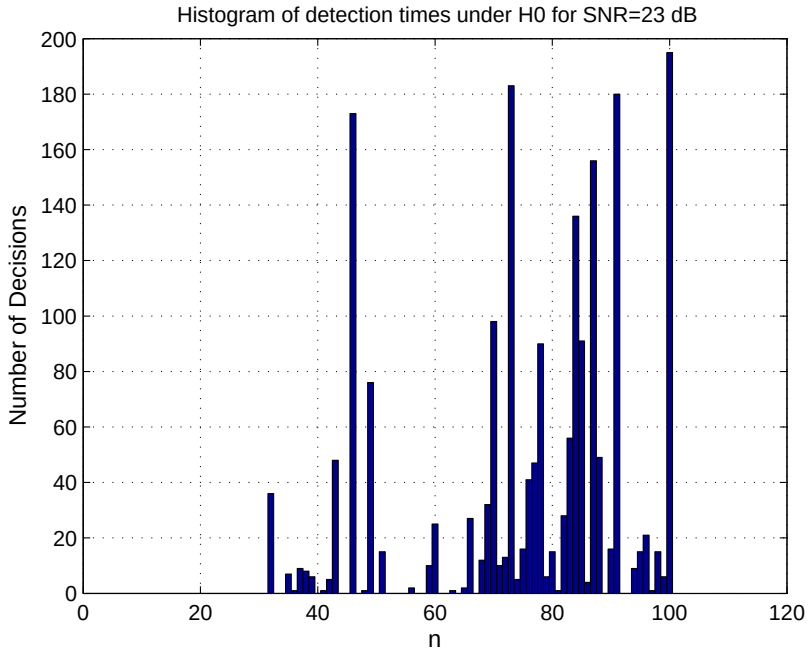


Figure 4.8: Helicopter classification(AR(5)) at 23 *dB* SNR, ASN=75.6795, FA=0, Unterminated Percent=%9.55

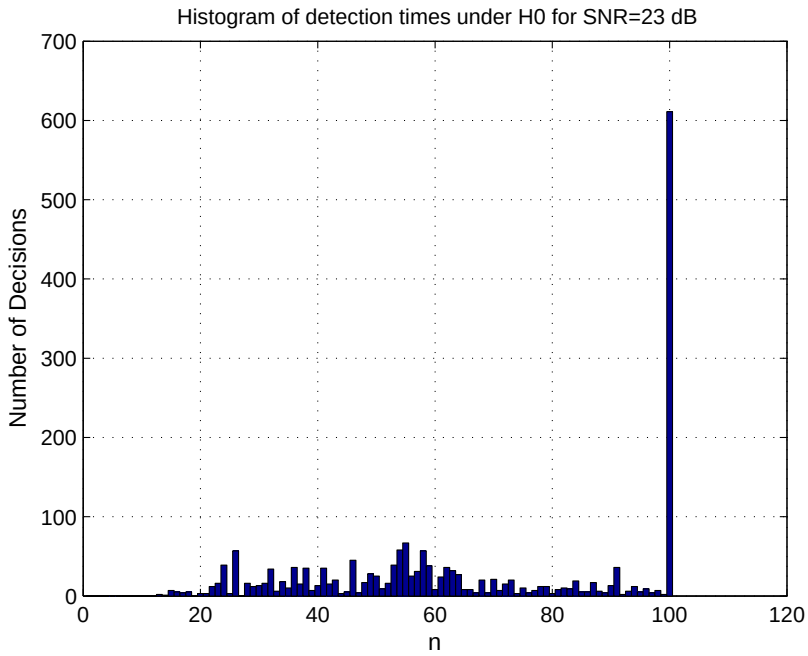


Figure 4.9: Helicopter classification(AR(10)) at 23 *dB* SNR, ASN=67.8885, FA=0.6822, Unterminated Percent=%30.15

comparison to other models. Addition of CWGN makes the classification algorithm more likely to decide plane hypothesis. (Figure 4.11) AR(10) model results in the fact that almost all of the trials end with wrong decision. (Figure 4.12)

In contrary to the expectations, the performance of the AR(10) model is not better than the two alternatives. The reason of this bad performance may be the fact that AR(10) model is modeling also the spurious peaks in the PSD estimate of the real helicopter data. These peaks do not stay in the portion of the used helicopter data and therefore this results in poor performance. All of these statistics show that the AR(5) model is more suitable for this classification algorithm among these three models. As a future work, the choice of AR model order will be studied.

For the plane hypothesis 25 *dB* SNR, the plane decisions are made in ASN of 22.7920. (Figure 4.13)

If the SNR value decreases to the 23 *dB* for the plane hypothesis, the ASN to make a decision decreases to 15.7410. (Figure 4.14)

In this chapter, the suggested method is applied on the problem of classification of rotary

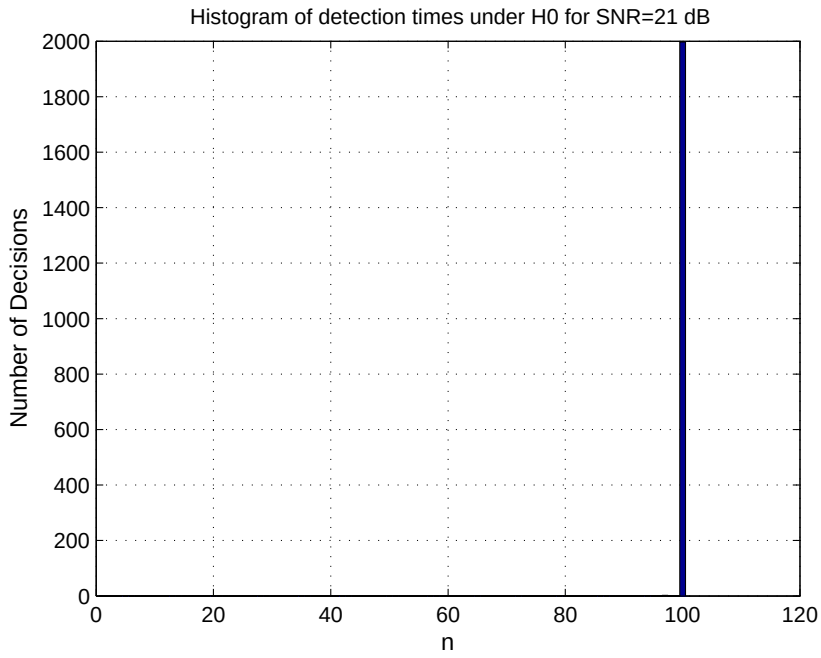


Figure 4.10: Helicopter classification(AR(3)) at 21 *dB* SNR, ASN=99.9895, FA=1, Unterminated Percent=%99.85

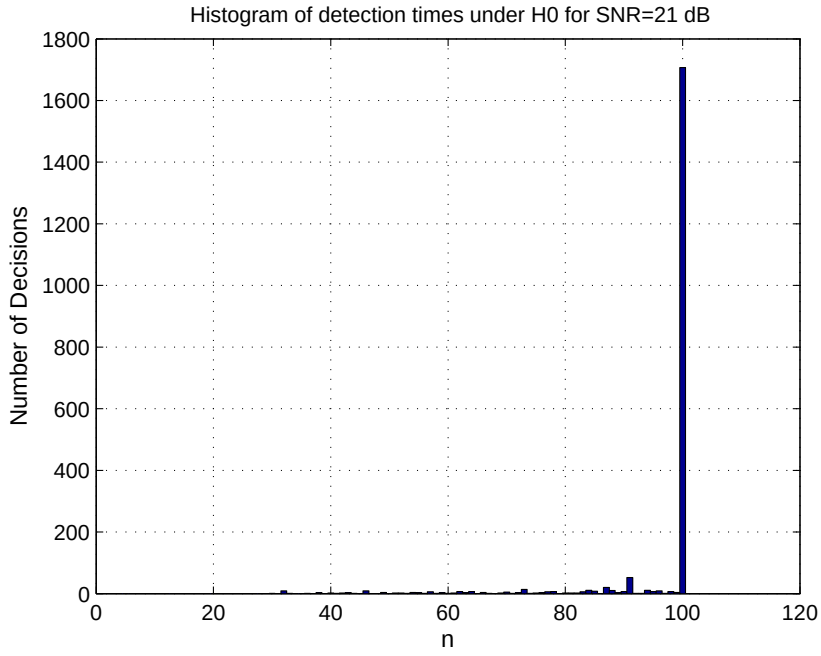


Figure 4.11: Helicopter classification(AR(5)) at 21 *dB* SNR, ASN=96.5335, FA=0.1684, Unterminated Percent=%85.15

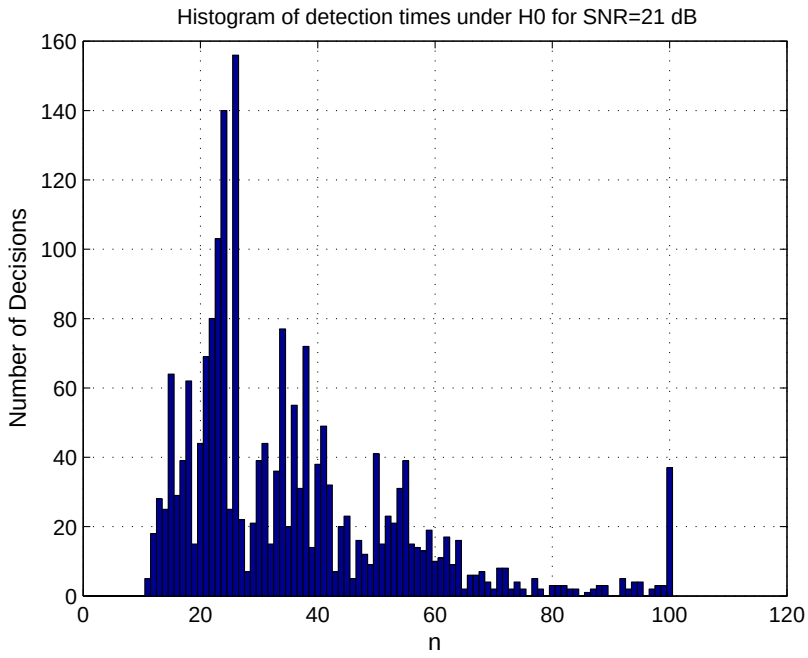


Figure 4.12: Helicopter classification(AR(10)) at 21 *dB* SNR, ASN=35.7990, FA=0.9903, Unterminated Percent=%1.75

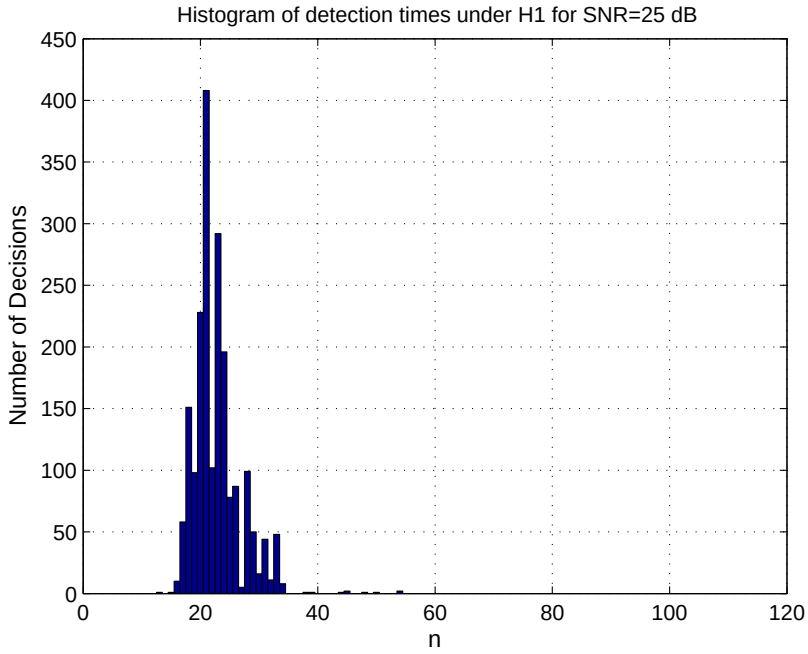


Figure 4.13: Plane classification at 25 *dB* SNR, ASN=22.7920, FA=0, Unterminated Percent=%0

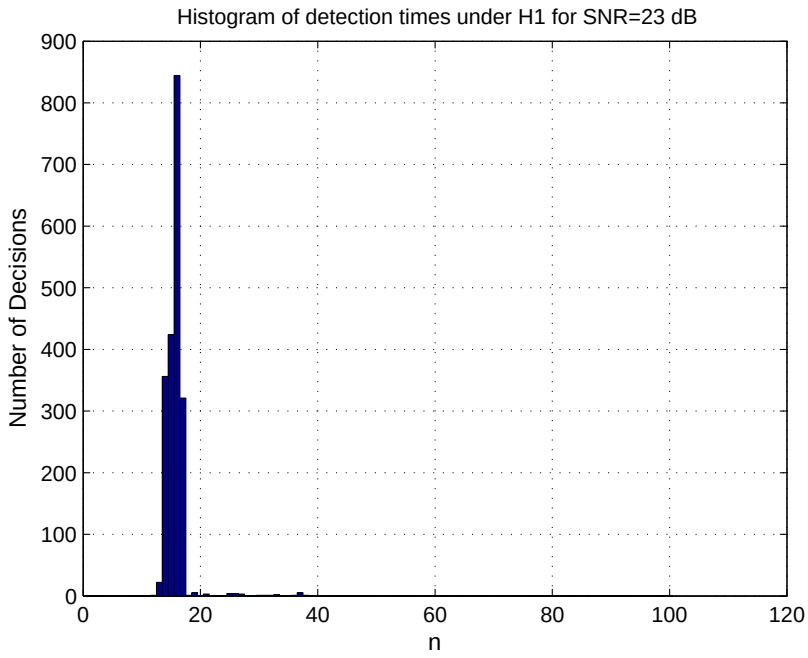


Figure 4.14: Plane classification at 23 *dB* SNR, ASN=15.7410, FA=0, Unterminated Percent=%0

- fixed wing target. To model the hypotheses with autoregressive models, we calculate the model coefficients by the help of Yule - Walker method. As the number of coefficient used in AR modeling increases, it is illustrated that the hypotheses are represented better. However, the increment in the number of coefficient also increases the complexity of the algorithm. The performance of the suggested classification algorithm is studied by using experimentally collected real data of an helicopter and simulated data of moving fixed wing target at different SNR levels. The CWGN has a negative effect on the classification of the helicopter by increasing the required samples to terminate the test while a positive effect on the moving fixed wing target hypothesis.

CHAPTER 5

CONCLUSION

In this thesis, a sequential method is presented for the classification of the autoregressive processes. This method requires fewer number of samples than the conventional systems having fixed sample sizes. The use of suggested method can result in a decrease in the dwell times in search systems. The simulation results show that the proposed sequential method requires about half the number of samples that the method with a predefined sample size.

The thresholds of the suggested method can be accurately calculated through Wald limits. An additional advantage of the sequential method is that the error probabilities of first and second type can be easily adjusted by two thresholds.

The complexity which arises from the need of inverting larger and larger matrices for each new coming sample is eliminated by using recursive algorithms to calculate inverse of the autocorrelation matrix and also the required determinant of the autocorrelation matrix value.

The method is examined for signals having different SNR values and it is observed that by increasing the decision start sample number (N_{start}) the algorithm is able to estimate the unknown process variance reasonably well and work with low SNR signals.

As an application example, the problem of fixed and rotary wing classification problem is examined. For this aim, the experimentally collected real data of an helicopter is modeled by different AR models and a fixed wing moving target is simulated. Using these models, the performance of the proposed classification algorithm is examined on experimental data at different SNR levels.

The future work related to the proposed method is the development of an analytical study

for the case of unknown process variance (a study on the accuracy of Wald's thresholds and the detection probabilities in the presence of estimation errors) and the development of the framework for AR signals under noise and an extension to the M-Ary hypothesis testing.

REFERENCES

[1] F. Gini and A. Farina, "Matched subspace CFAR detection of hovering helicopters," *IEEE Trans. Aerospace and Electronic Systems*, vol. 35, pp. 1293–1305, Oct 1999.

[2] H. L. Van Trees, *Detection, Estimation and Modulation Theory, part 1*. John Wiley - Sons, 1971.

[3] A. Wald, *Sequential Analysis*. New York: Wiley, 1947.

[4] A. Wald and J. Wolfowitz, "Optimum Character of the Sequential Probability Ratio Test," *Ann. Math. Stat*, vol. 19, 1948.

[5] R. Niu and P. Varshney, "Sampling Schemes for Sequential Detection With Dependent Observations," *IEEE Trans. Signal Processing*, vol. 58, pp. 1469 –1481, March 2010.

[6] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes*. New Jersey: Prentice Hall, 1993.

[7] E. S. Page, "Continuous Inspection Schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.

[8] W. El Falou, M. Khalil, and J. Duchene, "AR-based method for change detection using dynamic cumulative sum ," in *The 7th IEEE International Conference on Electronics, Circuits and Systems*, vol. 1, pp. 157 –160 vol.1, 2000.

[9] S. Dey and S. Marcus, "Change detection in Markov-modulated time series," in *Information, Decision and Control, 1999. IDC 99. Proceedings. 1999*, pp. 21 –24, 1999.

[10] E. Gombay, "Change detection in autoregressive time series," *Journal of Multivariate Analysis*, vol. 99, no. 3, pp. 451 – 464, 2008.

[11] J. Wang and A. Nehorai, "Sequential Detection for a Target in Compound-Gaussian Clutter ," in *Fortieth Asimolar Conference on Signals, Systems and Computers, 2006.*, 2006.

[12] T. Sastri, "A Recursive Algorithm for Adaptive Estimation and Parameter Change Detection of Time Series Models ," in *J. Opl. Res. Soc.*, vol. 37, pp. 987 –999, 1986.

[13] R. Cmejla and P. Sovka, "Recursive Bayesian Autoregressive Changepoint Detector for Sequential Signal Segmentation," *In EUSIPCO-2004*, 2004.

[14] F. Gustafsson, *Adaptive Filtering and Change Detection*. Wiley, 2000.

[15] W. Chung and C. Un, "Iterative autoregressive parameter estimation in presence of additive white noise," *Electronics Letters*, vol. 27, pp. 1800 –1802, Sep 1991.

[16] S. de Waele and P. M. T. Broersen, "Modeling Radar Data with Time Series Models," *In EUSIPCO-2000*, 2000.

[17] C.-K. Yeh and P.-C. Li, "Doppler angle estimation using AR modeling," *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 49, pp. 683 –692, Jun 2002.

[18] D. Abraham and P. Willett, "Active sonar detection in shallow water using the Page test," *IEEE Journal of Oceanic Engineering*, vol. 27, pp. 35 –46, Jan 2002.

[19] X. Chen and U. Tureli, "Passive Acoustic Detection of Divers Using Single Hydrophone," in *Fortieth Asilomar Conference on Signals, Systems and Computers, 2006.*, pp. 554 –558, Nov 2006.

[20] X. Chen, R. Wang, and U. Tureli, "Passive Acoustic Detection of Divers Under Strong Interference," in *OCEANS 2006*, pp. 1 –6, Sep 2006.

[21] S. Kay and J. Salisbury, "Improved active sonar detection using autoregressive prewhiteners," *Journal of the Acoustical Society of America*, vol. 87, pp. 1603–1611, Apr 1990.

[22] V. Carmillet and G. Jourdain, "Low-speed targets sonar detection using autoregressive models in reverberation; experimental performances for wideband signals," in *OCEANS '98 Conference Proceedings*, vol. 3, pp. 1285 –1289 vol.3, Oct 1998.

[23] T. Thayaparan, S. Abrol, E. Riseborough, L. Stankovic, D. Lamothe, and G. Duff, "Analysis of radar micro-Doppler signatures from experimental helicopter and human data," *IET Radar, Sonar Navigation*, vol. 1, pp. 289 –299, Aug 2007.

[24] S. Yang, S. Yeh, S. Bor, S. Huang, and C. Hwang, "Electromagnetic backscattering from aircraft propeller blades," *IEEE Trans. on Magnetics*, vol. 33, pp. 1432 –1435, Mar 1997.

[25] J. Misiurewicz, K. Kulpa, and Z. Czekala, "Analysis of recorded helicopter echo," in *Radar 97 (Conf. Publ. No. 449)*, pp. 449 –453, Oct 1997.

[26] J. Misiurewicz, K. Kulpa, and Z. Czekala, "Analysis of radar echo from a helicopter rotor hub," in *Microwaves and Radar, 1998. MIKON '98., 12th International Conference on*, vol. 3, pp. 866 –870 vol.3, May 1998.

[27] H. Haghshenas and M. Nayebi, "A novel method to detect rotor blades echo," in *Radar Conference, 2010 IEEE*, pp. 1331 –1334, May 2010.

[28] T. J. Abatzoglou, "Helicopter Recognition Radar Processor," *United States Patent*, Jul 2006.

[29] F. Gini, M. Greco, and F. Farina, "Radar detection and preclassification based on multiple hypothesis," *IEEE Trans. Aerospace and Electronic Systems*, vol. 40, pp. 1046 –1059, Jul 2004.

[30] C. W. Therrien, "A Sequential Approach to Target Discrimination," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-14, pp. 433 –440, May 1978.

[31] Q. Zhang, K. Wong, and S. Haykin, "New parametric method of detecting random signals in unknown noise. 2. Application to radar," *IEE Proceedings F Radar and Signal Processing*, vol. 139, pp. 359 –364, Oct 1992.

[32] S. M. Kay, *Modern Spectral Estimation*. New Jersey: Prentice Hall, 1988.

[33] M. H. Hayes, *Statistical Signal Processing and Modeling*. Wiley, 1996.

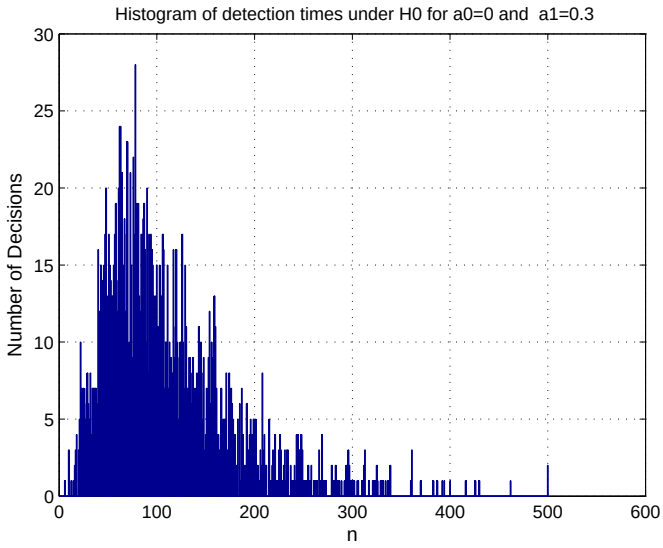
[34] J. Francos and B. Friedlander, "Bounds for estimation of complex exponentials in unknown colored noise," *IEEE Trans. Signal Processing*, vol. 43, pp. 2176 –2185, Sep 1995.

APPENDIX A

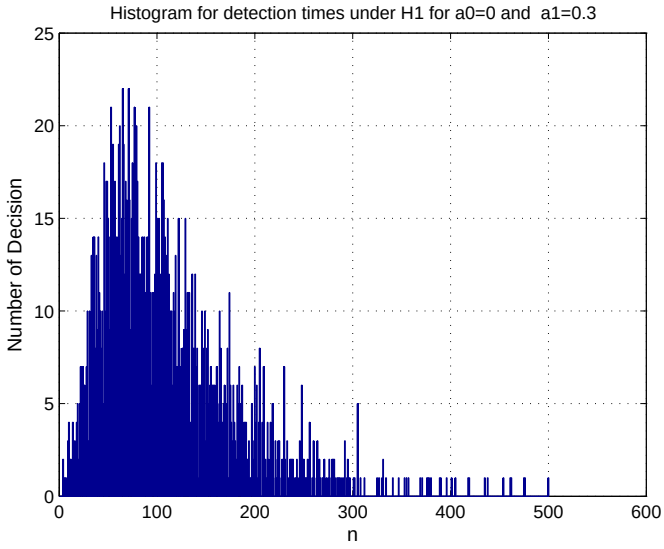
EFFECTS OF THE VALUE OF THE AR COEFFICIENTS TO THE PERFORMANCE OF THE ALGORITHM

A.1 Known Power Case

In this section, the effect of chosen AR coefficients on the false alarm and miss rates, ASN and the number of unterminated trials are investigated by means of numerical comparisons through Monte Carlo simulations for the known power case. The parameters are set such that the false alarm rate $\alpha = 0.01$ and the miss rate $\beta = 0.01$. Also Monte Carlo simulations are done for 2000 trials and each data vector have a length of 500. Firstly, the effect of the AR coefficients on the performance of the algorithm for the known power case is examined and the simulation results are illustrated.

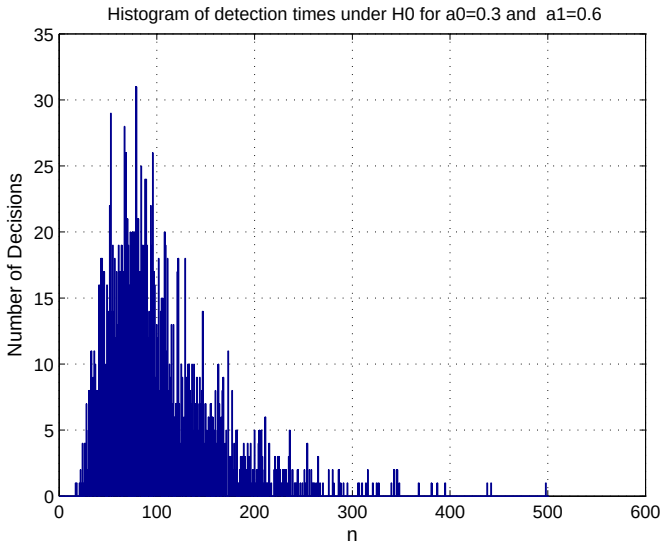


(a) $a_0 = 0$, a_1=0.3, FA=0.0065, ASN=110.1410, Unterminated Percent 0.1000%

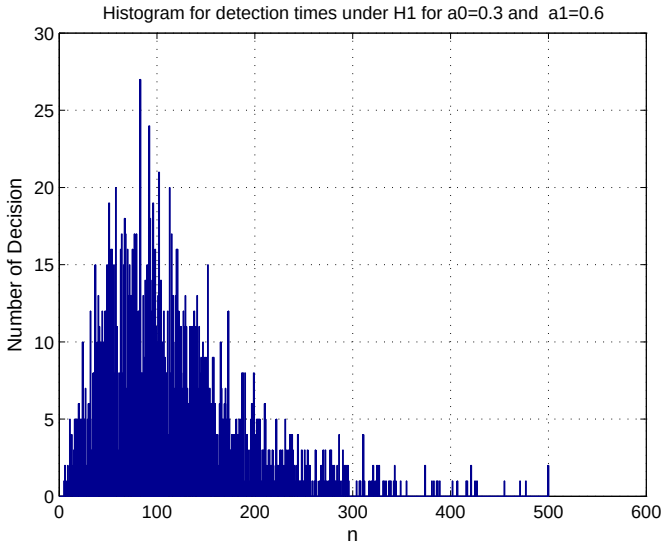


(b) $a_0 = 0$, a_1=0.3, MISS=0.0050, ASN=111.3980, Unterminated Percent 0.0500%

Figure A.1: Algorithm performance for known variance case, $a_0 = 0$, a_1=0.3

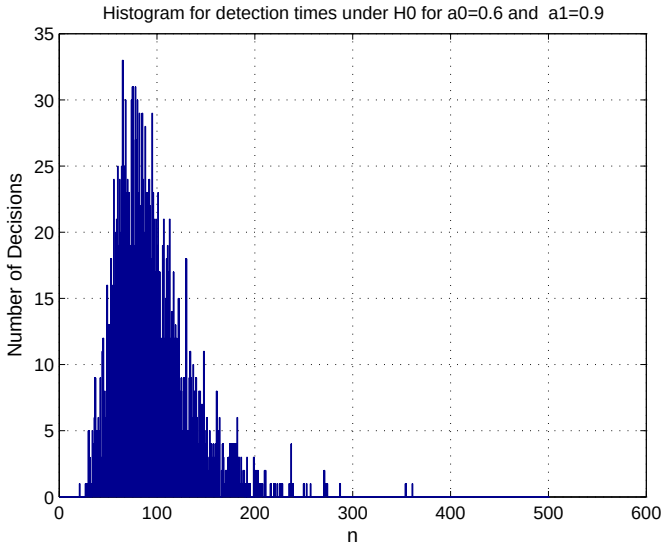


(a) a_0 = 0.3, a_1=0.6, FA=0.0040, ASN=104.8395, Unterminated Percent 0%

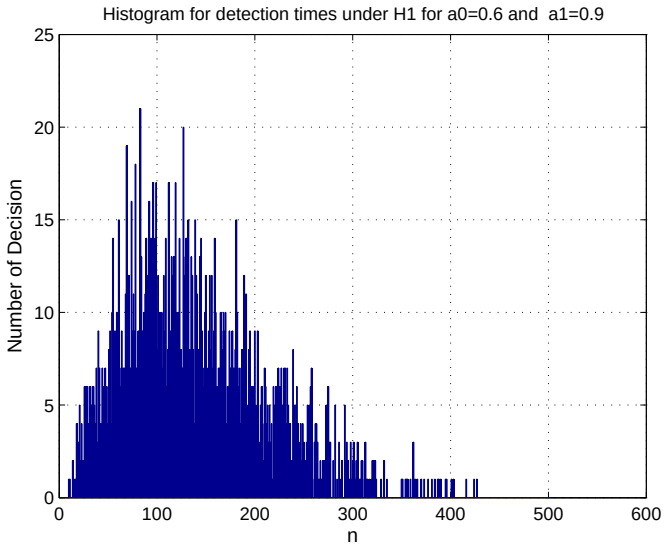


(b) a_0 = 0.3, a_1=0.6, MISS=0.0005, ASN=117.5130, Unterminated Percent 0.0500%

Figure A.2: Algorithm performance for known variance case, a_0 = 0.3, a_1=0.6



(a) a_0 = 0.6, a_1=0.9, FA=0, ASN=95.5640, Unterminated Percent 0%

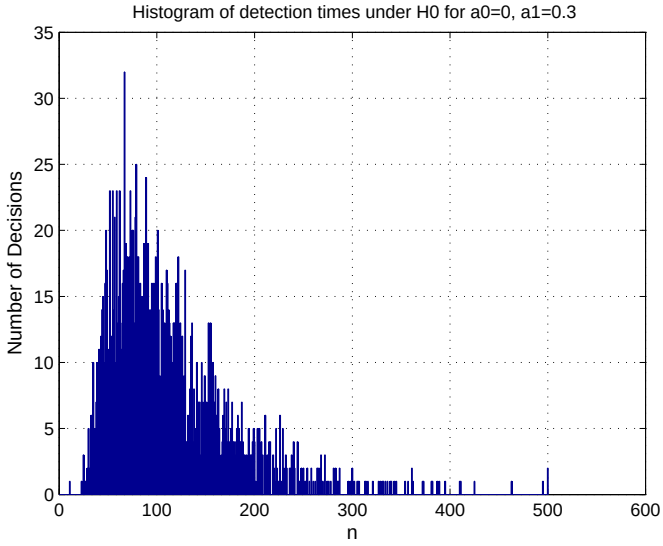


(b) a_0 = 0.6, a_1=0.9, MISS=0, ASN=140.9580, Unterminated Percent 0%

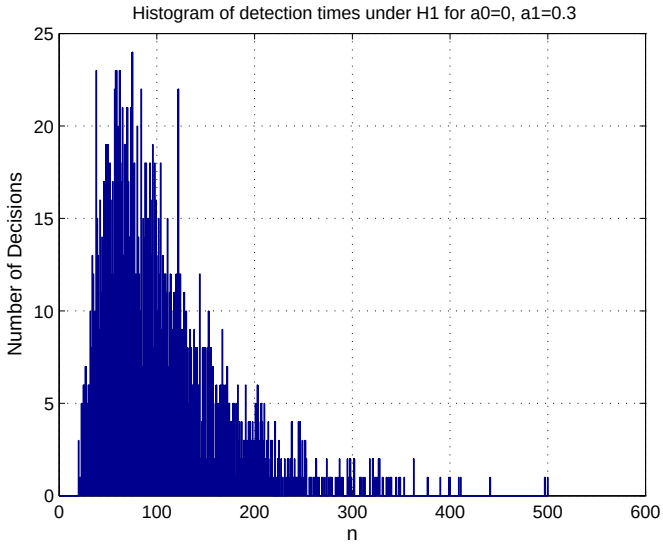
Figure A.3: Algorithm performance for known variance case, a_0 = 0.6, a_1=0.9

A.2 Unknown Power Case

In the previous section, the effect of chosen AR coefficients on the false alarm and miss rates, ASN and the number of unterminated trials are investigated by means of numerical comparisons through Monte Carlo simulations for the known power case. In this section, The effect of the AR coefficients on the performance of the algorithm is examined for the unknown power case.

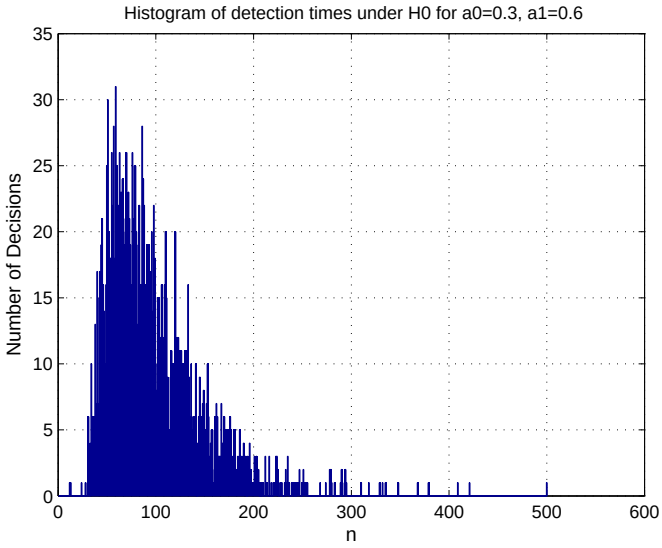


(a) $a_0 = 0$, a_1=0.3, FA=0.0070, ASN=112.8660, Unterminated Percent 0.1000%

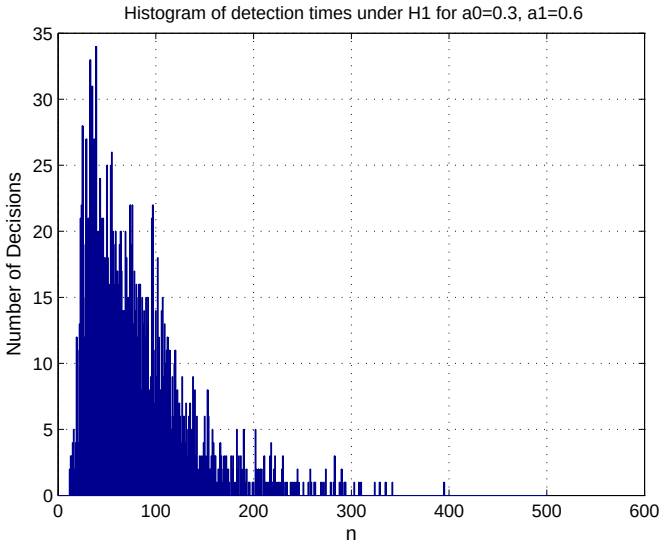


(b) $a_0 = 0$, a_1=0.3, MISS=0.0075, ASN=107.6575, Unterminated Percent 0.0500%

Figure A.4: Algorithm performance for unknown variance case, $a_0 = 0$, a_1=0.3

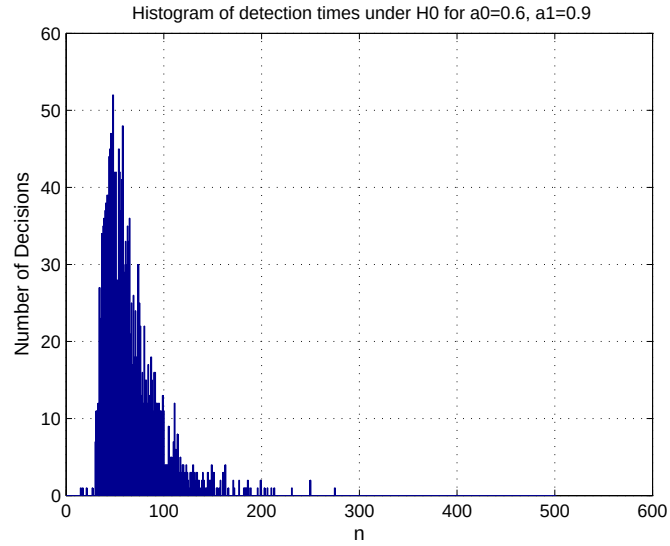


(a) a_0 = 0.3, a_1=0.6, FA=0.0080, ASN=97.5770, Unterminated Percent 0.0500%

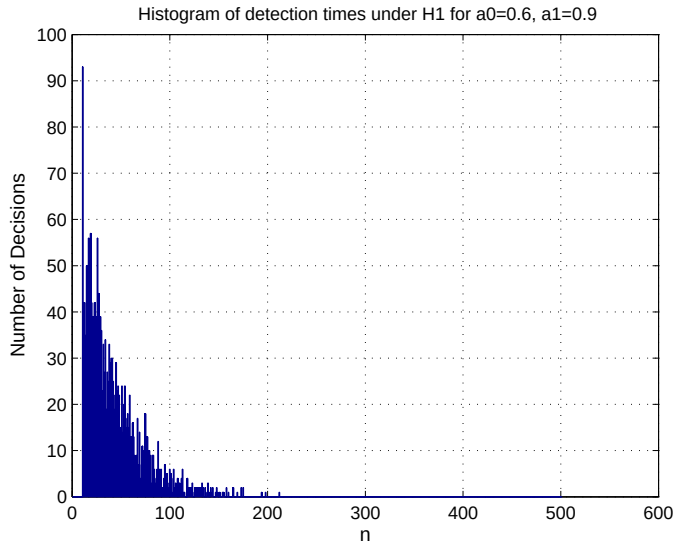


(b) a_0 = 0.3, a_1=0.6, MISS=0.0080, ASN=81.4765, Unterminated Percent 0%

Figure A.5: Algorithm performance for unknown variance case, a_0 = 0.3, a_1=0.6



(a) a_0 = 0.6, a_1=0.9, FA=0.0040, ASN=66.5160, Unterminated Percent 0%



(b) a_0 = 0.6, a_1=0.9, MISS=0.0085, ASN=42.7900, Unterminated Percent 0%

Figure A.6: Algorithm performance for unknown variance case, a_0 = 0.6, a_1=0.9