Master's Thesis

Design of a Constant False Alarm Rate (CFAR) detection scheme

Using the sequential likelihood ratio test in a TBD approach

Tom Sniekers

August 14, 2015

Supervisors
Dr. R. Bakker (Thales)
Dr. M. Bocquel (Thales/UT)
Dr. P.K. Mandal (UT)

Assessment committee
Prof.dr. A.A. Stoorvogel (UT)
Dr. P.K. Mandal (UT)
Dr. R. Bakker (Thales)
Dr. M. Bocquel (Thales/UT)
Dr. G.J. Still (UT)

UNIVERSITY OF TWENTE.

Department: Applied Mathematics

Chair: Hybrid Systems
Enschede

THALES
Sensors Development

System Engineering
Hengelo

ABSTRACT

In this Master's Thesis the problem of variation in the probability of false alarm is treated, when testing the null hypothesis "Target not present" against the alternative hypothesis "Target present", using sequential likelihood ratio tests in noisy and clutter background environments. The hypothesis testing is performed with a Track Before Detect (TBD) approach, where the tracking is performed by a Sequential Markov Chain Monte Carlo (SMCMC) filter. Insurmountable problems with applying Wald's theory with filtered data are discussed. A new sequential detection scheme is designed that uses Neyman-Pearson theory in a sequential setting. Several models are proposed to control the probability of false alarm of this detection scheme by integrating these models in the threshold calculations. The models can be applied within a Gaussian noise environment. Furthermore, expressions are given to properly apply the models in a (sea) clutter. The results for applying the new sequential detection scheme, with integrated models, fulfill the desired Constant False Alarm Rate (CFAR) property in noise. Finally, a proposed implementation of the detection scheme in the SMCMC filter is given.

PREFACE

This Master's Thesis is a results of the research done during my final project of the master Applied Mathematics at the University of Twente. This final project is the last step in achieving the academic title Master of Science (Msc). The research is performed at Thales B.V. in Hengelo and took place between January 2015 and August 2015. The purpose of the research was to design a sequential detection scheme with a constant false alarm rate.

First of all, I would like to thank my supervisors at Thales, Rienk Bakker and Mélanie Bocquel for their intensive and passioned support during my research. I also would like to thank my supervisor from the University of Twente, Pranab Mandal, for his guidance. Apart form my supervisors, I want to sincerely thank my colleague master student, Sijmen de Bruijn, with whom I spend the coffee and lunch breaks. A last word of thanks go to employees at Thales, who were very friendly and helpful when I asked for help.

Contents

1 Introduction **1**
1.1 Problem description . . . 1
1.2 Outline . . . 1

2 Background knowledge **3**
2.1 Radar Detection . . . 3
2.1.1 Classical Radar detection . . . 5
2.1.2 Track Before Detect . . . 5
2.2 Filtering . . . 6
2.2.1 Bayes Filter . . . 6
2.2.2 Particle Filtering . . . 7
2.2.3 Sequential Markov Chain Monte Carlo . . . 10
2.3 Important specifications of the filter . . . 13
2.3.1 Background models . . . 13
2.3.2 Target model . . . 14
2.3.3 Target + background model . . . 15
2.3.4 Measurement model . . . 18
2.3.5 Metropolis-Hastings algorithm . . . 18

3 Hypothesis Testing **21**
3.1 Sequential Likelihood Ratio Test . . . 22
3.1.1 Wald's theory . . . 22
3.1.2 Neyman-Pearson testing . . . 28
3.2 Simulation results i.i.d. data . . . 30
3.2.1 Results Wald's theory with i.i.d. data . . . 30
3.2.2 Results Neyman-Pearson theory with i.i.d. data . . . 34
3.3 Simulation results with data from particle filter . . . 37
3.3.1 Results Wald's theory with filtered data . . . 37
3.3.2 Results Neyman-Pearson theory with filtered data . . . 37

4 Problems with filtered data and Solutions **41**
4.1 Observations in filter . . . 41
4.1.1 RCS fluctuations . . . 46
4.1.2 Spreading particles . . . 46
4.1.3 MCMC step . . . 47
4.1.4 Swerling Case . . . 48
4.2 Models as solution . . . 52
4.2.1 Handling the RCS fluctuations . . . 52

4.2.2 Model spreading particles . . . 52
4.2.3 Proposed model MCMC step . . . 57
4.3 Results proposed models . . . 65
4.4 Proposed CFAR detection scheme . . . 71

5 Conclusions and recommendations **79**
5.1 Discussion . . . 79
5.2 Conclusions . . . 81
5.3 Contribution to Thales . . . 82
5.4 Recommendations for further research . . . 82

Bibliography **87**

A Different domain calculations **89**
A.1 Gaussian Noise . . . 89
A.2 Generalized Pareto . . . 93

B Distributions Target + Noise/Clutter **95**
B.1 Target + Noise . . . 95
B.2 Target and Clutter . . . 96
B.2.1 Analytical . . . 96
B.2.2 Parameter estimation . . . 97

C Distribution of sample mean of i.i.d Pareto distributed r.v.'s **99**

D Histograms from the Filter **105**

List of Figures

2.1 Radar principles 4
2.2 Classical data and signal processing (separate boxes) and TBD (large box) . . . 6

3.1 P_{FA} and P_D curves with Wald's criteria for i.i.d. noise samples 32
3.2 P_{FA} and P_D curves with Wald's criteria for i.i.d. clutter samples 33
3.3 Threshold $\ln(\tau)$ for Neyman-Pearson theory 35
3.4 P_{FA} and P_D curves for Neyman-Pearson theory, with i.i.d. samples 36
3.5 P_{FA} and P_D curves Wald's theory in noise with data from filter 38
3.6 P_{FA} and P_D curves Wald's theory in clutter with data from filter 39
3.7 P_{FA} and P_D curves Neyman-Pearson theory in noise with data from filter 40

4.1 Filter output track in noise 43
4.2 Filter output track in clutter 44
4.3 Filter output track on target in noise 45
4.4 Difference P_{FA} curves between predicted and posterior likelihood ratios 49
4.5 Difference P_D curves between predicted and posterior likelihood ratios 50
4.6 Difference in P_D between Swerling I and Swerling 0 target. 51
4.7 Graphical representation for model MCMC step 62
4.8 Predicted distributions 63
4.9 Posterior distributions 64
4.10 P_{FA} curves for predicted likelihood ratios with and without models implemented 67
4.11 P_D curves for predicted likelihood ratios with and without models implemented . 68
4.12 Thresholds $\ln(\tau)$ using Neyman-Pearson theory with and without models implemented 69
4.13 Results for empirically tuned values of c_n 70
4.14 P_{FA} curves proposed detection scheme 75
4.15 P_D curves proposed detection scheme 76
4.16 P_{FA} curves with scaling of $\lambda(\mathbf{x}_n)$ 77
4.17 P_D curves with scaling of $\lambda(\mathbf{x}_n)$ 78

B.1 Accuracy parameter estimation of distribution target+clutter 98

D.1 Predicted likelihood ratio distribution 106
D.2 Posterior likelihood ratio distribution 107

List of Algorithms

1 SIR particle filter 9
2 SMCMC 11
3 Metropolis-Hastings algorithm 12
4 Metropolis-Hastings algorithm in SMCMC 19

5 Metropolis-Hastings algorithm with segments 58
6 SMCMC with integrated CFAR detection scheme. 74

List of Acronyms

a.s.	almost surely
cdf	cumulative distribution function
CFAR	Constant False Alarm Rate
CLT	Central Limit Theorem
i.i.d.	independent identically distributed
LRT	Likelihood Ratio Test
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
MH	Metropolis-Hastings
pdf	probability density function
pmf	probability mass function
PSF	Point Spread Function
Radar	RAdio Detection And Ranging
RCS	Radar Cross Section
Rf	Radio frequency
ROC	Receiver Operating Characteristics
r.v.	random variable
SCR	Signal to Clutter Ratio
SIR	Sampling Importance Re sampling
SLLN	Strong Law Large Numbers
SLRT	Sequential Likelihood Ratio Test
SMC	Sequential Monte Carlo
SMCMC	Sequential Markov Chain Monte Carlo
SNR	Signal to Noise Ratio
TBD	Track Before Detect

List of Symbols

$\mathbb{N}, \mathbb{R}, \mathbb{C}$	Sets of respectively positive integers, real numbers and complex numbers
$\mathbb{R}^n$	n dimensional Euclidean space
$\mathbb{R}^+$	Positive real numbers (including 0)
$\mathbb{E}[\cdot]$	Expectation operator
$\mathrm{Var}[\cdot]$	Variance operator
$\sum(\cdot)$	Sum operator
$\prod(\cdot)$	Product operator
$\emptyset$	Empty set
$\lvert\cdot\rvert$	Cardinality of a set
$\mathcal{H}_i$	Hypothesis i, i = 0,1
$\mathcal{N}ull$	Null space of a vector
$\mathcal{N}(\cdot)$	Normal distribution
$\mathcal{U}[\cdot]$	Uniform distribution
$\mathcal{X}$	State space
$\mathcal{Z}$	Observation space
α	Significance level of a hypothesis test
$1-\beta$	Power of a hypothesis test
B	Number of Burn in iteration
$\delta_x(\cdot)$	Delta Dirac function, centered at x
$\lambda(x_n)$	Likelihood ratio at time n
$\lambda(\mathbf{x}_n)$	Likelihood ratio up to and including time n
$\lambda^T(\tau_n)$	Threshold likelihood ratio at time n
$\lambda^D(x_n)$	Detection likelihood ratio at time n
$\lambda^D(\mathbf{x}_n)$	Detection likelihood ratio up to and including time n
NoP	Number of Particles
P	Probability measure
p	Probability density function
p_n	Posterior density at time n
$p_{n\|n-1}$	Predicted density from time $n-1$ to n
τ	Threshold of hypothesis test
τ_n	Threshold of hypothesis test at time n
x_n	State at time n
$\mathbf{x}_n$	State up to and including time n
w_n^j	Weight of particle j at time n
z_n	Observation at time n
$z_{1:n}$	Sequence of observations up to and including time n

Chapter 1

Introduction

Radars are used to detect and track moving or fixed objects. Detection schemes are used to decide whether or not an object (target) is present. Measures of the performance of a detection scheme are the probability of detection and the probability of a false alarm. Detection schemes used for detecting large objects, show good performances in general. The detection of small objects is in general a lot more difficult, especially in a clutter background. Special tracking filters are used to track possible small objects in order to make an informed decision over the presence of targets. Current detection schemes, employed within Thales, for filtering based Radar detection of small targets in sea clutter can show a large variation in the false alarm rate with varying sea conditions. Such a detection scheme is the subject of this thesis.

1.1 Problem description

The goal of this thesis is to design a Constant False Alarm Rate (CFAR) detector. This means that regardless of the background noise/(sea) clutter, the detector exhibits a constant (fixed) probability of false alarm. The detector is meant to sequentially detect within a Track Before Detect (TBD) approach, using a sequential likelihood ratio test to test the null hypothesis "Target not present" against the alternative hypothesis "Target present". A particle filter is used to approximate target and clutter distributions.

1.2 Outline

In Chapter 2, the reader is provided with knowledge about basic Radar detection procedures, filtering processes and target and environment modulations. In Chapter 3, Wald's theory and Neyman-Pearson theory are discussed and tested with both i.i.d. generated data and filtered data from the particle filter. Also a newly designed sequential detection scheme will be introduced that uses Neyman-Pearson theory in a sequential setting. In Chapter 4, problems when applying current detection schemes are discussed, together with possible explanations. These problems consist of applying Wald's theory and Neyman-Pearson theory with filtered data, due to particle filter characteristics. Also models to avoid these problems and to improve the performance of the detection schemes are discussed. At the end of Chapter 4, the results of these models are presented with ultimately the presentation of the designed CFAR sequential detection scheme. In Chapter 5, the research is summarized with in Section 5.1 some analyses and comments about the results. In Section 5.2, the conclusions of this research are presented. In Section 5.3 the

contributions of the research to Thales are presented. In Section 5.4, several recommendations for further research are proposed. These recommendations are interesting for Thales when applying the derived models/methods in practical Radar applications. Appendices A to D will provide you with some further knowledge about the used distributions, made assumptions and obtained results.

Chapter 2

Background knowledge

RAdio Detection And Ranging (RADAR) systems are widely used for safety purposes, for example at airports to safely regulate the air traffic or in a military context to defend against hostile missiles.
This chapter provides the basics of Radar tracking and detecting. Firstly, the basic principles of the Radar and the differences between classical Radar detection and a Track Before Detect (TBD) approach are discussed. Secondly, the filtering process used in the (TBD) approach is discussed, where theoretical filters are distinguished from filters that use numerical approximations for practical implementations purposes. At last, some important properties of the filter are discussed in detail to provide the information needed in this thesis.

2.1 Radar Detection

A Radar uses electromagnetic pulses. Radio-frequency (Rf) energy is transmitted to, and reflected from, a reflecting object. The reflected energy is called echo. Only a small portion of the reflected energy is received by the radar. The process from generating Rf energy to the video on the screen is shown in Figure 2.1. A Radar transmitter produces short high power Rf pulses. These pulses are transmitted towards a flying or cruising object by the antenna. The reflected energy (echo signal) is also received by the antenna. However, the antenna can not transmit and receive at the same time. When the antenna is transmitting, the antenna is considered blind. The duplexer manages the switching between transmitter and receiver. The received echo signal is processed by a signal processor and finally displayed as a video on the screen.

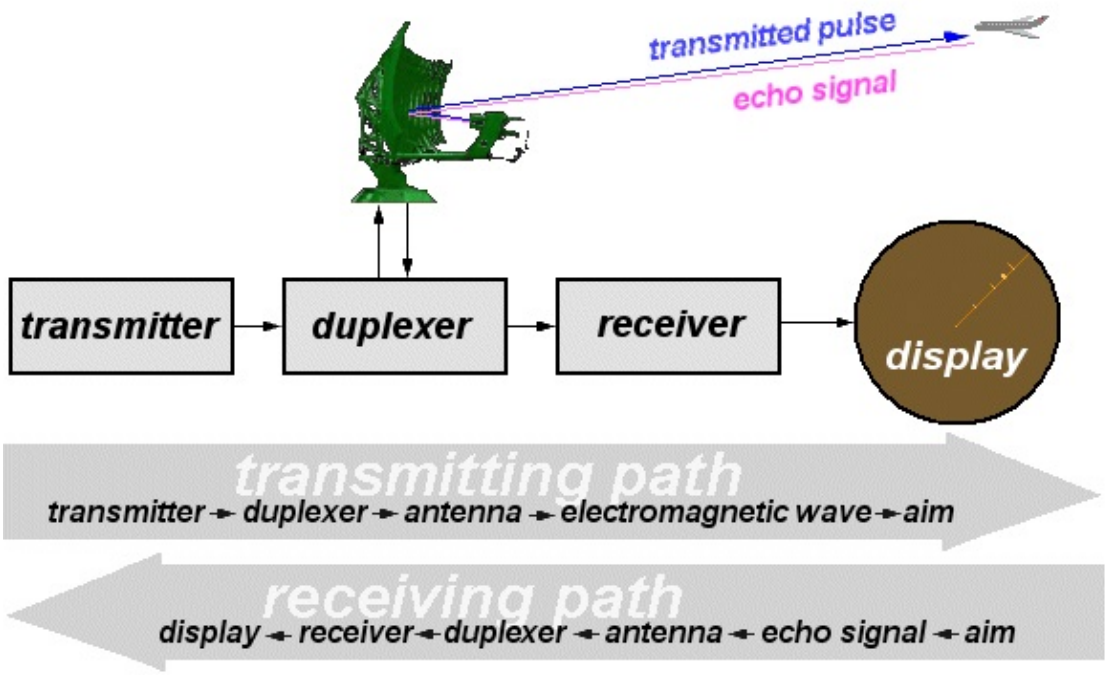


Figure 2.1: Radar principles

Based on the time between transmitting the pulse and receiving the echo, the distance to the object can be measured. The power of a reflected echo signal, P_r, can be calculated with the Radar equation [1]:

$$P_r = \frac{P_t \cdot G \cdot A_e \cdot \sigma}{(4\pi)^2 \cdot R^4}, \tag{2.1}$$

where

- P_t is the transmitted power.
- G is the Radar gain.
- A_e is the effective area.
- σ is the Radar Cross Section (RCS).
- R is the range of the object, determined by

$$R = \frac{c \cdot T_r}{2},$$

where

- c is the speed of light ($\approx 2.998 \cdot 10^8$).
- T_r is the time between transmitting and receiving the pulse.

2.1.1 Classical Radar detection

The classical detection approach consists of three stages: the detection stage, the cluster/extraction stage and the tracking stage. In the following, these stages are briefly discussed.

Detection: detection on hit level
In the first stage, the surface area of the Radar is divided in a number of cells. The reflected energy (measurement) in each cell is compared to a predetermined threshold. When the power of the measurement exceeds the threshold, a hit is declared. This process is called detection on hit level. In practice, a single scan consists of multiple bursts. Detection of the scan is performed by integration over the burst. In this thesis, this is simplified by assuming one burst per scan.

Clustering and Extraction: detection on plot level
In the second stage, all neighboring cells where a hit is declared are clustered. The power in these cells should match in range, Doppler, bearing and elevation. The clustered cells are called a plot. Detection on plot level is performed by comparing the number of clustered cells with a predetermined threshold. When the number of cells exceeds this threshold, extraction is performed by determining the center of a plot.

Tracking: detection on track level
In the third stage, a track filter is initialized at each plot. The filter prediction for successive scans is used to classify possible target plots that can be assigned to a track. After a number of scans, the number of plots assigned to a track is compared to a predetermined threshold. When the number of plots assigned to a track exceeds this threshold, a target is declared. This is called detection on track level.

The thresholds used in the three stages are usually chosen such that the probability of false declaration of a target is smaller than a desired false alarm rate. This classical approach works well when target reflections consist of high power compared to the reflections from background noise, i.e. a high Signal to Noise Ratio (SNR). However, for targets with a low SNR the probability of detection using this classical approach is low. Because the purpose of this thesis is to satisfy a constant false alarm rate detection scheme while detecting small targets, ideally in a sea clutter environment, the Track Before Detect (TBD) [2, 3] approach is used.

2.1.2 Track Before Detect

In the Track Before Detect approach, the detection decision is postponed until after the tracking stage. This allows the user to gather more information before the detection decision is made. The filters that perform the tracking are discussed in Section 2.2. Because time integration is performed over the filtered information, i.e. the information is constantly updated when new measurements are received, TBD can be seen as an integrated processing method.

The measurements that are compared to a threshold in the classical approach, serve in the TBD approach only as indication where a target might be. Because it will be computationally inefficient to start tracks all over the state space, tracks are only initialized at measurement cells where the power exceeds a startup threshold, denoted by τ_{start}. The purpose of the integrated processing is to make a detection decision only when enough information is available to make it reliable.

In the TBD approach, weaker targets will be tracked for a longer time and will have more time to produce reflections towards the Radar. As a result, the probability of detection for targets with a low SNR significantly increases when using the TBD approach compared to the classical approach. Figure 2.2 shows the difference between classical and TBD approach schematically.

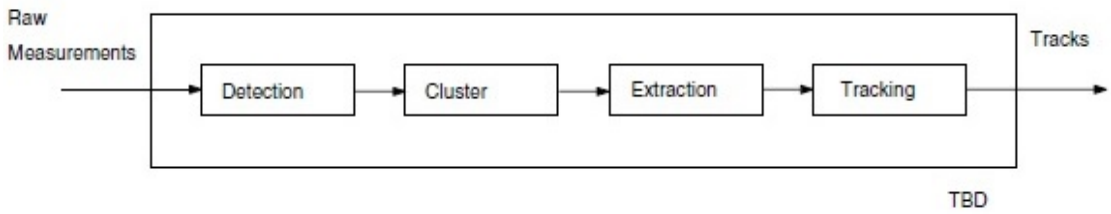


Figure 2.2: Classical data and signal processing (separate boxes) and TBD (large box)

2.2 Filtering

In this section, first the Bayes filter is discussed and later the particle filter and the Sequential Markov Chain Monte Carlo (SMCMC) filter, which are implementations of the Bayes filter based on point mass approximations, are discussed.

2.2.1 Bayes Filter

Bayesian filtering [4][5] is a probabilistic approach for estimating unknown dynamical systems recursively over time using prior knowledge and incoming evidence. Probability distributions are used to describe all relevant unknown quantities by measuring the probability of an event conditioned on the stochastic occurrence of the event. Consider the discrete-time stochastic state-space model of a dynamical system. At each time instant n, information about the system is described by state vector $x_n \in \mathbb{R}^{d_x}$. The time evolution of the state vector is described by a stochastic state space model in form of a Markov transition:

$$x_n = p_n(x_{n-1}, v_{n-1}), \qquad n \in \mathbb{N}.$$

This gives the transformation of a given state vector x_{n-1} and system noise v_{n-1} at time $n-1$ into a new state vector x_n at time n where,

- v_{n-1} is an independent identically distributed (i.i.d.) process noise vector,
- $p_n : \mathbb{R}^{d_x} \times \mathbb{R}^{d_v} \mapsto \mathbb{R}^{d_x}$ is a function (possible non-linear) of the state x_{n-1} and v_{n-1},
- d_x and d_v are the dimensions of the state space and the process noise vectors.

The system state is indirectly observed via a noisy measurement vector taking values in an observation space $\mathcal{Z}$, $z_n \in \mathbb{R}^{d_z}$. The measurement model is described by the observation equation:

$$z_n = h_n(x_n, w_n), \qquad n \in \mathbb{N}.$$

This specifies at time n the transformation of a given state vector x_n and measurement noise sample w_n into a measurement vector z_n, where

- w_n is an i.i.d. measurement noise sample,

- $h_n : \mathbb{R}^{d_x} \times \mathbb{R}^{d_w} \mapsto \mathbb{R}^{d_z}$ is a function (possible non-linear) of the state x_n and w_n,
- d_z and d_w are the dimensions of the measurement and measurement noise vectors.

From a Bayesian perspective, the tracking problem consists in inferring knowledge about the unobserved state x_n of a dynamic system, which changes over time, using a sequence of noisy measurements $z_{1:n}$. Bayesian estimation is used to recursively construct a time involving posterior distribution $p(x_n|z_{1:n})$, which describes state x_n given all measurements up to time n, $z_{1:n}$. The initial prior density of the state vector

$$p(x_0) \triangleq p_0(x_0|z_0), \qquad z_0 \in \emptyset,$$

is assumed to be known. The required pdf, $p_n(x_n|z_{1:n})$, can be recursively propagated by the Bayes recursion in two steps; the prediction step and the update step. Suppose that at time $n-1$, $p_{n-1}(x_{n-1}|z_{1:n-1})$ is available. The prediction step is performed by using the Chapman-Kolmogorov equation:

$$p(x_n|z_{1:n-1}) = \int p(x_n|x_{n-1})p(x_{n-1}|z_{1:n-1})\mathrm{d}x_{n-1}, \tag{2.2}$$

where,

- $p(x_n|x_{n-1})$ is the Markov process of order one that models the between-measurements target motion,
- $p(x_{n-1}|z_{1:n-1})$ is the posterior distribution at time $n-1$ conditioned on all measurements up to time and including $n-1$, $z_{1:n-1}$.

When at time n a new measurement z_n becomes available, the update step is performed using Bayes' rule

$$p(x_n|z_{1:n}) = \frac{p(z_n|x_n)p(x_n|z_{1:n-1})}{\int p(z_n|x_n)p(x_n|z_{1:n-1})\mathrm{d}x_n} \tag{2.3a}$$

$$= \frac{p(z_n|x_n)p(x_n|z_{1:n-1})}{p(z_n|z_{1:n-1})}, \tag{2.3b}$$

where, $p(z_n|x_n)$ is the likelihood function, which measures the adequacy of the guessed state x_n with respect to the measurement z_n, given state x_n. The Bayesian discrete-time recursive nonlinear filtering equations eqs.(2.2)-(2.3b) constitute the theoretical foundation for optimal single-sensor, single-target detection, tracking and identification. However, due to the multiple integrations in the Bayes recursion (2.2)-(2.3b), the full implementation of the Bayes filter is generally intractable in practice. Therefore, in case of strong non-linear system dynamics and non-Gaussian noises, numerical approximation performed by Sequential Monte Carlo (SMC) methods are preferred.

2.2.2 Particle Filtering

The Particle Filter [4, 6–8] performs SMC estimations based on point mass approximation of the probability densities in the Bayes filter. Assume it is possible to draw a set of i.i.d. samples, called particles, with the number of particles (NoP) $\gg 1$, according to an arbitrary probability

density $\pi(x)$, denoted by $\{x^{(j)}\}_{j=1}^{NoP}$. Then the particle filter approximates the density $\pi(x)$ by the point mass representation

$$\pi(x) \approx \frac{1}{NoP} \sum_{j=1}^{NoP} \delta_{x^j}(x), \tag{2.4}$$

where $\delta(\cdot)$ is the delta Dirac function. Following the Strong Law of Large Numbers (SLLN) [9] for i.i.d. random samples, the approximation of eq.(2.4) ensures asymptotic convergence almost surely in the sense that

$$\frac{1}{NoP} \sum_{j=1}^{NoP} p(x^j) \xrightarrow[NoP \to \infty]{a.s.} \int \pi(x) p(x) \mathrm{d}x. \tag{2.5}$$

See [10, 11] for more details about convergence results for particle filtering.
Unfortunately, it is usually impossible to sample efficiently from the posterior distribution $\pi(x)$, because $\pi(x)$ is multivariate, non standard and often only known up to a proportionally constant, i.e. $\pi(x) \propto p(x)$. In this case, an alternative solution consists of using the Bayesian Importance Sampling method. The idea is to generate samples from a known density $q(\cdot)$, referred to as the proposal density, which is close to $\pi(\cdot)$. These samples are weighted to construct a weighted point mass approximation of $\pi(\cdot)$. Let $\{x^j, w^j\}_{j=1}^{NoP}$ be the set of particles with corresponding weights that approximated the posterior density $\pi(x)$ by a weighted point mass representation:

$$\pi(x) \approx \sum_{j=1}^{NoP} w(x^j) \delta_{x^j}(x),$$

where the weights

$$\tilde{w}(x^j) = \frac{p(x^j)}{q(x^j)}, \qquad w(x^j) = \frac{\tilde{w}(x^j)}{\sum_{j=1}^{NoP} \tilde{w}(x^j)},$$

are the importance weights and the normalized importance weights. In the sequential case, the approximation of the posterior distribution $\pi(x)$ is calculated iteratively. Suppose at time $n-1$, the posterior distribution $p_{n-1}(x_{n-1}|z_{1:n-1})$ is approximated by a set of particles $\{x_{n-1}^j, w_{n-1}^j\}_{j=1}^{NoP}$:

$$p_{n-1}(x_{n-1}|z_{1:n-1}) \approx \sum_{j=1}^{NoP} w_{n-1}^j \delta_{x_{n-1}^j}(x_{n-1}).$$

Then for a given proposal density

$$x_n^j \sim q_n(x_n|x_{n-1}^j, z_n)$$

and the weight update is given by

$$w_n^j = \frac{\tilde{w}_n^j}{\sum_{j=1}^{NoP} \tilde{w}_n^j},$$

where

$$\tilde{w}_n^j = w_{n-1}^j \frac{p_n(z_n|x_n^j) p_{n|n-1}(x_n^j|x_{n-1}^j)}{q_n(x_n^j|x_{n-1}^j, z_n)}$$

such that the new set of weighted particles $\{x_n^j, w_n^j\}_{j=1}^{NoP}$ approximates $p_n(x_n|z_{1:n})$:

$$p_n(x_n|z_{1:n}) \approx \sum_{j=1}^{NoP} w_n^j \delta_{x_n^j}(x_n).$$

After a certain number of iterations, the weights of many particles will be negligible. This is called the Degeneracy problem. The degeneracy implies that a lot of computational time is lost (inefficient), because the contribution of many particles to the approximate of $\pi(x)$ is zero. Therefore, in Sampling Importance Re-sampling (SIR) filters a re-sample step is introduced. In [12] an effective number of particles is defined as:

$$N_{eff} = \frac{NoP}{1 + \mathrm{Var}[\tilde{w}_n^j]},$$

where

$$\tilde{w}_n^j = \frac{p_n(x_n^j|z_{1:n})}{q_n(x_n^j|x_{n-1}^j, z_n)}.$$

Because this expression can not be calculated exactly, an approximation is given by

$$\hat{N}_{eff} = \frac{1}{\sum_{j=1}^{Nop}(w_n^j)^2}.$$

In SIR filters, the re-sample step is performed if $\hat{N}_{eff} \leq N_{thr}$, where N_{thr} is a predetermined value. In Algorithm 1, the SIR particle filter [6, 8] is summarized.

Algorithm 1: SIR particle filter

Input :$\{x_{n-1}^j, w_{n-1}^j\}_{j=1}^{NoP}$, new measurement z_n
Output :$\{x_n^j, w_n^j\}_{j=1}^{NoP}$

Generate initial particles $\{x_0^j\}_{j=1}^{NoP}$ from $p_0(x)$;
Set weights w_0^j at $\frac{1}{NoP}$;
1 - Prediction;
for *j:=1* ***to*** *NoP* **do**
  Draw particle $\tilde{x}_n^j \sim q_n(x_n|x_{n-1}^j, z_n)$;
end
2 - Update;
for *j:=1* ***to*** *NoP* **do**
  Compute weights: $\tilde{w}_n^j \propto w_{n-1}^j \frac{p_n(z_n|\tilde{x}_n^j)p_{n|n-1}(\tilde{x}_n^j|x_{n-1}^j)}{q_n(\tilde{x}_n^j|x_{n-1}^j, z_n)}$;
end
Normalize weights : $w_n^j = \frac{\tilde{w}_n^j}{\sum_{j=1}^{NoP}\tilde{w}_n^j}$;
3 - Compute particle degeneracy: $\hat{N}_{eff} = \frac{1}{\sum_{j=1}^{Nop}(w_n^j)^2}$;
4 - Re-sample;
if $\hat{N}_{eff} \leq N_{thr}$ **then**
  Generate a new set of particles $\{x_n^j\}_{j=1}^{NoP}$, from the set $\{\tilde{x}_n^j\}_{j=1}^{NoP}$ according to normalized importance weights w_n^j;
  Set weights w_n^j at $\frac{1}{NoP}$;
else
  New set particles $\{x_{n-1}^j, w_{n-1}^j\}_{j=1}^{NoP} = \{\tilde{x}_{n-1}^j, w_{n-1}^j\}_{j=1}^{NoP}$;
end

2.2.3 Sequential Markov Chain Monte Carlo

When the dimension of the state space grows and multiple targets are present, the performance of SIR filters decreases. The main problem with these filters is the need to perform importance sampling in very high dimensional spaces if many targets are present. First, if too few particles are used, all but a few importance weights will be near zero. Re-sampling will then lead to a loss of diversity among the particles, known as sample impoverishment problem. Second, it can be difficult to find an efficient proposal distribution. The re-sample step causes a loss of diversity in high dimensional state spaces and also finding a good proposal distribution is difficult. Therefore, although the dimension of the state space is low and only one possible target is considered, a Sequential Markov Chain Monte Carlo (SMCMC) filter [13, 14] is used in this thesis. In an SMCMC filter, the inefficient importance sampling step of the standard SIR particle filter implementation is replaced by an efficient MCMC sampling step.
Within the Bayesian framework, our aim is to compute the filtering pdf $p_n(x_n|z_{1:n})$, recursively by

$$p_{n|n-1}(x_n|z_{1:n-1}) = \int p_{n|n-1}(x_n|x_{n-1})p_{n-1}(x_{n-1}|z_{1:n-1})\mathrm{d}x_{n-1}. \tag{2.6}$$

and

$$p_n(x_n|z_{1:n}) \propto p_n(z_n|x_n)p_{n|n-1}(x_n|z_{1:n-1}). \tag{2.7}$$

Let us define the set of unweighted particles $\{x^j_{n-1}\}_{j=1}^{NoP}$ that approximates the posterior distribution $p_{n-1}(x_{n-1}|z_{1:n-1})$ by the point mass, i.e.

$$\frac{1}{NoP}\sum_{j=1}^{NoP}\delta_{x^j_{n-1}}(x_{n-1}) \approx p_{n-1}(x_{n-1}|z_{1:n-1}). \tag{2.8}$$

By substituting eq.(2.6) and eq.(2.8) in eq.(2.7), $p_n(x_n|z_{1:n})$ can be written as

$$p_n(x_n|z_{1:n}) \approx \frac{1}{NoP}p_n(z_n|x_n)\sum_{j=1}^{NoP}p_{n|n-1}(x_n|x^j_{n-1}). \tag{2.9}$$

Once the filtered distribution is approximated by eq.(2.9), the MCMC sampling step is performed to obtain a better approximation of $p_n(x_n|z_{1:n})$. The set of particles with corresponding likelihoods that serve as input of the MCMC step, is considered the predicted posterior set with corresponding predicted posterior likelihoods. For simplicity, the set of predicted posterior particles with corresponding likelihoods will be referred to as the set of predicted particles in the thesis, denoted by $\{x^j_{pred}\}_{j=1}^{NoP}$, with corresponding predicted likelihoods, denoted by $p(z|x^j_{pred})$. The output of the MCMC will be referred to as the set of posterior particles, denoted by $\{x^j_{post}\}_{j=1}^{NoP}$, with corresponding posterior likelihoods $p(z|x^j_{post})$. A summary of the SMCMC is given in Algorithm 2.

Remark 1. It is important not to confuse the predicted set of particles in this thesis (in a SMCMC sense) with the set of predicted particles in a SIR [6, 8]. The important difference is that the set of predicted particles with likelihoods at time n, in the way it is defined in this thesis, contains the measurements up to and including time n, where the set of predicted particles at time n in a SIR only contains the measurements up to time and including time $n-1$.

Algorithm 2: SMCMC

Input :$\{x_{n-1}^j\}_{j=1}^{NoP}$, new measurement z_n
Output :$\{x_n^j\}_{j=1}^{NoP}$

Generate initial particles $\{x_0^j\}_{j=1}^{NoP}$ from $p_0(x)$;
1 - Prediction;
for *j:=1* ***to*** *NoP* **do**
  Draw particle $x_{n_{pred}}^j \sim q_n(x_n|x_{n-1}^j, z_n)$;
end
2 - Update;
for *j:=1* ***to*** *NoP* **do**
  Compute predicted likelihood ratios $\lambda(x_{n_{pred}}^j) = \frac{p(z_n|x_{n_{pred}}^j)}{p(z_n)}$;
end
3 - Perform MCMC step: Do MH algorithm; Obtain set of posterior particles $\{x_{n_{post}}^j\}_{j=1}^{NoP}$ with posterior likelihood ratios $\lambda(x_{n_{post}}^j)$;
4 - Copy set of posterior particles: $\{x_n^j\}_{j=1}^{NoP} = \{x_{n_{post}}^j\}_{j=1}^{NoP}$;

Metropolis-Hastings algorithm

The Metropolis-Hastings (MH) algorithm is used to perform the MCMC step. The MH algorithm sets up a Markov chain which must be aperiodic, irreducible and positive recurrent in order to be ergodic [15, 16], such that it will eventually reach all points of the space and that the time average equals the space average. Hence, the use of the MH algorithm allows us to simulate a Markov chain such that its stationary distribution is in fact the desired target distribution. Hence, the samples obtained from the Markov chain are equivalent to correlated samples drawn from the target distribution. This can be implemented, even if the normalization constant is unknown. The algorithm works as follows:

1. Given state x, draw proposed state x^* from transition density $q(x^*|x)$.

2. Calculate acceptance probability

$$\alpha(x^*|x) = \min\left(1, \frac{p(x^*|z)q(x|x^*)}{p(x|z)q(x^*|x)}\right) \tag{2.10}$$

$$= \min\left(1, \frac{p(z|x^*)p(x^*)q(x|x^*)}{p(z|x)p(x)q(x^*|x)}\right). \tag{2.11}$$

When the proposal distribution is chosen to be the predicted distribution, i.e. $q(x^*|x) \triangleq p(x^*)$, eq.(2.11) can be rewritten as

$$\alpha(x^*|x) = \min\left(1, \frac{p(z|x^*)}{p(z|x)}\right). \tag{2.12}$$

3. Draw $u \sim \mathcal{U}[0,1]$

4. Let

$$x = \begin{cases} x^*, & \alpha(x^*|x) \geq u \\ x, & \alpha(x^*|x) < u \end{cases}$$

The MH algorithm is designed in a sequential setting in order to approximate the filtering distribution eq.(2.7). This is achieved by using a conditional density $q(x^*|x)$, also known as proposal distribution, to generate a Markov chain with invariant distribution the approximate posterior eq(2.9). Then to compensate a probability of move and to satisfy reversibility condition, an acceptance $\alpha(x^*|x)$ is introduced. The MH algorithm is discussed further in Section 2.3.5. See [14, 17, 18] for complete details. The MH algorithm is summarized in Algorithm 3.

Algorithm 3: Metropolis-Hastings algorithm

Input :x_0, number of particles NoP, burn in period B
Output :$\{x^j\}_{j=1}^{NoP}$

for *j:=1* ***to*** *NoP+B* **do**
  Draw state $x^* \sim q(x^*|x^{j-1})$;
  Calculate acceptance ratio: $\alpha(x^*|x^{j-1}) = \min\left(1, \frac{p(z|x^*)}{p(z|x^{j-1})}\right)$;
  Draw $u \sim \mathcal{U}[0,1]$;
  if $\alpha \geq u$ **then**
    Accept x^*: $x^j = x^*$;
  else
    Reject x^*: $x^j = x^{j-1}$;
  end
end
Discard first B samples, store other NoP samples.

2.3 Important specifications of the filter

In this section, a few important properties of the SMCMC are discussed. These properties will be important during the rest of this thesis. First, the background models are discussed. Second, the target model. Then the combined target+background model is described, followed by the measurement model. Finally, a more detailed description of the MH algorithm in the SMCMC is given.

2.3.1 Background models

When a Radar antenna transmits Rf pulses, besides a target object, also background environment reflects echo signals back to the receiver. Hence not only targets echo samples, but also background samples (also referred to as background scatter) will appear on the video display.

For good performance of the filter, a good background model $p_0(x)$ is needed. A reflected signal is converted into two electrical components, a real valued component I and an imaginary valued component Q, given by

$$I = A\cos(\theta),$$

$$Q = A\sin(\theta),$$

where

- A is the amplitude of the complex signal.
- θ is the phase angle of the complex signal.

The power of the signal is the amplitude squared:

$$P = A^2.$$

When a Radar is transmitting pulses towards, for example, a peaceful sky, the background noise is often modeled as Gaussian noise. When the background is modeled by Gaussian noise, it means that the echo signals are Gaussian distributed in the I/Q domain, i.e.

$$z_I \sim \mathcal{N}(0, \sigma_n^2), \quad p(z_I) = \frac{1}{\sigma_n\sqrt{2\pi}}\mathrm{e}^{-z_I^2/2\sigma_n^2}, \qquad z_I \in \mathbb{R}, \tag{2.13a}$$

$$z_Q \sim \mathcal{N}(0, \sigma_n^2), \quad p(z_Q) = \frac{1}{\sigma_n\sqrt{2\pi}}\mathrm{e}^{-z_Q^2/2\sigma_n^2}, \quad z_Q \in \mathbb{R}. \tag{2.13b}$$

In Appendix A, the calculations for going from the I/Q domain to power domain and vice versa are performed. The amplitude distribution is given by

$$p(z_A) = \frac{z_A}{\sigma_n^2}\mathrm{e}^{-z_A^2/2\sigma_n^2}, \qquad z_A \in \mathbb{R}^+, \tag{2.14}$$

also known as the Rayleigh distribution. The power distribution is

$$p(z_P) = \frac{1}{2\sigma_n^2}\mathrm{e}^{-z_P/2\sigma_n^2}, \qquad z_P \in \mathbb{R}^+, \tag{2.15}$$

which is an exponential distribution with mean $2\sigma_n^2$. When a Radar is transmitting towards a rough sea, the exponential power distribution is not an accurate model. In this thesis, the power samples from sea (sea clutter) are modeled by a generalized Pareto distribution, i.e.

$$p(z_P) = \frac{1}{s}\left(1 + k\frac{z_P - \theta}{s}\right)^{-1-1/k}, \qquad z_P \in \mathbb{R}^+, \tag{2.16}$$

with

- shape parameter $k > 0$,
- scale parameter $s > 0$,
- location parameter θ, assumed to be 0,
- measurement $z_P > \theta = 0$.

The distribution of the amplitude of a clutter sample is given by (see Appendix A.2)

$$p(z_A) = \frac{2z_A}{s}\left(1 + k\frac{z_A^2}{s}\right)^{-1-1/k}, \quad z_A \geq 0, k > 0, s > 0 \tag{2.17}$$

and the distributions for the I and Q samples are (see Appendix A.2)

$$p(z_I) = \frac{1}{\sqrt{\pi s}}\left(1 + k\frac{z_I^2}{s}\right)^{-1-1/k}, \quad z_I^2 \geq 0, k > 0, s > 0, \tag{2.18a}$$

$$p(z_Q) = \frac{1}{\sqrt{\pi s}}\left(1 + k\frac{z_Q^2}{s}\right)^{-1-1/k}, \quad z_Q^2 \geq 0, k > 0, s > 0. \tag{2.18b}$$

2.3.2 Target model

Tracked targets do not always reflect an echo signal of the same power. It depends on the surface where transmitted pulses hit, the neighborhood of the target and on many more variables. Small targets may not even reflect an echo signal back, for example when a small sea mine is positioned behind a wave. Swerling models [19–21] describe the target reflection characteristics of the RCS, denoted by σ, of a target for different scenarios, varying from Swerling I, to a Swerling V model (also referred to as Swerling 0). According to the Swerling model, the RCS fluctuation of a target is modeled by a chi-square distribution with m degrees of freedom, i.e. $p(\sigma) \sim \chi^2_{2m}(\sigma)$:

$$p(\sigma) = \frac{m}{\Gamma(m)\sigma_{avg}}\left(\frac{m\sigma}{\sigma_{avg}}\right)^{m-1} \mathrm{e}^{-m\sigma/\sigma_{avg}}, \quad \sigma \in \mathbb{R}^+, \tag{2.19}$$

where

- σ_{avg} is the mean value of σ,
- $m = 1$ represents Swerling I and Swerling II targets where $p(\sigma)$ reduces to an exponential distribution,
- $m = 2$ represents Swerling III and Swerling IV targets,
- $m \to \infty$ represents a Swerling 0 target.

In this thesis, a Swerling I model is used. It means that the backscattered signal varies independent from scan to scan and it varies according to an exponential distribution. Practically, it means that sometimes the target is visible on the video and sometimes it is not. More difficult models embrace variations from burst to burst (Swerling II and IV) or embrace larger variation by using four degrees of freedom in eq. (2.19) (Swerling III and IV). A Swerling 0 will be discussed at some point in this thesis as well. The Swerling 0 target has a constant RCS and is used as reference to compare the probability of detection between a Swerling I and a Swerling 0 target.
Because the Swerling I model indicates that the target does not reflects an echo signal at any time, the I/Q samples of the target are modeled by a normal distribution with zero mean. This means that the target samples are similarly distributed as the Gaussian noise samples. i.e.

$$p(z_I) = \frac{1}{\sigma_t\sqrt{2\pi}}\mathrm{e}^{-z_I^2/2\sigma_t^2}, \quad z_I \in \mathbb{R} \tag{2.20a}$$

$$p(z_Q) = \frac{1}{\sigma_t\sqrt{2\pi}}\mathrm{e}^{-z_Q^2/2\sigma_t^2}, \quad z_Q \in \mathbb{R} \tag{2.20b}$$

$$p(z_A) = \frac{z_A}{\sigma_t^2}\mathrm{e}^{-z_A^2/2\sigma_t^2}, \qquad z_A \in \mathbb{R}^+ \tag{2.20c}$$

$$p(z_P) = \frac{1}{2\sigma_t^2}\mathrm{e}^{-z_P/2\sigma_t^2}, \qquad z_P \in \mathbb{R}^+, \tag{2.20d}$$

where σ_t is the standard deviation of the target I/Q samples. From Appendix A, the power distribution can be rewritten as

$$p(z_P) = \frac{1}{\mu_t}\mathrm{e}^{-z_P/\mu_t}, \quad z_P \in \mathbb{R}^+,$$

with $\mu_t = 2\sigma_t^2$.

Within the particle filter, the time evolution of a possible target through the state space is predicted according to a dynamic model with a certain error. This model is iteratively updated when new information from measurements are available.

2.3.3 Target + background model

In Sections 2.3.1 and 2.3.2, the distribution of noise/clutter and target samples are given in three different domains. Since a reflected target echo signal will partly consist of background echo as well, the distribution of the sum of target and background samples is needed. In this section, these distributions are given with a brief motivation. For further details and proofs, let us refer to Appendix B.

Target + noise

The target and noise distributions for the three domains are given in eqs.(2.13)-(2.15) and eq.(2.20). Physically, target and noise samples are added in the I/Q domain. Hence to determine the distribution of the sum of target and noise samples, the distribution of eqs.(2.13a) and (2.20a) and eqs.(2.13b) and (2.20b) are needed. Since the I samples are independently normally distributed r.v.'s, their sum is normally distributed with the mean being the sum of the

means and the variance being the sum of the two variances (see Appendix B.1), i.e.

$$p(z_I) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-z_I^2/2\sigma^2}, \quad z_I \in \mathbb{R}, \tag{2.21}$$

where $\sigma^2 = \sigma_t^2 + \sigma_n^2$. The I and Q samples are identically distributed, so

$$p(z_Q) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-z_Q^2/2\sigma^2}, \quad z_Q \in \mathbb{R}, \tag{2.22}$$

Given the calculations in Appendix A, the amplitude and power distributions for the sum of target and noise samples are

$$p(z_A) = \frac{z_A}{\sigma^2} \mathrm{e}^{-z_A^2/2\sigma^2}, \quad z_A \in \mathbb{R}^+ \tag{2.23a}$$

$$p(z_P) = \frac{1}{2\sigma^2} \mathrm{e}^{-z_P/2\sigma^2}, \quad z_P \in \mathbb{R}^+, \tag{2.23b}$$

From eq.(2.23b), the power distribution for target+noise is an exponential distribution with mean μ, i.e.

$$p(x) = \frac{1}{\mu} \mathrm{e}^{-x/\mu}, \quad x \in \mathbb{R}^+, \tag{2.24}$$

with

$$\mu = 2\sigma^2 = 2\sigma_n^2 + 2\sigma_t^2 = \mu_n + \mu_t.$$

Target + clutter (analytical)

Now, some calculations about the distribution of target + clutter are discussed. Contrary to the distributions of target+noise, the distribution of target+clutter is very difficult to find. Let us refer to Appendix B.2.1 for more details and calculations. Similar to the case of target+noise, the distribution of the I samples is used, given by eqs. (2.18a) and (2.20a). It seems hopeless to calculate the convolution of eqs.(2.18a) and (2.20a) directly. Therefore the characteristic functions are introduced. Unfortunately, the characteristic function of eq.(2.18a) is not helpful either (see Appendix B.2.1). Therefore, instead of calculating the sum of the I and Q samples, a transformation to the complex signal is proposed:

$$\bar{z} = x + iy, \tag{2.25}$$

where

- $x = z_I$,
- $y = z_Q$.

This is proper because the I and Q samples are uncorrelated. First the distribution for the target complex signal is given. Using the transform of eq.(2.25) in the joint distribution function of the I and Q samples

$$p(x, y) = \frac{1}{2\pi\sigma_t^2} \mathrm{e}^{-(x^2+y^2)/(2\sigma_t^2)}, \quad x, y \in \mathbb{R}$$

transforms into

$$p(\bar{z}) = \frac{1}{2\pi\sigma_t^2} \mathrm{e}^{-|\bar{z}|^2/(2\sigma_t^2)}, \quad \bar{z} \in \mathbb{C} \tag{2.26}$$

Likewise, using the transform of eq.(2.25), the joint distribution of clutter I and Q samples

$$p(x,y) = \frac{1}{2\pi s^2}\left(1 + k\frac{x^2+y^2}{2s^2}\right)^{-1-1/k}, \quad x,y \in \mathbb{R}$$

transforms into

$$p(\bar{z}) = \frac{1}{2\pi s^2}\left(1 + k\frac{|\bar{z}|^2}{2s^2}\right)^{-1-1/k}, \quad \bar{z} \in \mathbb{C}. \tag{2.27}$$

Now, to find the characteristic function of eqs.(2.26) and (2.27) a transformation to polar coordinates is used. The characteristic function is given by

$$\varphi_{\bar{z}}(t) = \mathbb{E}[\mathrm{e}^{it\bar{z}}] = \int_{-\infty}^{\infty} \mathrm{e}^{it\bar{z}} p(\bar{z})\mathrm{d}\bar{z} = \int_0^{2\pi}\int_0^r \mathrm{e}^{itr} p(r,\theta)\mathrm{d}r\mathrm{d}\theta, \tag{2.28}$$

where $p(r,\theta)$ for target and clutter are given by respectively

$$p(r,\theta) = \frac{r}{2\pi\sigma_t^2}\mathrm{e}^{-r^2/(2\sigma_t^2)}, \qquad r \in \mathbb{R}^+, \theta \in [0,2\pi], \tag{2.29a}$$

$$p(r,\theta) = \frac{r}{2\pi s^2}\left(1 + k\frac{r^2}{2s^2}\right)^{-1-1/k}, \quad r \in \mathbb{R}^+, \theta \in [0,2\pi]. \tag{2.29b}$$

Unfortunately, calculations of $\varphi_{\bar{z}}(t)$ with $p(r,\theta)$ from eq.(2.29b) does not give a useful expression either (see Appendix B.2.1). Therefore, instead of calculating the distribution of target+clutter analytically, parameter estimation is performed to approximate this density.

Target + clutter (parameter estimation)

The parameter estimation is based on performed numerical approximations. For details, let us refer to Appendix B.2.2.
Let the joint distributions of the I and Q samples for target and clutter be given by

$$p(z_I, z_Q) = \frac{1}{2\pi\sigma_t^2}\mathrm{e}^{-(z_I^2+z_Q^2)/(2\sigma_t^2)}, \qquad z_I, z_Q \in \mathbb{R} \tag{2.30a}$$

$$p(z_I, z_Q) = \frac{1}{2\pi s^2}\left(1 + k\frac{z_I^2+z_Q^2}{2s^2}\right)^{-1-1/k}, \quad z_I, z_Q \in \mathbb{R}. \tag{2.30b}$$

Based on numerical evaluation, the convolution of a generalized Pareto distributed r.v. and a normal distributed r.v. seems to be again (generalized) Pareto distributed. The questions remains, what the shape parameter k and the scale parameter s of this new distribution are (θ is still assumed to be 0). The proposed distribution joint distribution of I and Q samples for target+clutter is given by

$$p(z_I, z_Q) = \frac{1}{\pi(\mu_t+\mu_c)}\left(1 + \tilde{k}\frac{z_I^2+z_Q^2}{\mu_t+\mu_c}\right)^{-1-1/\tilde{k}}, \quad z_I, z_Q \in \mathbb{R}, \tag{2.31}$$

where

- $\tilde{k} = k\frac{\mu_c}{\mu_c+\mu_t}$,

- μ_t is the mean of the target power distribution, equal to $2\sigma_t^2$,
- μ_c is the mean of the clutter power distribution, given by $\frac{s}{1-k}$.

By using, the transformations in Appendix A.2, eq.(2.31) yields to

$$p(z_P) = \frac{1}{\tilde{s}} \left(1 + \tilde{k}\frac{z_P}{\tilde{s}}\right)^{-1-1/\tilde{k}}, \quad z_P \in \mathbb{R}^+ \tag{2.32}$$

where

- $\tilde{s} = s + \mu_t$,
- $\tilde{k} = \frac{k \cdot s}{\tilde{s}} = \frac{k \cdot s}{s+\mu_t}$.

The figures in Appendix B.2.2 show the accuracy of the new parameters in eq.(2.32).

2.3.4 Measurement model

In this section, the measurement model is discussed. First the impulse response (point spread function) from a reflecting object is discussed. Second, the processing of the point spread function in the likelihood function is explained.

Point spread function

Transmitted pulses are very short but have extremely high power. They can be described mathematically by a delta pulse. A reflected echo signal has a different character. The point spread function (PSF), in more general terms the impulse response, is a way to describe the response of an imaging system to a point source. Instead of a delta pulse, the PSF can be seen as a blurred signal. Physically, it comes from the fact that a signal from a point source is spread out such that, when it is received by an antenna, it can be measured in multiple cells. The higher the spread (beam width), the vaguer the image of the object.
To model the impulse response of a point source, the point spread function is used to predict particles in the measurement model. The PSF is assigned to the particles around their predicted position. In general, a -3dB beam width is used, which means a beam width, such that the amplitude of the beam is decreased by 50%, i.e. power by 75%, when it is absorbed by the Radar. To avoid bias, the point spread function should be a density (integrating to one), to ensure proper results. For computational efficiency, the PSF is often discretized over the measurement cells. Hence the cells around the particles will get a (point spread) value, where the values sum up to 1.

Likelihood function

The blurred signals measured by the radar are compared to the PSF around the particles. In 2D, the resulting measurement of the particle is given by the inner product of the received signal vector and the predicted point spread vector. The resulting measurement will be used to calculated the likelihood function, $p(z|x^j)$, of the particles.

2.3.5 Metropolis-Hastings algorithm

In Section 2.2.3, the MH algorithm has been introduced. The algorithm is used in a sequential setting in order to approximate the posterior distribution eq.(2.9). Within the hypothesis testing,

thoroughly discussed in Chapter 3, eq.(2.9) approximates the target (+background) distribution, denoted by $p_1(x)$. This distribution is compared to a background hypothesis with a background distribution, denoted by $p_0(x)$. This distribution is independent of the particle cloud. The likelihood functions of $p_1(x)$ and $p_0(x)$, denoted by $p_1(z|x)$ and $p_0(z|x)$ are used to construct a likelihood ratio, denoted by $\lambda(x)$ in the following way

$$\lambda(x) = \frac{p_1(z|x)}{p_0(z|x)}.$$

This likelihood ratio is used to assign a track to a target or to noise. Without loss of generality the scaling by $p_0(x)$ can be included in the MH algorithm, because $p_0(x)$ does not depend on the particle cloud. Hence, instead of using the likelihood function $p_1(x)$, the likelihood ratio $\lambda(x)$ is used in the MH algorithm. Therefore eq.(2.11) at time n, becomes

$$\alpha(x_n^*|x_n) = \min\left(1, \frac{\lambda(x_n^*)p_n(x_n^*)q_n(x_n|x_n^*)}{\lambda(x_n)p_n(x_n)q_n(x_n^*|x_n)}\right). \tag{2.33}$$

Again, when the proposal distribution is chosen to be the predicted distribution, i.e. $q_n(x_n^*|x_n) = p_n(x_n^*)$, where $p_n(x_n^*)$ in our case is given by $p_n(x_n^*) = p_{n|n-1}(x_n^*|x_{n-1})p_{n-1}(x_{n-1})$, eq.(2.33) can be rewritten as

$$\alpha(x_n^*|x_n) = \min\left(1, \frac{\lambda(x_n^*)}{\lambda(x_n)}\right). \tag{2.34}$$

The MH algorithm as used in the SMCMC is summarized in Algorithm 4.

Algorithm 4: Metropolis-Hastings algorithm in SMCMC

Input :$\{x_{pred}^j; \lambda(x_{pred}^j)\}_{j=1}^{NoP}$, number of particles NoP, burn in period B
Output :$\{x_{post}^m, \lambda(x_{post}^m)\}_{m=1}^{NoP}$

```
for m:=1 to NoP+B do
    if m = 1 then
        Randomly select particle x^j_pred out {x^j_pred}_{j=1}^{NoP};
        Store x^j_pred: x^m_post = x^j_pred,   λ(x^m_post) = λ(x^j_pred);
    else
        Draw proposed particle x^{j*}_pred ~ q(x^{j*}|x^{j-1}), with associate likelihood ratio
        λ(x^{j*}_pred);
        Calculate acceptance ratio: α(x*|x) = min(1, λ(x^{j*}_pred)/λ(x^{m-1}_post));
        Draw u ~ U[0,1];
        if α ≥ u then
            Accept x^{j*}: x^m_post = x^{j*},   λ(x^m_post) = λ(x^{j*}_pred);
        else
            Reject x^{j*}: x^m_post = x^{m-1}_post,   λ(x^m_post) = λ(x^{m-1}_post);
        end
    end
end
Discard first B samples, store other NoP samples: {x^m_post}_{m=B+1}^{NoP+B} → {x^m_post}_{m=1}^{NoP}
```

Chapter 3

Hypothesis Testing

In Chapter 2, the basic principles of Radar detection and filtering are explained. Particle filtering is used to perform the tracking of possible targets. Based on the information from the particle filter, a decision is made about the presence of a target. In this chapter, the theory behind hypothesis testing is explained. Let us first consider simple hypotheses of the form

$$\mathcal{H}_i : \quad x_1, x_2, \ldots, x_N \overset{iid}{\sim}, \quad i = 0, 1. \tag{3.1}$$

$\mathcal{H}_0$ corresponds to the hypothesis "Target is not present" and is called the null hypothesis. $\mathcal{H}_1$ is called the alternative hypothesis and refers to the hypothesis "Target is present". Two possible errors can be made while testing the null hypothesis $\mathcal{H}_0$ against the alternative hypothesis $\mathcal{H}_1$. A false-positive or false alarm occurs when concluding $\mathcal{H}_1$ while in fact $\mathcal{H}_0$ is true. The probability of false alarm (P_{FA}) is given by

$$P_{FA} = P(\text{Conclude } \mathcal{H}_1 | \mathcal{H}_0) = \alpha,$$

where α is called the significance level of the test. A false-negative or miss detection occurs when concluding $\mathcal{H}_0$ while in fact $\mathcal{H}_1$ is true. The probability of miss detection (P_{MD}) is given by

$$P_{MD} = P(\text{Conclude } \mathcal{H}_0 | \mathcal{H}_1) = \beta.$$

A correct detection is made when the test decides $\mathcal{H}_1$ when $\mathcal{H}_1$ is true. The probability of correct detection (P_D) is given by

$$P_D = P(\text{Conclude } \mathcal{H}_1 | \mathcal{H}_1) = 1 - \beta,$$

where $1 - \beta$ is called the power of the test and where the last equality holds if always a decision is made at a certain point. The error probabilities decrease as N (the number of i.i.d. observations) increases, and the minimum number N needed to achieve desired levels of error can be characterized. Rather than fixing N ahead of time, it is natural to consider a sequential approach to testing which continues to gather samples until a confident decision can be made. This idea is usually referred to as a sequential likelihood ratio test (SLRT). In [22] the Receiver Operating Characteristic (ROC) curves of the SLRT show a better performance than the ROC curves of the likelihood ratio test (LRT) in terms of P_D against P_{FA} at an arbitrary (finite) time. Hence following [22], the SLRT is a more powerful test than the LRT. Therefore, a sequential likelihood ratio test will be used in the detection scheme in this thesis. Furthermore, the Neyman-Pearson lemma [23], which gives the most powerful test for a fixed sample LRT, is applied and integrated in a sequential detection scheme.

3.1 Sequential Likelihood Ratio Test

Again consider the hypothesis (3.1). The SLRT is based on considering the likelihood ratio as a function of the number of observations. Define

$$\lambda(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \frac{p_1(x_i)}{p_0(x_i)}, \quad n = 1, 2, \ldots$$

To simplify the notation, let $\mathbf{x}_n := (x_1, x_2, \ldots, x_n)$. The goal of the SLRT is to decide which hypothesis is correct as soon as possible (i.e., for the smallest value of N). To do this the SLRT requires two thresholds, an upper threshold τ_u and a lower threshold τ_l, with

$$\tau_u > \tau_l.$$

The SLRT is defined as:

Definition 2 (Sequential likelihood ratio test)**.**

1. If $\lambda(\mathbf{x}_n) \geq \tau_u$, choose $\mathcal{H}_1$, else go to 2.
2. If $\lambda(\mathbf{x}_n) \leq \tau_l$, choose $\mathcal{H}_0$, else n=n+1 go to 1.

Let us define S as the stopping (decision) time of a test. Assume that the SLRT will eventually decide (stop), i.e.

$$P_i(S < \infty) = 1, \quad i = 0, 1.$$

At time n, P_{FA} and P_{MD} are given by

$$P_{FA}^n = P(\text{Conclude } \mathcal{H}_1 | \mathcal{H}_0) = P(\lambda(\mathbf{x}_n) \geq \tau_u | \mathcal{H}_0) = \alpha,$$

$$P_{MD}^n = P(\text{Conclude } \mathcal{H}_0 | \mathcal{H}_1) = P(\lambda(\mathbf{x}_n) \leq \tau_l | \mathcal{H}_1) = \beta.$$

The challenge is to set the thresholds so that a certain level of errors is guaranteed. Making τ_u larger and τ_l smaller yields a test that will tend to stop later and produce more accurate decisions. On the other hand, making τ_u smaller and τ_l larger yields to a test that makes a quicker decision, but yields to a test which produces more errors. Hence, the values for τ_u and τ_l are crucial for the P_{FA} and P_D. Section 3.1.1 discusses some theorems and results from Wald's theory [24]. Section 3.1.2 discusses Neyman-Pearson theory [23], and highlights important difference with respect to Wald's theory. Finally, both theories are applied to define a SLRT and compared, first in Section 3.2 with i.i.d. samples and second, in Section 3.3, with samples from a particle filter.

3.1.1 Wald's theory

In [24], expressions for the upper and lower thresholds, τ_u and τ_l, are given as well as an expression for the expected stopping time denoted by $\mathbb{E}[S]$. In what follows, expressions for τ_u and τ_l are derived differently using decision sets/critical regions of the test. These expressions are derived with the aim to provide a test that satisfies desired probabilities of detection, P_D, and false alarm, P_{FA}. Also an expression for $\mathbb{E}[S]$ is derived using Wald's equation [25][26] and the Kullback-Leibler Divergence [27]. To simplify the notation, let $\mathbf{x}_S := (x_1, \ldots, x_S)$ and write $p_j(x_S) := \prod_{i=1}^{S} p_j(x_i), j = 0, 1$. At time S, P_D can be written in terms of the critical region $C_1 := \{\mathbf{x}_S : \lambda(\mathbf{x}_S) \geq \tau_l\}$ as follows

$$P_D = \int_{C_1} p_1(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S = \int_{C_1} \frac{p_1(\mathbf{x}_S)}{p_0(\mathbf{x}_S)} p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S \tag{3.2a}$$

$$= \int_{C_1} \lambda(\mathbf{x}_S) p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S \tag{3.2b}$$

$$= \int_{\{\mathbf{x}_S : \lambda(\mathbf{x}_S) \geq \tau_u\}} \lambda(\mathbf{x}_S) p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S + \int_{\{\mathbf{x}_S : \tau_l \leq \lambda(\mathbf{x}_S) < \tau_u\}} \lambda(\mathbf{x}_S) p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S \tag{3.2c}$$

$$= \int_{\{\mathbf{x}_S : \lambda(\mathbf{x}_S) \geq \tau_u\}} \lambda(\mathbf{x}_S) p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S \tag{3.2d}$$

$$\geq \tau_u P_{FA}, \tag{3.2e}$$

where eq.(3.2d) follows from the assumption $P_i(S < \infty) = 1, i = 0, 1$. Via same procedure, $1 - P_{FA}$ at time S can be written as

$$1 - P_{FA} = 1 - \int_{C_1} p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S = \int_{C_0} p_0(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S \tag{3.3a}$$

$$= \int_{C_0} \frac{p_0(\mathbf{x}_S)}{p_1(\mathbf{x}_S)} p_1(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S = \int_{C_0} \frac{p_1(\mathbf{x}_S)}{\lambda(\mathbf{x}_S)} \mathrm{d}\mathbf{x}_S \geq \frac{1}{\tau_l} \int_{C_0} p_1(\mathbf{x}_S) \mathrm{d}\mathbf{x}_S \tag{3.3b}$$

$$\geq \frac{1 - P_D}{\tau_l}, \tag{3.3c}$$

where

$$C_0 = \{\mathbf{x}_S : \lambda(\mathbf{x}_S) \leq \tau_l\}.$$

Equations eqs.(3.2e)-(3.3c) give us bounds on the thresholds necessary to achieve the desired P_D and P_{FA}:

$$\tau_u \leq \frac{P_D}{P_{FA}}, \qquad \tau_l \geq \frac{1 - P_D}{1 - P_{FA}}.$$

Let us set

$$\tau_u = \frac{1 - \beta}{\alpha}, \tag{3.4}$$

$$\tau_l = \frac{\beta}{1 - \alpha}. \tag{3.5}$$

These thresholds guarantee that the error probabilities of the test will be at least as small as specified by choice of P_D and P_{FA}, i.e. by choice of α and $1 - \beta$. Let us refer to [24] for more details. From now on, let us denote $P_{FA} \leq \alpha$ and $P_D \geq 1 - \beta$ as Wald's criteria and denote τ_u and τ_l, given by eqs.(3.4) and (3.5) as Wald's thresholds. Despite the fact that Wald's criteria are guaranteed, Wald's thresholds may be too conservative. The expected stopping time $\mathbb{E}[S]$ can be found by using Wald's Equation [25][26]:

Theorem 3 (Wald's Equation)**.** *If S is a stopping time with respect to an i.i.d. sequence $\{X_i : i \geq 1\}$ and if $\mathbb{E}[S] < \infty$ and $\mathbb{E}[|X|] < \infty$, then*

$$\mathbb{E}\left[\sum_{i=1}^{S} X_i\right] = \mathbb{E}[S]\mathbb{E}[X]. \tag{3.6}$$

For later purposes, eq.(3.6) is rewritten as

$$\mathbb{E}\left[\sum_{i=1}^{S} \ln(X_i)\right] = \mathbb{E}[S]\mathbb{E}[\ln(X)]. \tag{3.7}$$

This is proper if $\mathbb{E}[|\ln(X)|] < \infty$ holds. Using that at stopping time S,

$$\lambda(\mathbf{x}_S) = \prod_{i=1}^{S} \frac{p_1(x_i)}{p_0(x_i)},$$

$\ln(\lambda(\mathbf{x}_S))$ can be written as

$$\ln(\lambda(\mathbf{x}_S)) = \sum_{i=1}^{S} \ln\left(\frac{p_1(x_i)}{p_0(x_i)}\right) = \sum_{i=1}^{S} \ln(X_i).$$

Let $\mathbb{E}_j$ denote the expectation with respect to p_j, $j = 0, 1$. $\mathbb{E}_j\left[\ln(\lambda(\mathbf{x}_S))\right] = \mathbb{E}_j\left[\sum_{i=1}^{S} \ln(X_i)\right]$ can be approximated with the following assumption: when $\ln(\lambda(\mathbf{x}_S)) \geq \ln(\tau_u)$, the SLRT decides $\mathcal{H}_1$ and $\ln(\lambda(\mathbf{x}_S))$ will be just slightly higher than $\ln(\tau_u)$ and also when $\ln(\lambda(\mathbf{x}_S)) \leq \ln(\tau_l)$, $\ln(\lambda(\mathbf{x}_S))$ will be just slightly lower than $\ln(\tau_l)$. Hence, expressions for the approximate values of $\mathbb{E}_j[\ln(\lambda(\mathbf{x}_S))]$ are given by:

$$\begin{aligned} E_1[\ln(\lambda(\mathbf{x}_S))] &\approx P_D \cdot \ln(\tau_u) + (1 - P_D) \cdot \ln(\tau_l) \\ &\approx P_D \cdot \ln\left(\frac{1-\beta}{\alpha}\right) + (1 - P_D) \cdot \ln\left(\frac{\beta}{1-\alpha}\right), \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} E_0[\ln(\lambda(\mathbf{x}_S))] &\approx P_{FA} \cdot \ln(\tau_u) + (1 - P_{FA}) \cdot \ln(\tau_l) \\ &\approx P_{FA} \cdot \ln\left(\frac{1-\beta}{\alpha}\right) + (1 - P_{FA}) \cdot \ln\left(\frac{\beta}{1-\alpha}\right). \end{aligned} \tag{3.9}$$

$\mathbb{E}_j[\ln(X)] = \mathbb{E}_j[\ln(\lambda(x))]$ can be calculated with the Kullback-Leibler divergence [27]:

Theorem 4 (Kullback-Leibler Divergence)**.** *For distributions P_1 and P_0 of a continuous random variable, the Kullback-Leibler divergence is defined to be the integral:*

$$D_{KL}(P_1||P_0) = \int_{-\infty}^{\infty} p_1(x) ln\left(\frac{p_1(x)}{p_0(x)}\right) \mathrm{d}x,$$

where p_1 and p_0 denote the densities of P_1 and P_0.

Observe that $\ln\left(\frac{p_1(x)}{p_0(x)}\right) = \ln(\lambda(x))$, such that $\mathbb{E}_j[\ln(\lambda(x))]$ is given by

$$\begin{aligned} \mathbb{E}_j[\ln(\lambda(x))] &= \mathbb{E}_j\left[\ln\left(\frac{p_1(x)}{p_0(x)}\right)\right] \\ &= \int_{-\infty}^{\infty} \ln\left(\frac{p_1(x)}{p_0(x)}\right) p_j(x) \mathrm{d}x \end{aligned}$$

$$= \begin{cases} \int_{-\infty}^{\infty} p_1(x) \ln\left(\frac{p_1(x)}{p_0(x)}\right) \mathrm{d}x, & j = 1 \\ -\int_{-\infty}^{\infty} p_0(x) \ln\left[\left(\frac{p_1(x)}{p_0(x)}\right)^{-1}\right] \mathrm{d}x, & j = 0 \end{cases}$$

$$= \begin{cases} D_{KL}(P_1||P_0), & j = 1 \\ -D_{KL}(P_0||P_1), & j = 0 \end{cases} \tag{3.10}$$

In (3.7), substitute $\mathbb{E}_j\left[\sum_{i=1}^{S} \ln(X_i)\right]$ by eqs.(3.8) and (3.9) and substitute $\mathbb{E}_j[\ln(X)]$ by eq.(3.10), such that the expected stopping time can be written as

$$\mathbb{E}[S] \approx \begin{cases} \frac{\mathbb{E}_1\left[\sum_{i=1}^{S} \ln(X_i)\right]}{\mathbb{E}_1[\ln(X)]}, & j = 1 \\ \frac{\mathbb{E}_0\left[\sum_{i=1}^{S} \ln(X_i)\right]}{\mathbb{E}_0[\ln(X)]}, & j = 0 \end{cases}$$

$$\approx \begin{cases} \frac{P_D \cdot \ln\left(\frac{1-\beta}{\alpha}\right) + (1-P_D) \cdot \ln\left(\frac{\beta}{1-\alpha}\right)}{D_{KL}(P_1||P_0)}, & j = 1 \\ -\frac{P_{FA} \cdot \ln\left(\frac{1-\beta}{\alpha}\right) + (1-P_{FA}) \cdot \ln\left(\frac{\beta}{1-\alpha}\right)}{D_{KL}(P_0||P_1)}, & j = 0 \end{cases} \tag{3.11}$$

Next, the expression for the expected stopping times for two different often used probability density functions are derived by calculating $D_{KL}(P_1||P_0)$ and $D_{KL}(P_0||P_1)$.

Example 5. Assume p_0 is Gaussian background noise whose power distribution is given by

$$p_0(x) = \frac{1}{\mu_b} \mathrm{e}^{-x/\mu_b}, \quad x \in \mathbb{R}^+$$

and that p_1 is the distribution of a target and background, given by

$$p_1(x) = \frac{1}{\mu_b + \mu_t} \mathrm{e}^{-x/(\mu_b+\mu_t)}, \quad x \in \mathbb{R}^+.$$

Then, $D_{KL}(P_1||P_0)$ is given by

$$\begin{aligned} D_{KL}(P_1||P_0) &= \int_{\infty}^{\infty} p_1(x) \ln\left(\frac{p_1(x)}{p_0(x)}\right) \mathrm{d}x \\ &= \int_0^{\infty} \frac{\mathrm{e}^{-x/(\mu_b+\mu_t)}}{\mu_b + \mu_t} \ln\left(\frac{\frac{1}{\mu_b+\mu_t}\mathrm{e}^{-x/(\mu_b+\mu_t)}}{\frac{1}{\mu_b}\mathrm{e}^{-x/\mu_b}}\right) \mathrm{d}x \\ &= \int_0^{\infty} \frac{\mathrm{e}^{-x/(\mu_b+\mu_t)}}{\mu_b + \mu_t} \ln\left(\left(\frac{\mu_b}{\mu_b + \mu_t}\right) \mathrm{e}^{-x/(\mu_b+\mu_t)+x/\mu_b}\right) \mathrm{d}x \\ &= \int_0^{\infty} \frac{\mathrm{e}^{-x/(\mu_b+\mu_t)}}{\mu_b + \mu_t} \left(\ln\left(\frac{\mu_b}{\mu_b + \mu_t}\right) + \frac{\mu_t x}{\mu_b(\mu_b + \mu_t)}\right) \mathrm{d}x \end{aligned}$$

$$= \ln\left(\frac{\mu_b}{\mu_b+\mu_t}\right)\left[-\mathrm{e}^{-x/(\mu_b+\mu_t)}\right]_0^\infty - \left[\frac{\mu_t x}{\mu_b(\mu_b+\mu_t)}\mathrm{e}^{-x/(\mu_b+\mu_t)}\right]_0^\infty$$

$$- \left[\frac{\mu_t}{\mu_b}\mathrm{e}^{-x/(\mu_b+\mu_t)}\right]_0^\infty$$

$$= \ln\left(\frac{\mu_b}{\mu_b+\mu_t}\right) + \frac{\mu_t}{\mu_b}. \tag{3.12}$$

Similar calculations are used to find $D_{KL}(P_0||P_1)$:

$$D_{KL}(P_0||P_1) = \int_\infty^\infty p_1(x)\ln\left(\frac{p_1(x)}{p_0(x)}\right)\mathrm{d}x$$

$$= \int_0^\infty \frac{\mathrm{e}^{-x/\mu_b}}{\mu_b}\ln\left(\frac{\frac{1}{\mu_b}\mathrm{e}^{-x/\mu_b}}{\frac{1}{\mu_b+\mu_t}\mathrm{e}^{-x/(\mu_b+\mu_t)}}\right)\mathrm{d}x$$

$$= \int_0^\infty \frac{\mathrm{e}^{-x/\mu_b}}{\mu_b}\left(\ln\left(1+\frac{\mu_t}{\mu_b}\right) - \frac{\mu_t x}{\mu_b(\mu_b+\mu_t)}\right)\mathrm{d}x$$

$$= \ln\left(1+\frac{\mu_t}{\mu_b}\right) - \frac{\mu_t}{\mu_b+\mu_t} \tag{3.13}$$

Plugging in eq.(3.12) and eq.(3.13) in eq.(3.11), the expected stopping time in Gaussian noise is given by

$$\mathbb{E}[S] \approx \begin{cases} \frac{P_D\cdot\ln\left(\frac{1-\beta}{\alpha}\right)+(1-P_D)\cdot\ln\left(\frac{\beta}{1-\alpha}\right)}{\ln\left(\frac{\mu_b}{\mu_b+\mu_t}\right)+\frac{\mu_t}{\mu_b}}, & j=1 \\ \\ -\frac{P_{FA}\cdot\ln\left(\frac{1-\beta}{\alpha}\right)+(1-P_{FA})\cdot\ln\left(\frac{\beta}{1-\alpha}\right)}{\ln\left(1+\frac{\mu_t}{\mu_b}\right)-\frac{\mu_t}{\mu_b+\mu_t}}, & j=0 \end{cases} \tag{3.14}$$

Example 6. A commonly used distribution to model (sea) clutter is a generalized Pareto distribution:

$$p_0(x) = \frac{1}{s}\left(1 + k\frac{x}{s}\right)^{-1-1/k}, \quad k, s > 0, \quad x \in \mathbb{R}^+.$$

In Section 2.3.3 it is assumed that $p_1(x)$ (target+clutter) is generalized Pareto distributed as well, with new parameters:

$$p_1(x) = \frac{1}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}} \quad \tilde{k}, \tilde{s} > 0, \quad x \in \mathbb{R}^+.$$

Now,

$$D_{KL}(P_1||P_0) = \int_{-\infty}^{\infty} p_1(x)\ln\left(\frac{p_1(x)}{p_0(x)}\right)\mathrm{d}x$$

$$= \int_0^{\infty} \frac{1}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}} \ln\left[\frac{s}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}}\left(1 + k\frac{x}{s}\right)^{1+1/k}\right]\mathrm{d}x$$

$$= \int_0^{\infty} \frac{1}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}} \left[\ln\left(\frac{s}{\tilde{s}}\right) - \left(1 + \frac{1}{\tilde{k}}\right)\ln\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right) + \left(1 + \frac{1}{k}\right)\ln\left(1 + k\frac{x}{s}\right)\right]\mathrm{d}x$$

$$= \int_0^{\infty} \frac{\ln(s/\tilde{s})}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}}\mathrm{d}x \qquad I_1$$

$$+ \int_0^{\infty} \frac{-1}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}} \left[\left(1 + \frac{1}{\tilde{k}}\right)\ln\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)\right]\mathrm{d}x \qquad I_2$$

$$+ \int_0^{\infty} \frac{1}{\tilde{s}}\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1-1/\tilde{k}} \left[\left(1 + \frac{1}{k}\right)\ln\left(1 + k\frac{x}{s}\right)\right]\mathrm{d}x \qquad I_3,$$

with

$$I_1 = \left[-\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1/\tilde{k}}\ln\left(\frac{s}{\tilde{s}}\right)\right]_0^{\infty} = \ln\left(\frac{s}{\tilde{s}}\right),$$

$$I_2 = \left[\frac{(\tilde{k}+1)\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1/\tilde{k}}\left(\ln\left(1 + \frac{\tilde{k}x}{\tilde{s}}\right) + \tilde{k}\right)}{\tilde{k}}\right]_0^{\infty} = -(\tilde{k}+1)$$

$$I_3 = \left[\left(\frac{1}{k} + 1\right)\left(1 + \tilde{k}\frac{x}{\tilde{s}}\right)^{-1/\tilde{k}}\left(\tilde{k}\cdot {}_2F_1\left(1; \frac{-1}{\tilde{k}}; \frac{\tilde{k}-1}{\tilde{k}}; \frac{k(\tilde{s}+\tilde{k}x)}{k\tilde{s}-\tilde{k}s}\right) - \ln\left(1 + \frac{kx}{s}\right) - \tilde{k}\right)\right]_0^{\infty}$$

$$= \left(\frac{1}{k} + 1\right)\left[\pi\left(\frac{k\tilde{s}}{\tilde{k}s}\right)^{1/\tilde{k}}\left(1 - \frac{k\tilde{s}}{\tilde{k}s}\right)^{-1/\tilde{k}}\csc\left(\pi/\tilde{k}\right) - \frac{k\tilde{s}\cdot {}_2F_1\left(1;1;2-\frac{1}{\tilde{k}};\frac{k\tilde{s}}{\tilde{k}s}\right)}{s(\tilde{k}-1)}\right],$$

where I_3 is calculated with Wolfram Mathematica [28], with

$$\csc(\theta) = \frac{1}{\sin(\theta)}$$

and where ${}_2F_1$ is the Hypergeometric function [29]. Unfortunately, the result of I_3 is not a useful expression as the value ${}_2F_1\left(1;1;2-\frac{1}{\tilde{k}};\frac{k\tilde{s}}{\tilde{k}s}\right)$ is not clear.

3.1.2 Neyman-Pearson testing

In this section the Neyman-Pearson theory is discussed. Consider the simple hypothesis

$$\mathcal{H}_i : \quad x_1, x_2, \ldots, x_N \sim p_i, \quad i = 0, 1, \tag{3.15}$$

and the likelihood ratio test

$$\lambda(x_1, x_2, \ldots, x_N) = \frac{p(x_1, x_2, \ldots, x_N|\mathcal{H}_1)}{p(x_1, x_2, \ldots, x_N|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau$$

with i.i.d. x_i and a predetermined decision time N. The LRT is summarized as follows:

Definition 7 (Likelihood ratio test)**.**
If $\lambda(\mathbf{x}_N) \geq \tau$, choose $\mathcal{H}_1$, else choose $\mathcal{H}_0$

The most powerful test is the test that maximizes P_D for a certain significance level $P_{FA} \leq \alpha$. Neyman-Pearson solves this optimization problem

$$\text{argmax}_x P_D(x), \qquad s.t. \qquad P_{FA}(x) \leq \alpha,$$

with Lagrange multipliers [30], which led to the following theorem [23]:

Theorem 8 (Neyman-Pearson theorem)**.** *To maximize P_D for a given $P_{FA} = \alpha$, decide $\mathcal{H}_1$ if*

$$\lambda(x) = \frac{p(x|\mathcal{H}_1)}{p(x|\mathcal{H}_0)} > \tau,$$

where the threshold τ is found by

$$P_{FA} = \int_{\{x:\lambda(x)>\tau\}} p(x|\mathcal{H}_0)\mathrm{d}x = \alpha.$$

Now a new SLRT is defined that uses Neyman-Pearson theory to calculate the threshold at each time iteratively to. Hence, in contrary to Wald's thresholds, Neyman-Pearson threshold is time dependent, because with each additional sample, $\lambda(\mathbf{x})$ and $p(\mathbf{x}|\mathcal{H}_0)$ change. Let us define the sequential likelihood ratio test (SLRT) with Neyman-Pearson as follows:

Definition 9 (SLRT with Neyman-Pearson)**.**

1. If $\lambda(\mathbf{x}_n) \geq \tau_n$, choose $\mathcal{H}_1$, else go to 2
2. If $n < N, n = n + 1$, go to 1, else go to 3
3. choose $\mathcal{H}_0$

The test at time n is given by

$$\lambda(\mathbf{x}_n) = \prod_{i=1}^{n} \lambda(x_i) \tag{3.16}$$

$$= \frac{\prod_{i=1}^{n} p(x_i|\mathcal{H}_1)}{\prod_{i=1}^{n} p(x_i|\mathcal{H}_0)} \tag{3.17}$$

$$= \frac{p(\mathbf{x}_n|\mathcal{H}_1)}{p(\mathbf{x}_n|\mathcal{H}_0)} > \tau_n, \quad \text{with} \quad \mathbf{x}_n \in \mathbb{R}^n$$

where the threshold τ_n is found by solving

$$P_{FA} = \int_{\{\mathbf{x}_n:\lambda(\mathbf{x}_n)>\tau_n\}} p(\mathbf{x}_n|\mathcal{H}_0)\mathrm{d}\mathbf{x}_n = \alpha, \tag{3.18}$$

Example 10. Again, assume p_0 and p_1 to be the power distributions as in Example 5. At time n, the critical region $\{\mathbf{x}_n : \lambda(x_1, x_2, \ldots, x_n) > \tau_n\}$ is given by

$$\begin{aligned}
\{\mathbf{x}_n : \lambda(x_1, x_2, \ldots, x_n) > \tau_n\} &= \left\{\mathbf{x}_n : \frac{p(\mathbf{x}_n|\mathcal{H}_1)}{p(\mathbf{x}_n|\mathcal{H}_0)} > \tau_n\right\} \\
&= \left\{\mathbf{x}_n : \left(\frac{\mu_b}{\mu_b+\mu_t}\right)^n \mathrm{e}^{\mu_t\left(\sum_{i=1}^n x_i\right)/(\mu_b(\mu_b+\mu_t))} > \tau_n\right\} \\
&= \left\{\mathbf{x}_n : \ln(\tau_n) < n\ln\left(\frac{\mu_b}{\mu_b+\mu_t}\right) + \frac{\mu_t\sum_{i=1}^n x_i}{\mu_b(\mu_b+\mu_t)}\right\} \\
&= \left\{\mathbf{x}_n : \sum_{i=1}^n x_i > \left(\ln(\tau_n) - n\ln\left(\frac{\mu_b}{\mu_b+\mu_t}\right)\right)\frac{\mu_b(\mu_b+\mu_t)}{\mu_t} = \tilde{\tau}_n\right\}
\end{aligned} \tag{3.19}$$

Lemma 11. *The sum of n independent identically exponentially distributed random variables with mean (μ) is Gamma distributed with parameters (n, μ). Furthermore, when n is an integer, the Gamma distribution is the same as the Erlang distribution. See [31, 32] for proof.*

Hence from Lemma 11, the pdf for the sum of n i.i.d. exponential distributed samples is given by

$$f(x, n, \mu) = \frac{1}{\mu^n\Gamma(n)}x^{n-1}\mathrm{e}^{-x/\mu}, \quad \mu > 0, \quad n \in \mathbb{N}^+, \quad x \in \mathbb{R}^+. \tag{3.20}$$

As a result, P_{FA} is given by

$$\begin{aligned}
P_{FA} &= \int_{\tilde{\tau}_n}^{\infty} \frac{1}{\Gamma(n)\mu_b^n}x^{n-1}\mathrm{e}^{-x/\mu_b}\mathrm{d}x = \alpha \\
&= -\left[\frac{1}{\Gamma(n)\mu_b^{n-1}}x^{n-1}\mathrm{e}^{-x/\mu_b}\right]_{\tilde{\tau}_n}^{\infty} + \int_{\tilde{\tau}_n}^{\infty} \frac{1}{\Gamma(n-1)\mu_b^{n-1}}x^{n-2}\mathrm{e}^{-x/\mu_b}\mathrm{d}x \\
&= \frac{\tilde{\tau}_n^{n-1}\mathrm{e}^{-x/\mu_b}}{\Gamma(n)\mu_b^{n-1}} - \left[\frac{1}{\Gamma(n-1)\mu_b^{n-2}}x^{n-2}\mathrm{e}^{-x/\mu_b}\right]_{\tilde{\tau}_n}^{\infty} + \int_{\tilde{\tau}_n}^{\infty} \frac{1}{\Gamma(n-2)\mu_b^{n-2}}x^{n-3}\mathrm{e}^{-x/\mu_b}\mathrm{d}x
\end{aligned}$$

After performing these calculations n times, P_{FA} will be given by

$$P_{FA} = \sum_{i=1}^n \left(\frac{(\tilde{\tau}_n/\mu_b)^{i-1}}{\Gamma(i)}\right)\mathrm{e}^{-\tilde{\tau}_n/\mu_b} = \alpha. \tag{3.21}$$

$\tilde{\tau}_n$ can be found numerically by setting α at a desired level and using the cumulative distribution function (cdf) of the Gamma distribution to find the maximum value for $\tilde{\tau}_n$ such that $P_{FA} \leq \alpha$. Once, $\tilde{\tau}$ is found, τ_n can be found from

$$\tau_n = \left(\frac{\mu_b}{\mu_b + \mu_t}\right)^n \mathrm{e}^{\tilde{\tau}_n \mu_t / (\mu_b(\mu_b+\mu_t))}. \tag{3.22}$$

Finding τ_n in eq.(3.18) for the clutter distribution requires the pdf for the sum of i.i.d. generalized Pareto distributed r.v's. In [33] the expression for the sum of two samples is derived. Unfortunately, the expression for multiple samples (> 3) is not clear. Therefore, no example of performing the calculations likewise Example 10 are given.

3.2 Simulation results i.i.d. data

In this section, the hypothesis testing theories described in Section 3.1 are tested with generated i.i.d. samples. The results will be thoroughly compared later with the results with the data from the particle filter. Also discussions regarding to the desired CFAR property are given.

Data samples are generated according to either $p_1(x)$ or $p_0(x)$. In this section, these are assumed to be known exactly. The likelihood ratios, $\lambda(x) = \frac{p_1(x)}{p_0(x)}$, are calculated for the two data sets. These likelihood ratios are tested against Wald's thresholds (3.4)-(3.5) and Neyman-Pearson threshold (3.18) to obtain the P_{FA} and P_D values.

3.2.1 Results Wald's theory with i.i.d. data

Gaussian Noise

In Gaussian noise, the expected stopping time $\mathbb{E}[S]$ can be calculated with eq.(3.14). First, two scenarios are given for which $\mathbb{E}[S]$ is calculated.
Scenario 1: suppose that $\mu_t = 10$, $\mu_b = 1$ (SNR = 10dB), $\alpha \leq 0.001$ and $1 - \beta \geq 0.9$, then $\mathbb{E}[S]$ is given by:

$$\mathbb{E}[S] \approx \begin{cases} \frac{0.9 \cdot \ln\left(\frac{0.9}{0.001}\right) + (1-0.9) \cdot \ln\left(\frac{0.1}{1-0.001}\right)}{\ln\left(\frac{1}{1+10}\right) + \frac{10}{1}}, & j = 1 \\ -\frac{0.001 \cdot \ln\left(\frac{0.9}{0.001}\right) + (1-0.001) \cdot \ln\left(\frac{0.1}{1-0.001}\right)}{\ln\left(1 + \frac{10}{1}\right) - \frac{10}{1+10}}, & j = 0 \end{cases}$$

$$\approx \begin{cases} 0.775, & j = 1 \\ 1.540, & j = 0 \end{cases} \tag{3.23}$$

Scenario 2: suppose that $\mu_t = 10$, $\mu_b = 5$ (SNR = 3dB), $\alpha \leq 0.001$ and $1 - \beta \geq 0.5$, then $\mathbb{E}[S]$ is

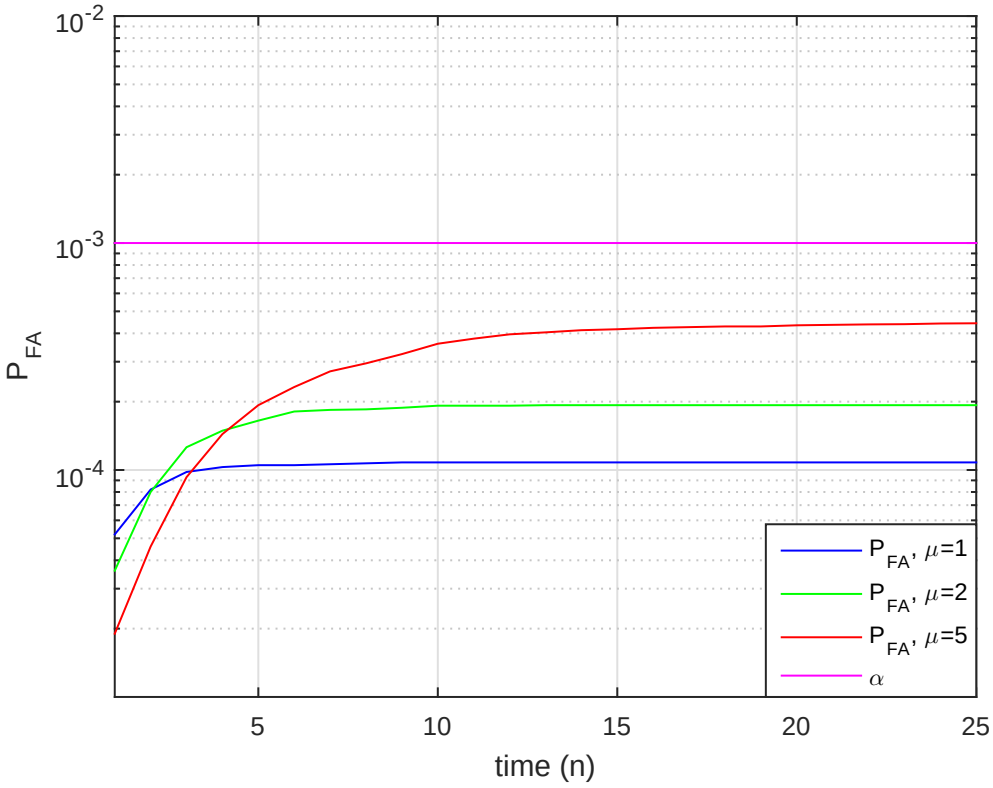
$$\mathbb{E}[S] \approx \begin{cases} \frac{0.5 \cdot \ln\left(\frac{0.5}{0.001}\right) + (1-0.5) \cdot \ln\left(\frac{0.5}{1-0.001}\right)}{\ln\left(\frac{5}{5+10}\right) + \frac{10}{5}}, & j = 1 \\ -\frac{0.001 \cdot \ln\left(\frac{0.5}{0.001}\right) + (1-0.001) \cdot \ln\left(\frac{0.5}{1-0.001}\right)}{\ln\left(1 + \frac{10}{5}\right) - \frac{10}{5+10}}, & j = 0 \end{cases}$$

$$\approx \begin{cases} 3.063, & j = 1 \\ 1.586, & j = 0. \end{cases} \tag{3.24}$$

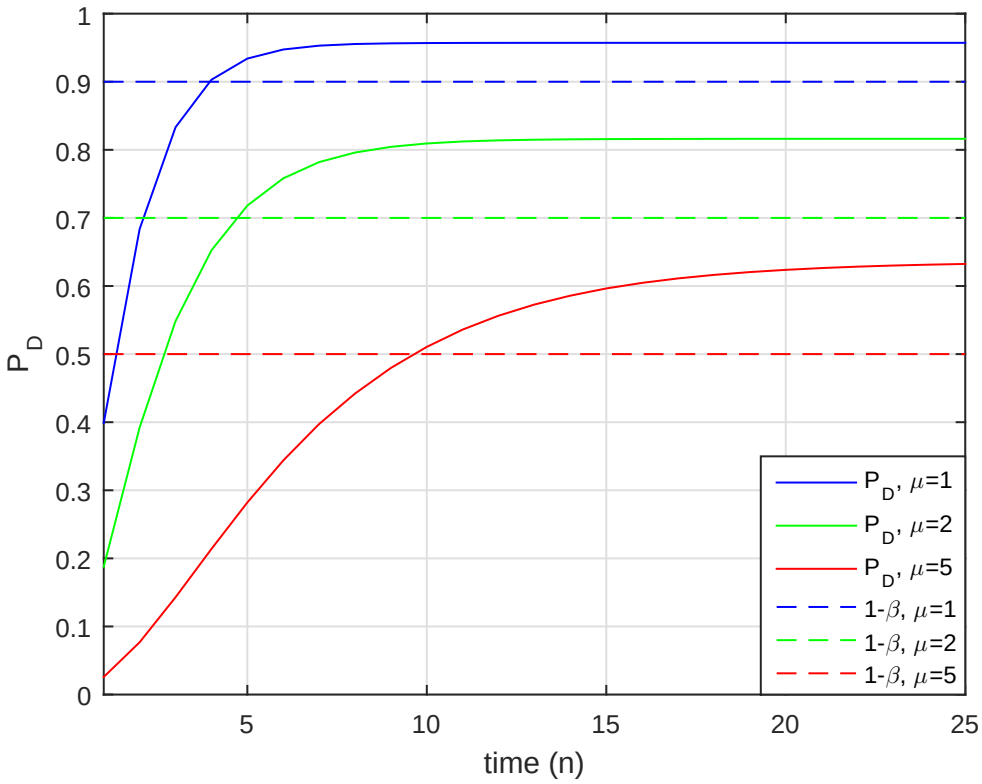
In Figure 3.1 the P_{FA} and P_D curves are plotted for three different noise powers ($\mu = 1, 2$ and 5) and target power $\mu_t = 10$ for 10^6 i.i.d. generated samples. First of all, following Appendix A in [22], this sample size gives very reliable results for the values plotted in the figure. Figure 3.1a shows that all three curves fulfill the P_{FA} criterion ($P_{FA} \leq \alpha$). However, the curves do not reach the purple line, which could indicate that the lower threshold τ_l is too high. Furthermore, since three different P_D criteria ($P_D \geq 1 - \beta$) are used for the three noise powers, the likelihoods for the corresponding noise powers are tested against different upper and lower thresholds. Hence, nothing general can be said about the CFAR property, because the CFAR property acquires the condition of the same threshold for all background conditions. Figure 3.1b shows that all P_D curves (solid lines) fulfill the corresponding P_D criterion (dotted lines) as well. It also shows that with a higher noise power level, it takes longer to reach the criterion, which coincides with the calculations of $\mathbb{E}[S]$ in Scenario 1 and 2.

Pareto Clutter

Although, the final expression for $\mathbb{E}[S]$ has not been derived in case of Pareto clutter, Wald's theory is tested regardless to see whether or not Wald's criteria are reached in finite time. In Figure 3.2 the P_{FA} and P_D curves are plotted for 10^6 i.i.d. clutter samples. The figure shows that for all three shape parameters k, Wald's P_D criterion are reached in finite time, while still fulfilling the P_{FA} criterion as well. Figure 3.2a indicates that the lower threshold might be a bit too high as no P_{FA} comes very close to the purple line.
Contrary to the Gaussian noise case where three different thresholds were used, in the Pareto case the thresholds for all three shape parameters k are the same. Unfortunately, the P_{FA} curves still do not show signs of the CFAR property as the values for different k, both for n is low and n is high, are very different.

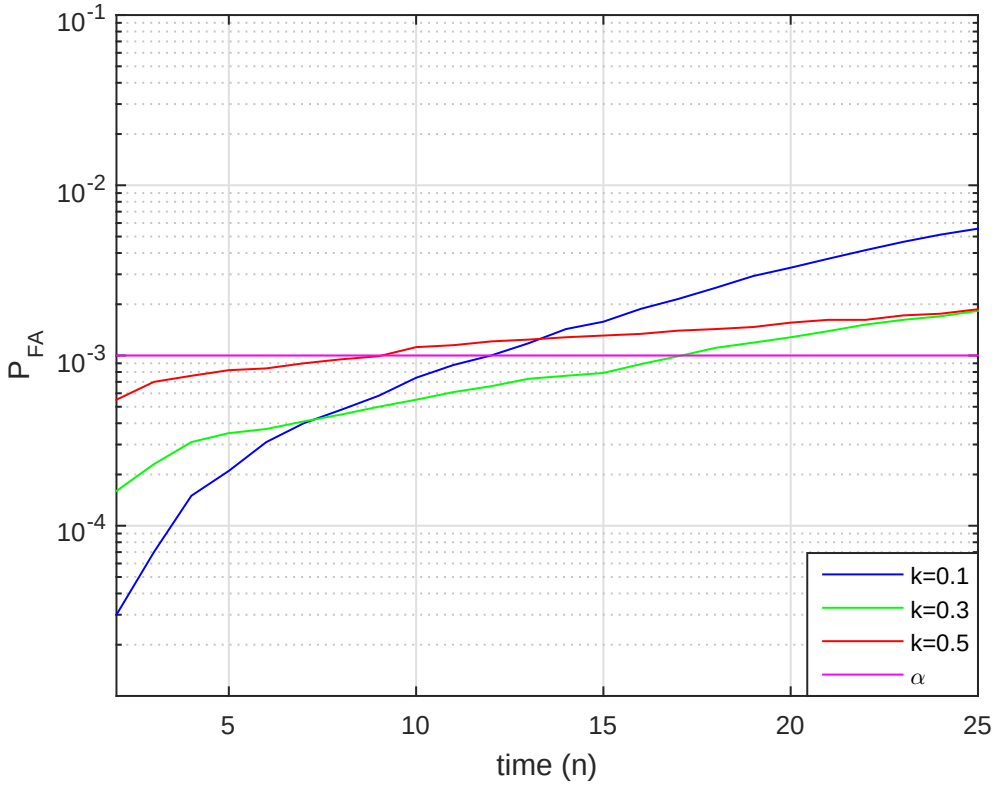


(a) False Alarms

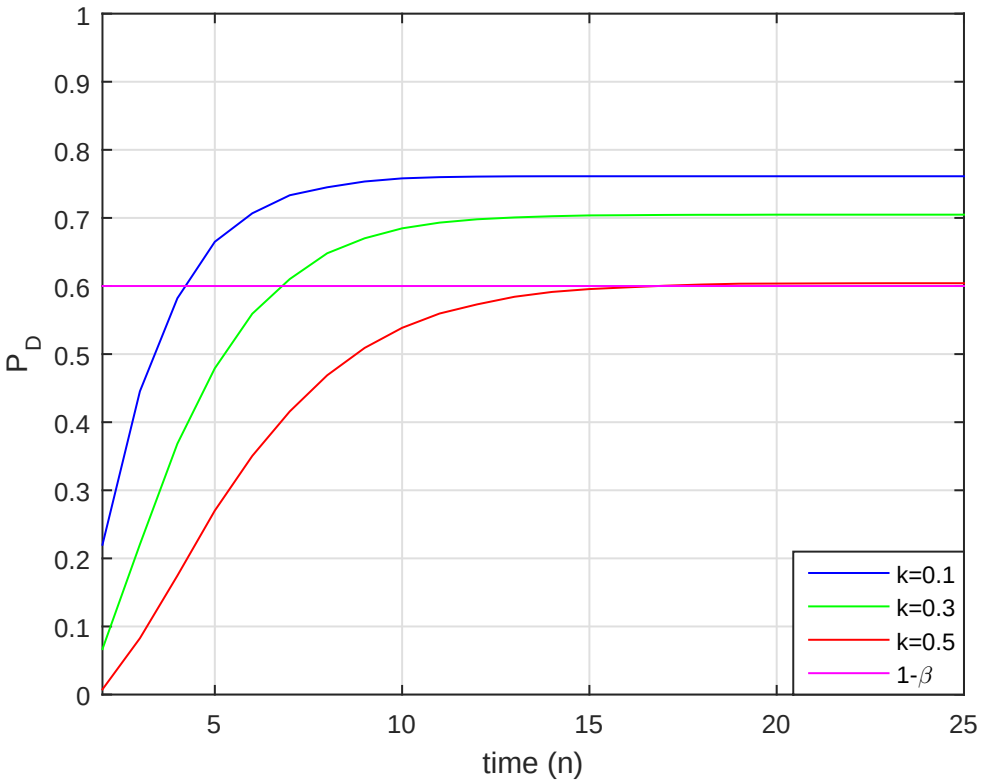


(b) Detections

Figure 3.1: P_{FA} and P_D curves with Wald's criteria for i.i.d. noise samples



(a) False Alarms



(b) Detections

Figure 3.2: P_{FA} and P_D curves with Wald's criteria for i.i.d. clutter samples

3.2.2 Results Neyman-Pearson theory with i.i.d. data

In this section, Neyman-Pearson Theory is tested and the results for the standard LRT detection scheme with Neyman-Pearson threshold are compared to the new sequential detection scheme 9. Note again that the threshold is varying in time, because at each time n, a new sample will be available and as a result, the threshold needs to be updated. Only Gaussian background noise is considered, because the expressions for the critical region $\{\mathbf{x}_n : \lambda(\mathbf{x}_n) > \tau_n\}$ and $p_0(\mathbf{x}_n)$ with $\mathbf{x}_n \in \mathbb{R}^n$ in case of Pareto clutter are not clear, as mentioned in Section 3.1.2.

The desired significance level α is constructed as follows: a final time horizon is set at $N = 25$. Hence, following detection scheme 9, the sequential test has 25 time steps to decide p_1. If p_1 is not decided after $N = 25$, automatically p_0 will be decided. At $N = 25$, Sum α is set at 10^{-3} in the sense that the summed P_{FA} values from 1 to N should satisfy $P_{FA} \leq \alpha$. Assuming that at each time, the probability of a false alarm is the same, the threshold at each time n is calculated using $\alpha = \frac{10^{-3}}{N}$ in eq.(3.18).

In Figure 3.4a, the P_{FA} curves of both the LRT and the SLRT are plotted for three different noise powers ($\mu = 1$, $\mu = 2$ and $\mu = 5$). 10^7 i.i.d. samples are used, which are generated as explained in the beginning of Section 3.2. The solid lines (–) represent the (integrated) P_{FA} values for the SLRT. The dotted lines (- -) represent the P_{FA} values for the LRT at each time, when a predetermined fixed horizon n would have been used. For proper comparison between the two detection schemes, α is set at $\frac{10^{-3}}{N}$ at each time in the LRT as well. The circled lines (-o-) show the summed P_{FA} values of the LRT, up to and including time n.

From Figure 3.4a, some interesting observations can be made. First, the dotted lines of the LRT circle closely around the pink line. This indicates that when the number of samples goes to infinity, the P_{FA} values for all three power levels will be exactly α. This shows that Neyman-Pearson theory indeed constructs the most powerful test (P_D is maximized when $P_{FA} = \alpha$, see [23] for more details). As a result of the P_{FA} values for the LRT at times 1 to N individually, the summed values (-o- lines) for all noise levels approximately match with the summed α (the purple -o- line). Ideally the solid lines from the SLRT should be close to this purple circled line as well. This would indicate that the time depending threshold in the SLRT, calculated by solving eq.(3.18), is not too high and that Neyman Pearson theory constructs the most powerful test for the sequential detection scheme, defined in 9 as well. However, the actual curves show a bit different behavior. Although the P_{FA} values for the three noise levels are again approximately the same, they do not reach the value of the desired P_{FA} (pink -o- line). The explanation is that once a false alarm is detected, the SLRT stops, whereas in the summed LRT curves, the possibility of detect a false alarm at for example time $n = 3$ and again at say time $n = 5$ is included. Hence a direct consequence of using the SLRT, is that the integrated P_{FA} values can not be controlled exactly by using Neyman-Pearson lemma. From now, the difference between the SLRT and LRT P_{FA} values will be referred to as Sequential Loss.

In Figure 3.4b, P_D curves are given for both noise levels and both the LRT and the SLRT. The target power is in both cases $\mu_t = 10$, which makes the SNR respectively 10dB (blue), 7dB(green) and 3dB (red). For the case when the SNR equals 10dB, the curves of the LRT and the SLRT are the same. Hence, for powerful targets, compared to the background, it does not matter which test is used. Note that when the SNR is really high, say bigger than 20dB, the TBD approach is not even needed and the classical detecting approach is preferred, as explained in Section 2.1.

For the target with SNR = 3dB, the LRT and the SLRT curves are different. The SLRT has a slightly higher P_D which shows the advantage of the SLRT to detect at each time: for example, the SLRT could have made a detection at $n = 4$, while at $n = 8$, the integrated likelihood ratio is lower than the threshold at time 8. This results in no target detection for the LRT, while the SLRT already made this detection at $n = 4$.

In Figure 3.3, the thresholds are given for all noise levels in the log domain. Again, these threshold are calculated by solving eq.(3.18), where all parameters are known. Clearly, the thresholds are very different. The threshold for $\mu = 1$ (blue) is decreasing from approximately 7 till -15 at $n = 25$. This means that linearly, this threshold converges to 0 very quickly. The threshold for $\mu = 2$ converges less quick to 0 and the threshold for $\mu = 5$ (red) is most stable. It increases a bit in the beginning, and decreases a bit when more samples are available. An important note is that although the P_{FA} curves in Figure 3.4a for the three power levels were approximately similar, the CFAR property is not fulfilled, because the likelihood ratios are tested against different threshold.

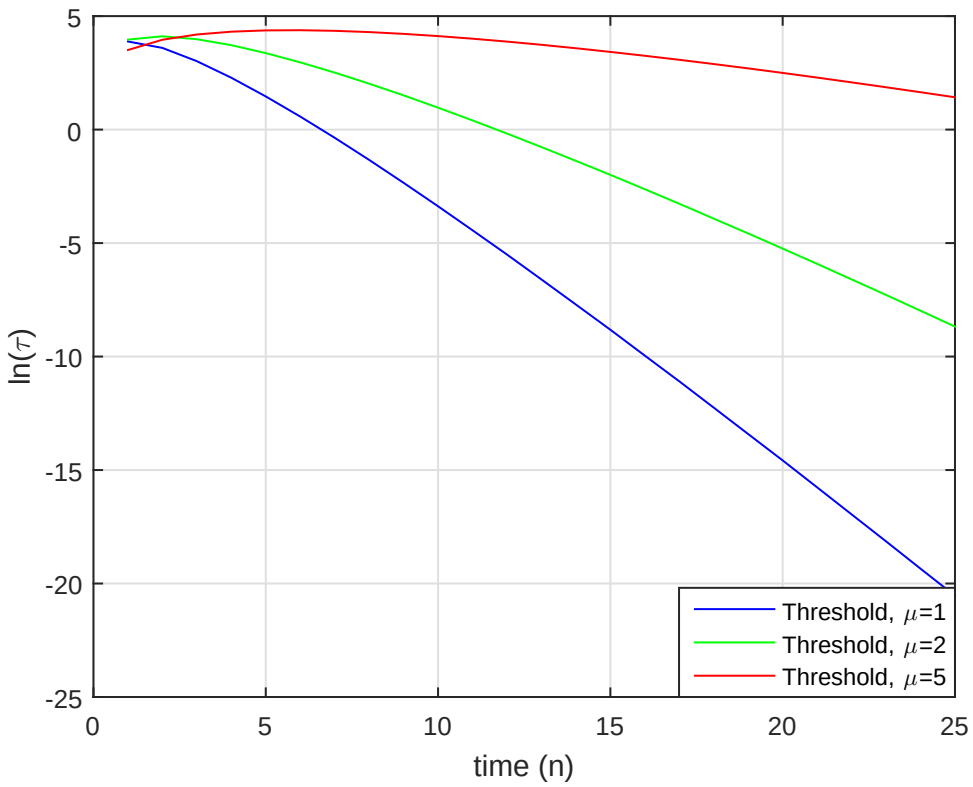
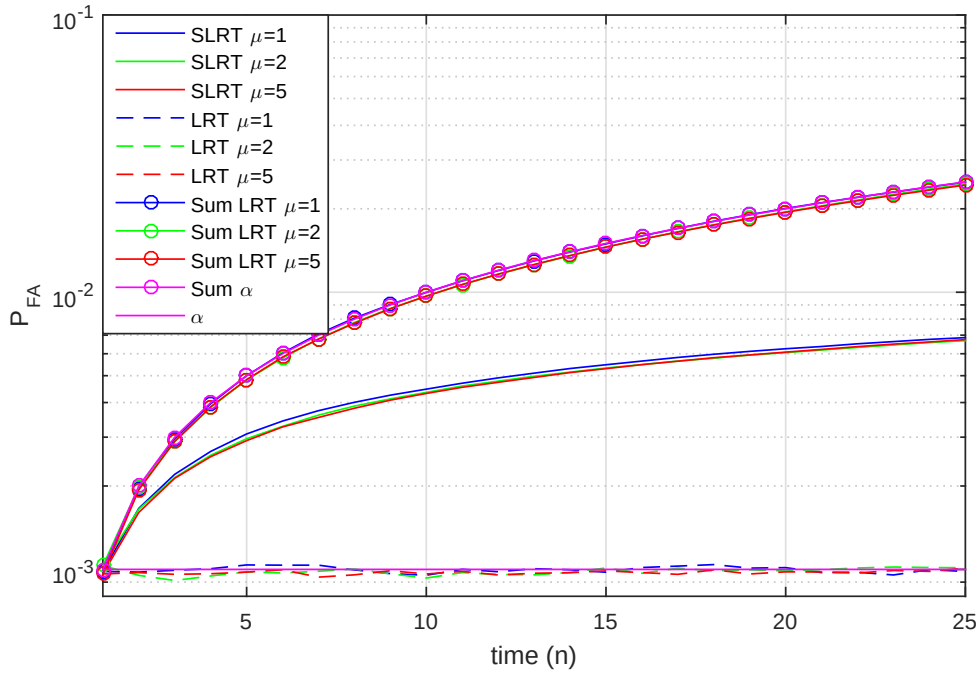
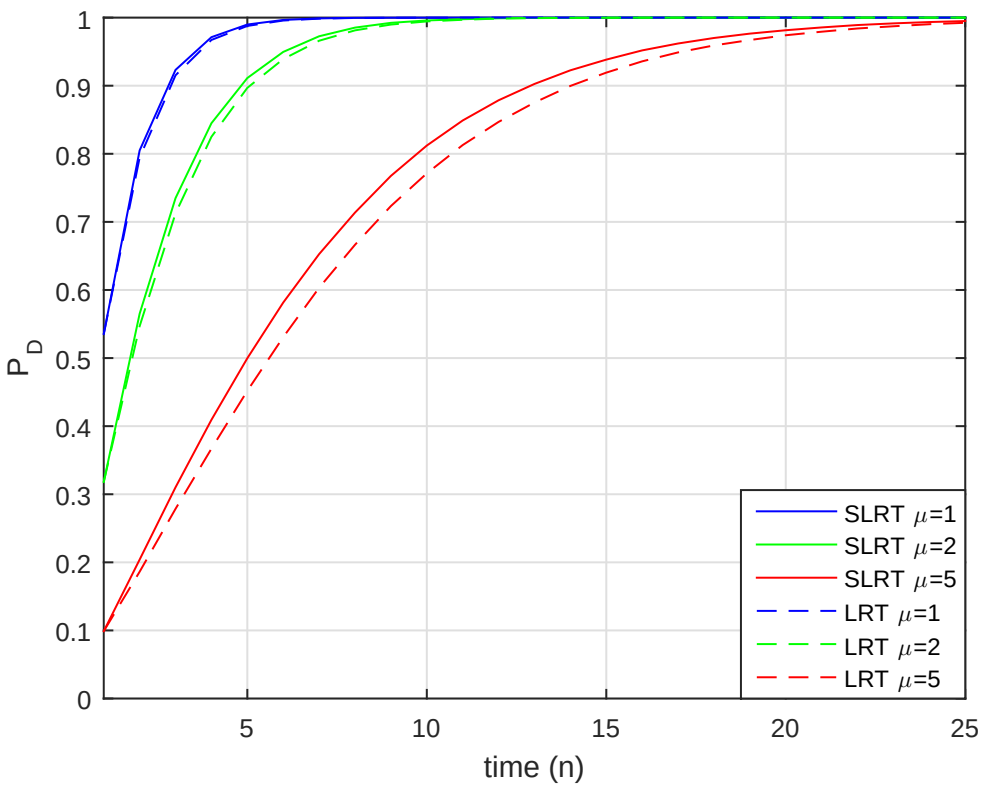


Figure 3.3: Threshold $\ln(\tau)$ for Neyman-Pearson theory



(a) P_{FA} curves



(b) P_D curves

Figure 3.4: P_{FA} and P_D curves for Neyman-Pearson theory, with i.i.d. samples

3.3 Simulation results with data from particle filter

In this section, the results for the hypothesis testing theories are shown for data acquired from the particle filter. The data samples (posterior likelihood ratios) are obtained from the filter with 10^5 Monte Carlo (MC) simulations. The results will be compared with those for the i.i.d. data in Section 3.2.

3.3.1 Results Wald's theory with filtered data

Gaussian Noise

In Figure 3.5, the results for Wald theory are shown in Gaussian noise. Figure 3.5b shows that the P_D criterion for all power levels are reached, similar to the i.i.d. data in Figure 3.1b. Also the stopping times and even the actual P_{FA} values show similarities to the results of the i.i.d. data. However, Figure 3.5a shows different results. All P_{FA} curves cross the P_{FA} criterion, which means that the lower threshold τ_l is too low to use with the filtered data. Also, at the time that the P_{FA} criterion is crossed (around $n = 7$), not all P_D criteria are fulfilled. Since the P_D criterion for $\mu = 5$ is not reached until $n = 11$, is seems that for high power levels, one Wald criterion can not be fulfilled without damaging the other.

Pareto Clutter

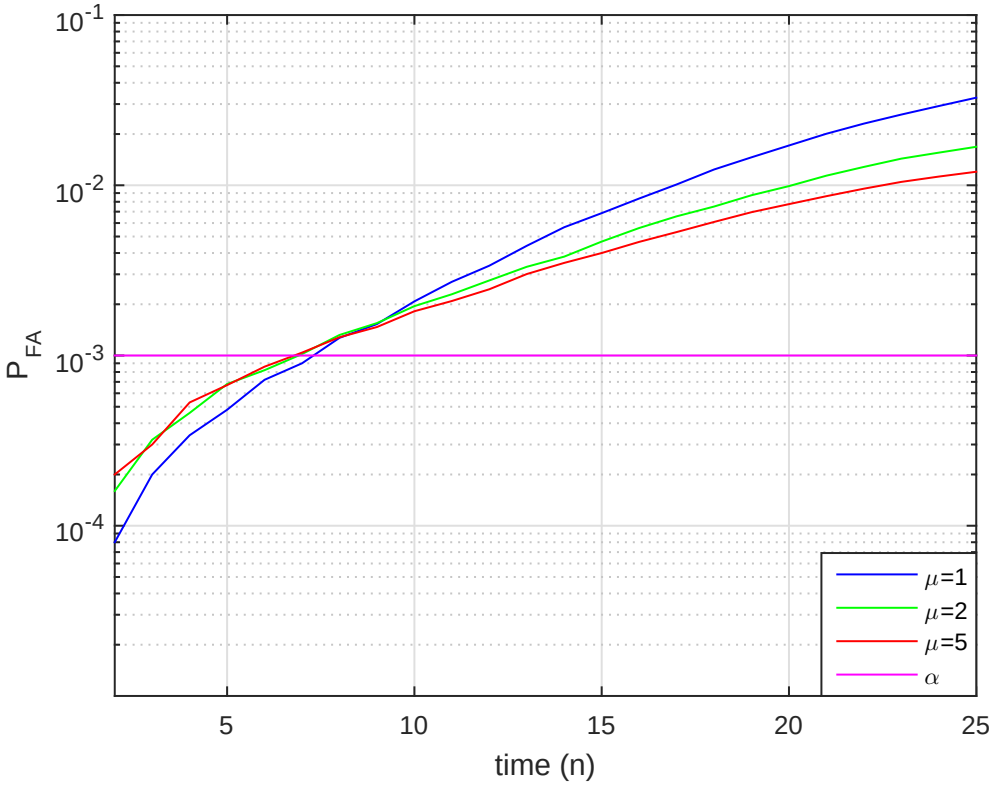
Figure 3.6 shows the results of Wald's theory for filtered samples in clutter. Similar to the results for Gaussian noise, the P_D curves satisfy the P_D criterion, although it is very close for $k = 0.5$ (red line). Also similar to Figure 3.5, the P_{FA} curves in Figure 3.6a do not fulfill Wald's P_{FA} criterion.

Overall, based on Figures 3.5 and 3.6, the observation is that Wald's theory does not work properly for filtered data from the particle filter. The explanation is given in Chapter 4.

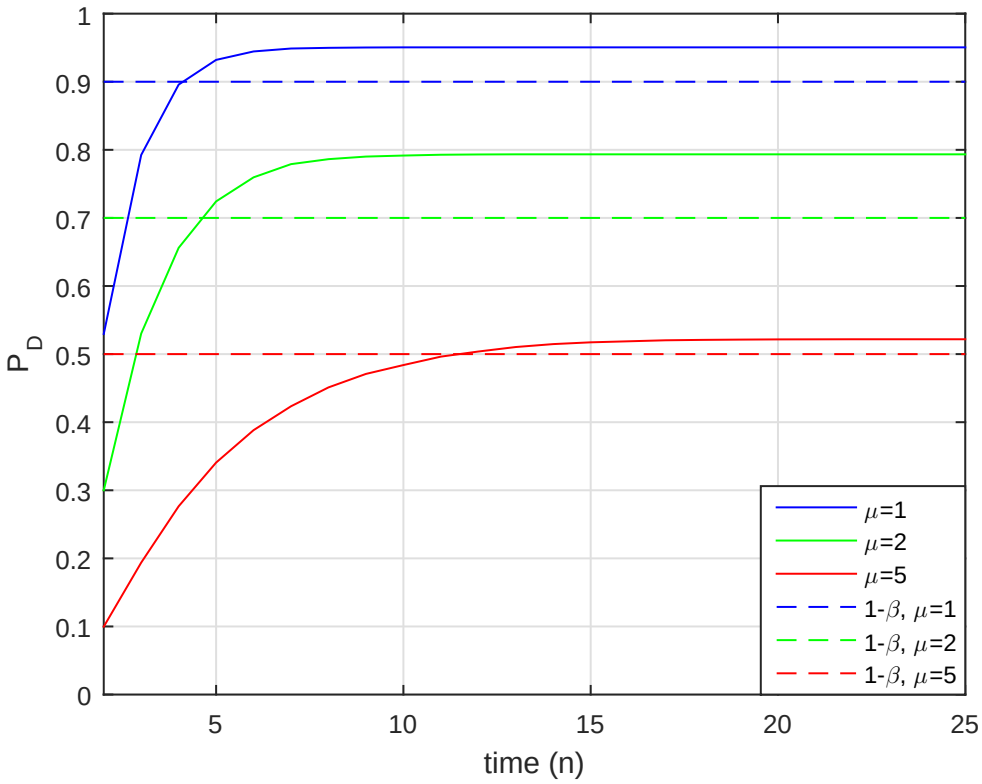
3.3.2 Results Neyman-Pearson theory with filtered data

In Figure 3.7, the results of Neyman-Pearson theory are shown with data from the particle filter. Just as in Section 3.2.2, only results for Gaussian noise are considered. Similar to the results for Wald's theory with filtered data, the P_D curves in Figure 3.7b converge to 1 for all power levels. Unfortunately, the P_{FA} curves in Figure 3.7a eventually converge to 1 as well. This is a major difference compared to the results for the i.i.d. data, where the P_{FA} curves did not even reach significance level α (the solid lines in Figure 3.4a). So the thresholds, calculated by solving eq.(3.18) and shown (for $\mu = 1$ and $\mu = 5$) in Figure 3.3, seem to be far too low. Furthermore, the false alarm rate in Figure 3.7a is certainly not constant as the P_{FA} curves for the different power levels are not close to each other. Based on these results, it seems that also Neyman-Pearson theory is not suitable for the data from the particle filter.

Remark 12. Because of the computational time of the filter, only 10^5 data samples are generated for each power level. Therefore at each time step, α is set at 10^{-3}, instead of $\frac{10^{-3}}{N}$ as in Section 3.2.2. Hence, the actual P_{FA} values in Figure 3.7 are not comparable to the values in Figures 3.4a and 3.4b. However, the behavior of the curves are comparable, which makes the comments on the results in Sections 3.3.1 and 3.3.2 proper.

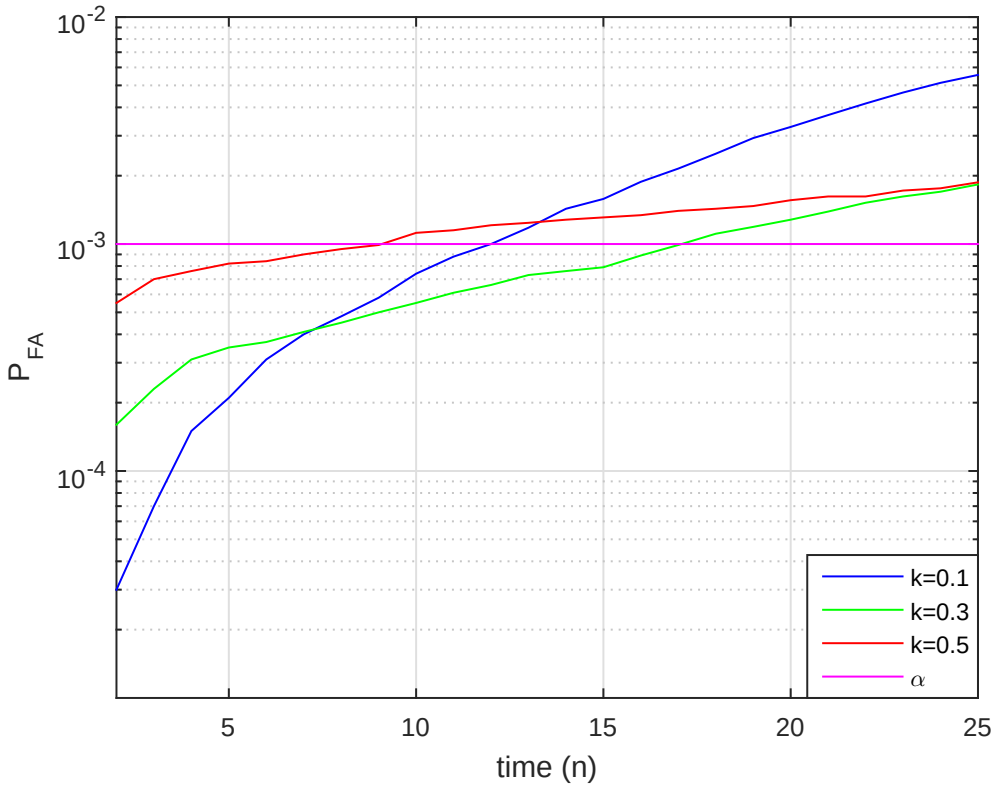


(a) P_{FA} curves

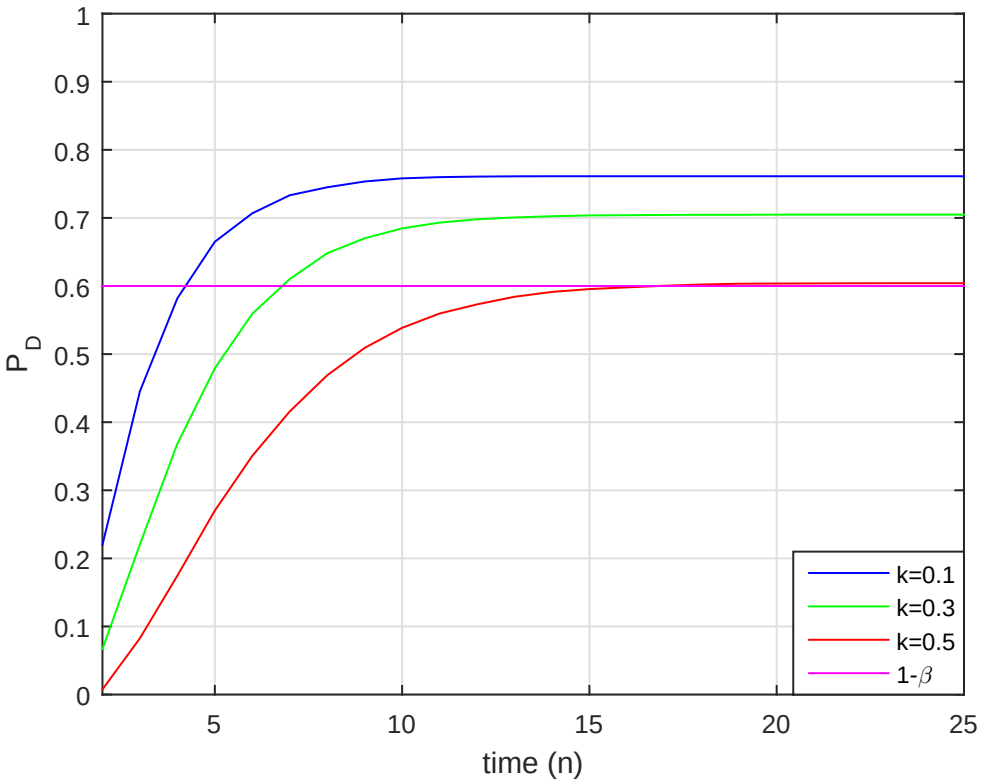


(b) P_D curves

Figure 3.5: P_{FA} and P_D curves Wald's theory in noise with data from filter

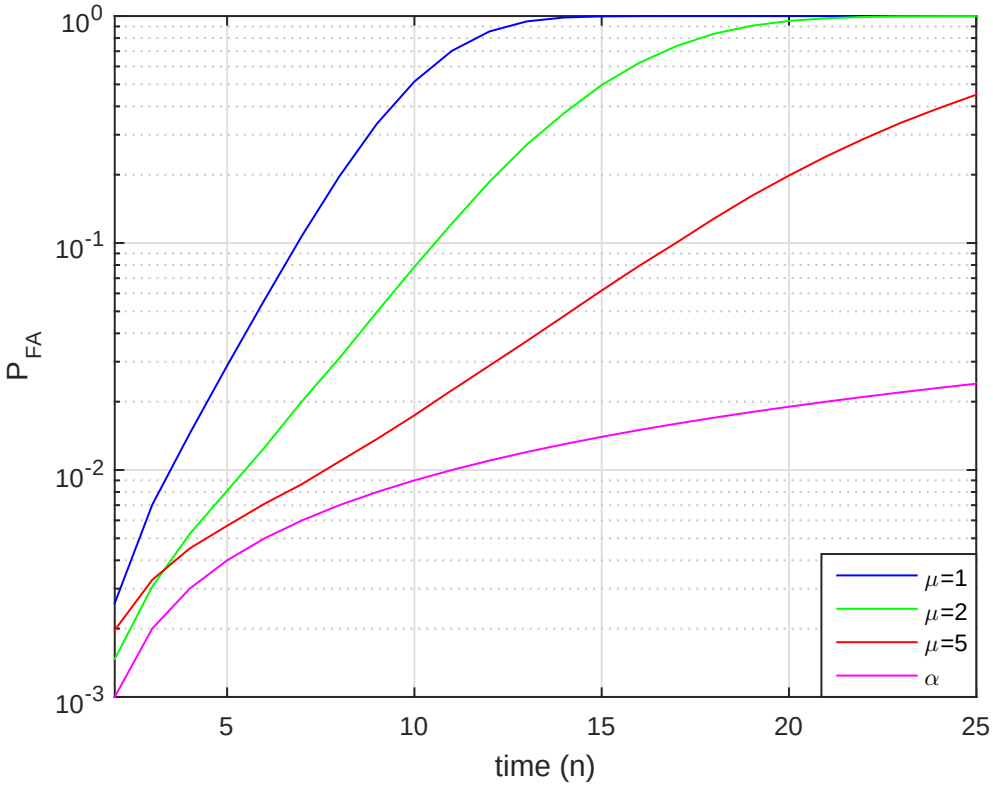


(a) P_{FA} curves

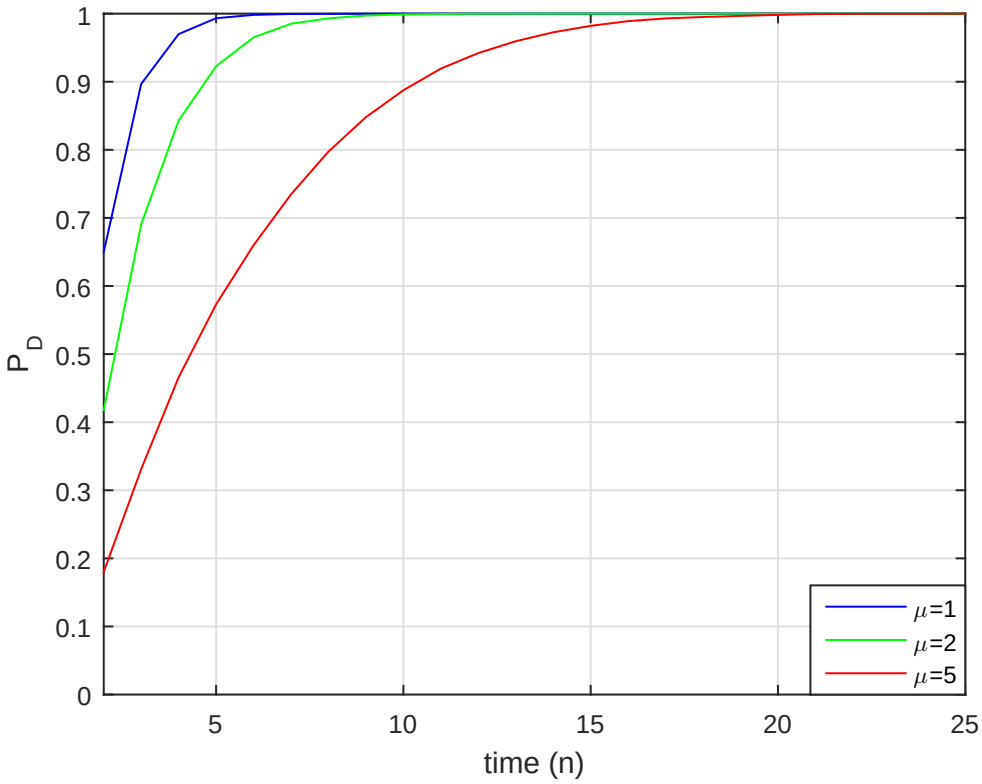


(b) P_D curves

Figure 3.6: P_{FA} and P_D curves Wald's theory in clutter with data from filter



(a) P_{FA} curves



(b) P_D curves

Figure 3.7: P_{FA} and P_D curves Neyman-Pearson theory in noise with data from filter

Chapter 4

Problems with filtered data and Solutions

The results in Sections 3.2 and 3.3 for the i.i.d. generated samples coincide with both Wald's theory and Neyman-Pearson theory. However, the results of the data acquired from the particle filter show significant differences. In this chapter, possible reasons for these differences are discussed and models are proposed to correct some wrong assumptions or to improve the performance of the detection schemes.

The most important assumption is that the samples are independent identically distributed. In Section 3.2 the samples are generated such that they are indeed i.i.d. and the results coincided with the theory. The results of the filtered data indicate that the samples might not be i.i.d. at all. Several observations are made that could explain these undesired results. In Section 4.1, these observations are discussed and in Section 4.2 models/solutions are proposed. In Section 4.3, the results will be presented for the new models. Finally in Section 4.4, a proposed implementation of the CFAR detection scheme within the TBD approach is presented.

Remark 13. In Section 3.1.1, Wald's threshold are derived with use of the assumption that the sequential test will decide eventually, i.e. $P(S < \infty) = 1$. Despite not giving proof that $P(S < \infty) = 1$ holds, by for example plotting P_{MD} together with P_D, it is verified that the test at $N = 25$ ensures

$$P(\text{Conclude } \mathcal{H}_0|\mathcal{H}_0) + P(\text{Conclude } \mathcal{H}_1|\mathcal{H}_0) = 1, \tag{4.1a}$$

$$P(\text{Conclude } \mathcal{H}_0|\mathcal{H}_1) + P(\text{Conclude } \mathcal{H}_1|\mathcal{H}_1) = 1. \tag{4.1b}$$

Hence, Wald's thresholds are derived while satisfying the condition of $P(S < \infty) = 1$.

4.1 Observations in filter

In this section, the reasons behind the differences between the results of the i.i.d. data and the results of the data from the particle filter are discussed. First, the influence of the spread of the particle cloud on the observed measurement by the cloud is discussed, as well as the influence on the RCS and velocity estimations. Also the result of the MCMC step on the tested data in Section 3.3 is discussed. Finally, some comments on the target model are given.

To start, three figures are presented that present the output for a single run of the filter in case of tracking respectively Gaussian noise, Pareto clutter and an actual target (in Gaussian noise), where all plotted particles represent the posterior estimations. The noise has mean power $\mu_n = 1$, the clutter parameters are $k = 0.3$, $s = 1$ which results in a mean power $\mu_c = 1.43$ and the target has mean power $\mu_t = 10$. In Figures 4.1a, 4.2a and 4.3a, the spread of the particle cloud over the range measurement cells is shown. The power of this measurement cell is forced to be above the startup threshold τ_{start}. The particle spread in Figures 4.1a and 4.2a is way more than in Figure 4.3a. It also seems that the spread in Figure 4.2a is a bit more compact than the spread in Figure 4.1a in the sense that the range that is covered by the cloud is less in Figure 4.2a than in Figure 4.1a. It also seems that there are less individually "living" particles in Figure 4.2a. This can be explained by the fact that the variance of clutter power samples is much higher than the variance of noise samples and hence there will be more high clutter samples that can initiate particles to track them. In Figure 4.3a, the true target position is plotted (in purple) alongside the particle cloud and corresponding estimates. The target has a relatively high SNR (10dB) and as a result the particles track the target very closely.

In Figures 4.1b, 4.1c, 4.2b, 4.2c, 4.3b and 4.3c, the RCS estimation and velocity estimation of the particles is given during the filter. The RCS is initialized based on the first power measurement, by solving the Radar equation (2.1). In (2.1), the transmitted power P_t, the Radar gain G are known and as soon as a measurement is received, the effective area A_e, the range of an reflecting object R and the reflected power P_r are estimated. Based on the parameters, the RCS can be estimated by

$$\sigma = \frac{P_r \cdot R^4 \cdot (4\pi)^2}{P_t \cdot G \cdot A_e}. \tag{4.2}$$

In Figures 4.1b and 4.2b, the RCS estimate decreases as the filter progresses. The RCS in clutter is initiated higher than the RCS in noise and also the estimation at $n = 25$ is higher in case of clutter, due to the higher variance of clutter samples and hence the higher probability of occurrence of high power samples. In Figure 4.3b, the true target RCS is plotted alongside the particle estimates. After a small decrease in the estimates during the first time steps (due to a too high initiation), the estimate of the particles cloud circles around the true value.

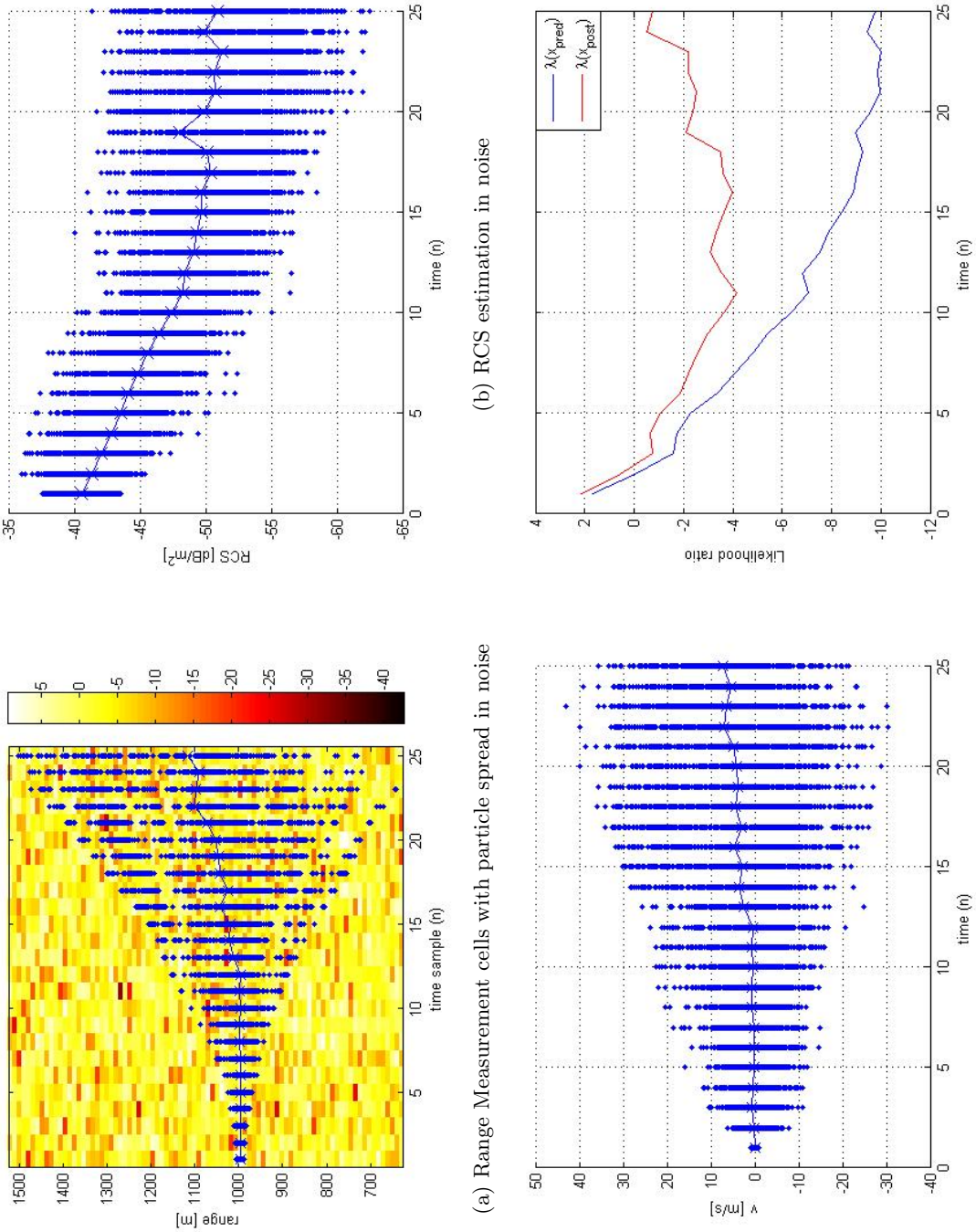
In Figures 4.1c, 4.2c and 4.3c, the estimate of the velocity of the target is plotted. The velocity is initialized around 0 and a dynamic state model is used where the velocity is estimated to be 0 as well. Hence, Figures 4.1 to 4.3 are results from perfect filters in the sense that the track position and the dynamic system model, match perfectly with the true position of the target. Again, the spread in velocity in noise and clutter is much bigger than the spread of the particles cloud that tracks the actual target. Furthermore, the spread in noise seems again to be a bit more than the spread in clutter by the same reason that high clutter samples are less rare than high noise samples.

The integrated predicted and posterior log likelihood ratios are shown in Figures 4.1d, 4.2d and 4.3d. The likelihood ratio at time n is given by the mean over all the particles, i.e.

$$\lambda(x_n) = \frac{1}{NoP} \sum_{j=1}^{NoP} \lambda(x_n^j). \tag{4.3}$$

Time integration is performed by

$$\lambda(\mathbf{x}_n) = \prod_{i=1}^{n} \lambda(x_i). \tag{4.4}$$



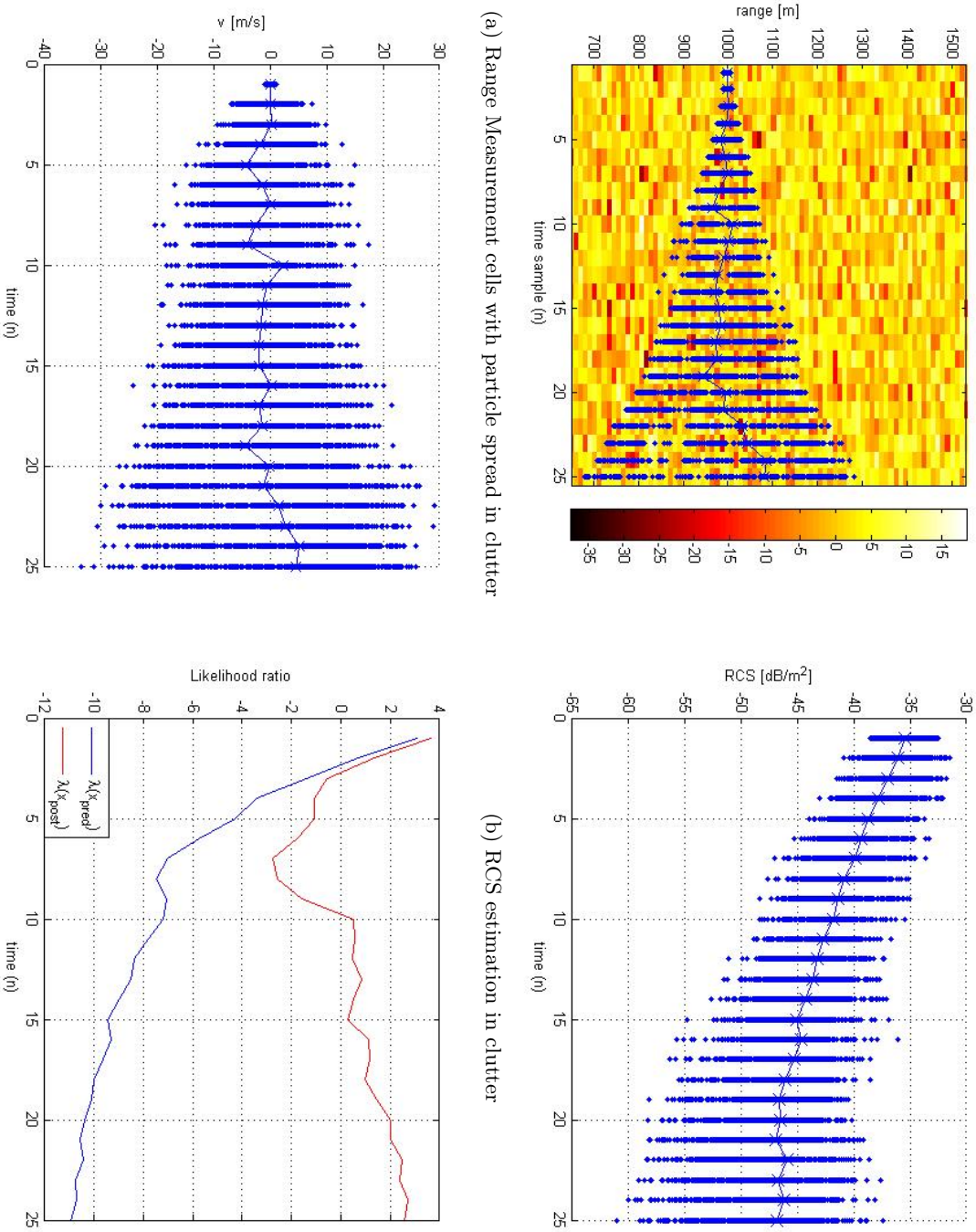
(a) Range Measurement cells with particle spread in noise

(b) RCS estimation in noise

(c) Velocity estimation in noise

(d) Likelihood ratio in noise

Figure 4.1: Filter output track in noise



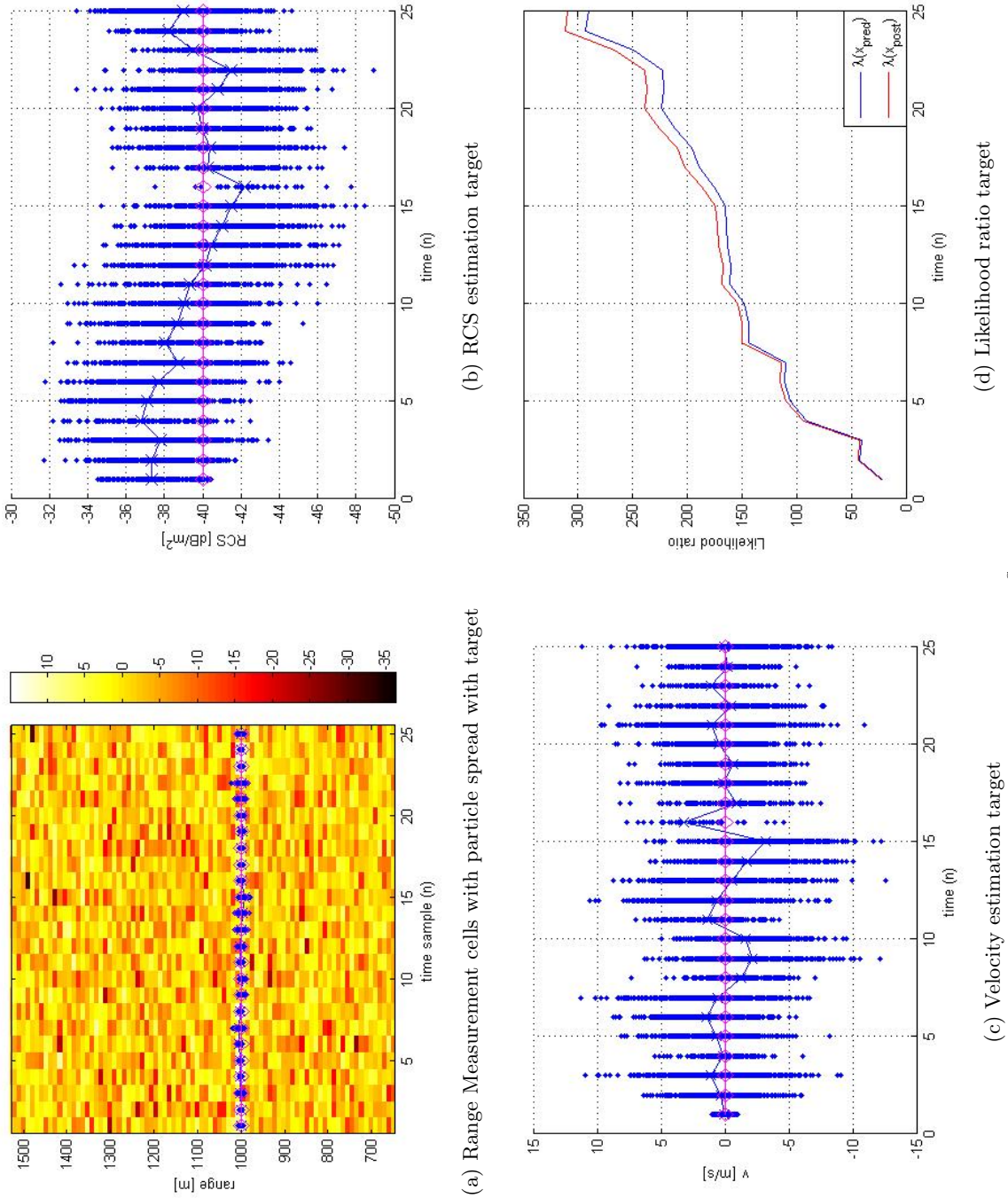
(a) Range Measurement cells with particle spread in clutter

(b) RCS estimation in clutter

(c) Velocity estimation in clutter

(d) Likelihood ratio in clutter

Figure 4.2: Filter output track in clutter



(a) Range Measurement cells with particle spread with target

(b) RCS estimation target

(c) Velocity estimation target

(d) Likelihood ratio target

Figure 4.3: Filter output track on target in noise

4.1.1 RCS fluctuations

The RCS of the particles is initialized based on solving eq.(4.2). Figures 4.1b, 4.2b and 4.3b show that the estimate of the RCS is far from constant. During the track, the estimates change based on incoming measurements. For example, when the first measurement turns out to be a fortuitous rare noise/clutter sample, the RCS estimate will drop heavily, which is the case in Figures 4.1b and 4.2b, because new measurements contain (far) less power. As a result, the estimate of the mean target power, μ_t, also changes with each incoming measurement. These fluctuations should be considered in threshold calculations. Furthermore, background estimates change during the filter as well and also these fluctuations should be considered in threshold calculations. In this research the background parameters, μ_b in case of Gaussian noise and k, s and θ in case of Pareto clutter, are known and held constant during the filtering.

4.1.2 Spreading particles

A particle cloud that tracks a possible target does not constantly have the same shape. During the prediction step in the particle filter, the particles are predicted according to a state model with a certain error. This error causes the particles to spread out at each time n. During the MCMC step the particles are judged and rescheduled based on their likelihood ratio. In summary, the higher the likelihood ratio of a particle, the more likely that particle is to be accepted. After the MCMC step, the accepted particles form the posterior particle cloud, which will be clustered towards high power measurements. When the particle cloud receives many high power measurement (probably from a target), the MCMC step ensures that the cloud tracks the target closely, because many particles will have a high likelihood ratio. This is illustrated in Figure 4.3a. However, when the cloud is actually tracking noise of clutter, not many particles will have a high likelihood ratio and after the MCMC step the cloud will not be clustered. The next prediction step causes the particles to be spread even more and this phenomenon repeats itself. Figures 4.1a and 4.2a illustrate the resulting divergent behavior of the particle cloud.

Due to the spreading particles, the cloud will cover a varying number of measurement cells during the track. In Figures 4.1a and 4.2a, the cloud only covers a few cells in the beginning, but is spread over many cells in the end. Consequently, the particle cloud contains a varying number of cell measurements during the track. Since the likelihood ratio at time n is the mean over the likelihood ratios of all particles (4.3), it seems that the input in the likelihood ratio expression is not identically distributed, because the number of measurement cells fluctuates. Let us denote the number of range measurement cells covered by the particle cloud by c and let us consider the following example for clearness.

Example 14. Let us assume for simplicity that the particle cloud during the track is equally spread over a varying number of cells, i.e. the number of cells covered by the cloud c varies, but at each time the number of particles in each cell is $\frac{NoP}{c}$. Furthermore, let us neglect the point spread function and let us only consider the measurement in the cell where the respective particle is positioned. Using these assumptions, each particle gets a likelihood function $p(z|x^j)$, where z is only the measurement cell where the respective particle is positioned. Let us now assume that at time $n = 1$, the whole particle cloud is positioned in one measurement cell. Then, the likelihood of all particles will be based on the same measurement cell. As a result, the input in the likelihood ratio expression (4.3) is practically one measurement cell. Assume that at time $n = 5$, the particle cloud is equally spread over 5 measurement cells, i.e. $c = 5$. Now the input in (4.3) contains 5 measurement cells. Hence, $\lambda(x_n)$ is based on the mean of the 5 measurement cells.

Because $\lambda(x_n)$ is the mean of all likelihood ratios of the particle cloud (eq.(4.3)), where the cloud covers a varying number of measurement cells c during the track, it seems that the resulting sample of the complete particle cloud is not identically distributed over time. This has consequences as for Wald's theory and Neyman-Pearson theory in Sections 3.1.1 and 3.1.2 is it assumed that the samples are i.i.d. The fact that this assumption might not hold causes immediate problems in applying Wald's theory, because the whole theory is based on the assumption that the samples are i.i.d. Also Neyman-Pearson theory is based on the assumption of i.i.d. samples. However, the theory is applied in a sequential detection scheme as Definition 9. The advantage of the design in Definition 9 is that the samples for different times do not have to be identically distributed as long as at each time, a proper distribution is used in the threshold calculations.

4.1.3 MCMC step

The MCMC step is an important step in the SMCMC. As explained in Sections 2.2.3 and 2.3.5, MCMC works by defining a Markov chain over the state space $\mathcal{X}$ such that the stationary distribution of the chain, $\pi(\mathcal{X})$, equals the posterior distribution $p_1(\mathcal{X}|\mathcal{Z})$ over the configuration $\mathcal{X}$, given measurements $\mathcal{Z}$. Because of the acceptance/rejection step in the MH algorithm, the posterior particle cloud is expected to be more clustered towards high measurements samples. As a direct consequence, despite the fact that both the predicted and the posterior particle cloud as defined in Section 2.2.3 approximate the same target distribution, the expected mean of the posterior likelihood ratio is bigger than the mean of the predicted likelihood ratio:

$$\mathbb{E}[\lambda_{post}(x)] = \mathbb{E}\left[\frac{1}{NoP}\sum_{j=1}^{NoP}\lambda_{post}(x^j)\right] > \frac{1}{NoP}\sum_{j=1}^{NoP}\lambda_{pred}(x^j) = \lambda_{pred}(x).$$

In practice, the posterior likelihood will always be bigger than the predicted likelihood, when the number of particles is big enough. In Figures 4.1d, 4.2d and 4.3d the differences between the time integrated (4.4) predicted and posterior likelihood ratio are illustrated. Clearly the posterior likelihood ratio is bigger than the predicted likelihood ratio in all three cases. Furthermore, it seems that the MCMC step has a relatively bigger influences when tracking noise or clutter (Figures 4.1d and 4.2d), than when tracking a target (Figure 4.3d). Also, the influence seems to depend on the shape of the particle cloud, because the effect of the MCMC is not the same for each time.

Since the posterior likelihood ratio is used in the hypothesis testing, the effect of the MCMC illustrated in Figures 4.1d and 4.2d could explain the undesired P_{FA} values in Section 3.3. In Figure 4.4, the P_{FA} curves illustrate the resulting effect of the MCMC on the false alarm probability. In this figure, the curves show the P_{FA} values for tracks in noise where the likelihood ratios are tested against Wald's thresholds (eqs.(3.4)-(3.5)), where $1-\beta$ is set to respectively 0.9 for $\mu_n = 1$, 0.7 for $\mu_n = 2$ and 0.5 for $\mu_n = 5$ and α is set to 10^{-3}. 10^5 Monte Carlo runs are performed, where the number of particles is 10^4, which is sufficient to show the effect of the MCMC. In Figure 4.4a, all P_{FA} curves fulfill Wald's P_{FA} criterion, while in Figure 4.4b all curves clearly cross the P_{FA} criterion. In Figure 4.5, the corresponding P_D curves are given. Contrary to the P_{FA} curves, the posterior values do fulfill Wald's criterion in Figure 4.5b and one predicted value in Figure 4.5a do not (the red line corresponding to $\mu = 5$). At first, this might seem confusing as the reasoning behind presenting the P_{FA} curves was to show that the predicted likelihood ratios fulfill Wald's P_{FA} criterion, while the posterior likelihood ratios do not. However, the fact that the P_D criterion is not reached for all power levels, is caused by another phenomenon, further discussed in Section 4.1.4.

Overall, it's important to know the impact of the MCMC step to obtain proper results. An accurate model is necessary to enclose the MCMC step, such that it can be integrated in the threshold calculations. A proposed model is discussed in Section 4.2.3.

4.1.4 Swerling Case

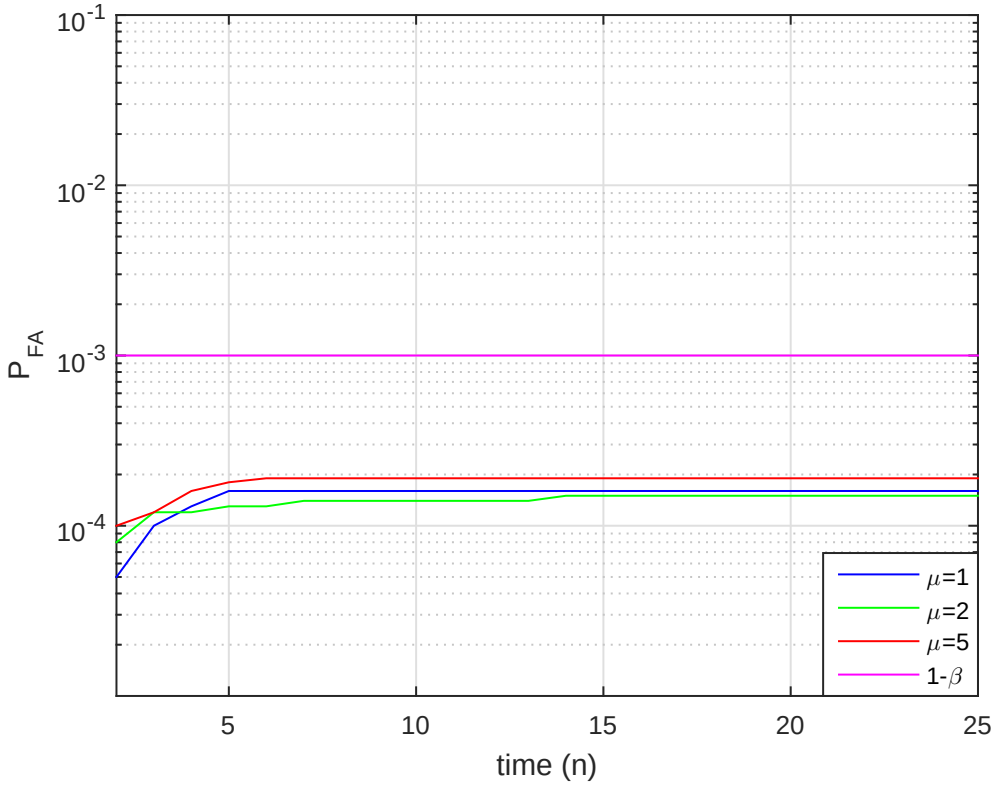
In Section 2.3.2, different Swerling models are discussed. Since a Swerling I target model is used in this theses, target power reflections are not always measured in the cells. Depending on whether or not the target reflections are measured, power samples from different distributions are measured. This immediately contradicts the assumption that the samples are identically distributed. In calculating Neyman-Pearson threshold in Algorithm 3, this is not a problem as long as the distributions p_0 and p_1 are well estimated. However, in applying Wald's theory this contradicts the most important assumptions of i.i.d. samples. Furthermore, a more practical problem could arise: Wald's thresholds do not depend on underlying distributions, but they do depend on the criteria $P_{FA} \leq \alpha$ and $P_D \geq 1 - \beta$. When the measurements from the targets behave according to a Swerling I model, it obviously takes longer to detect the target. Hence, the expected stopping time $\mathbb{E}[S]$ will increase. Theoretically, this is not a big problem as long as decision can be made within a finite time horizon. However, the particle cloud will spread out when it does not get high power measurements from a target as illustrated for example in Figures 4.1a and 4.2a. This makes it even harder to detect a possible target. As a result, the P_D for targets with a low SNR might not fulfill the P_D criterion.

In Figure 4.6 the differences in P_D between a Swerling I and a Swerling 0 target are shown, tested against Wald's thresholds. The target has a mean power of $\mu_t = 10$, and the desired criteria are

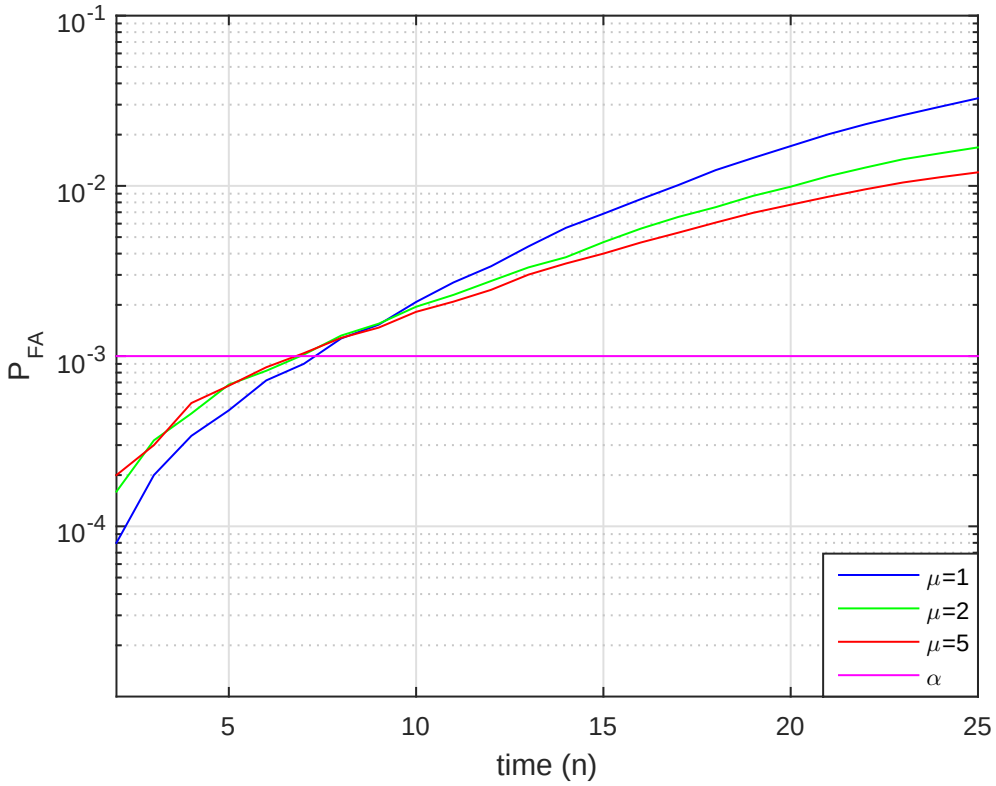
$$P_{FA} \leq \alpha = 10^{-3} \qquad P_D \geq 1 - \beta = 0.6.$$

In Figure 4.6b, the P_D criterion are quickly reached for all three shape parameters k, which suggests that also the P_D for targets with less power will fulfill the criterion. Figure 4.6a shows the Swerling I model used during this thesis. For $k = 0.5$ the P_D criterion is barely reached. This suggests that for targets with lower Signal to Clutter Ratio (SCR) , the P_D criterion can not longer be reached. Hence, it seems that by using a Swerling I target model, Wald's theory can no longer be applied, because the assumption of i.i.d. samples is not fulfilled which results in a P_D criterion that can not be reached.

Remark 15. Although it is discussed that Wald theory can not be applied properly when using a Swerling I model, a Swerling 0 target model is often not realistic in practice. In some fortunate scenarios a target can be modeled with a Swerling 0, for example when transmitting towards the nose of a missile in clear weather. However, because the application of this thesis will be primarily tracking small targets in (heavy) sea clutter, a Swerling 0 target model is not representative. Therefore, a Swerling I target model will still be used and no adjustments to this model will be discussed in Section 4.2.

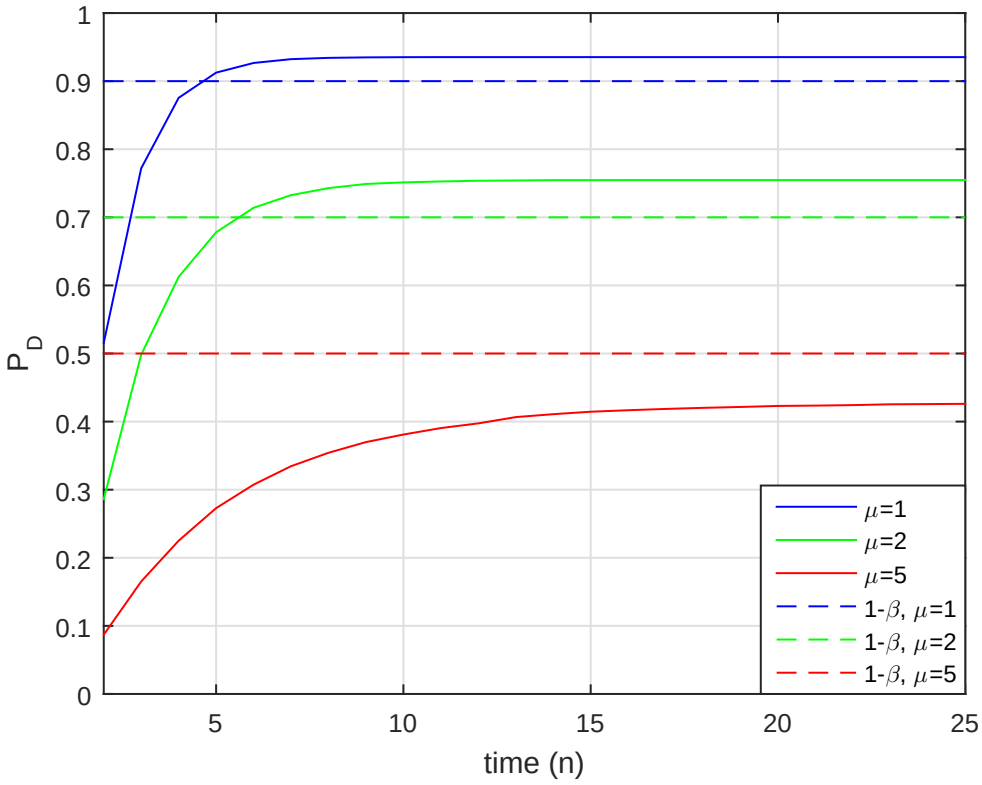


(a) Predicted likelihood ratio

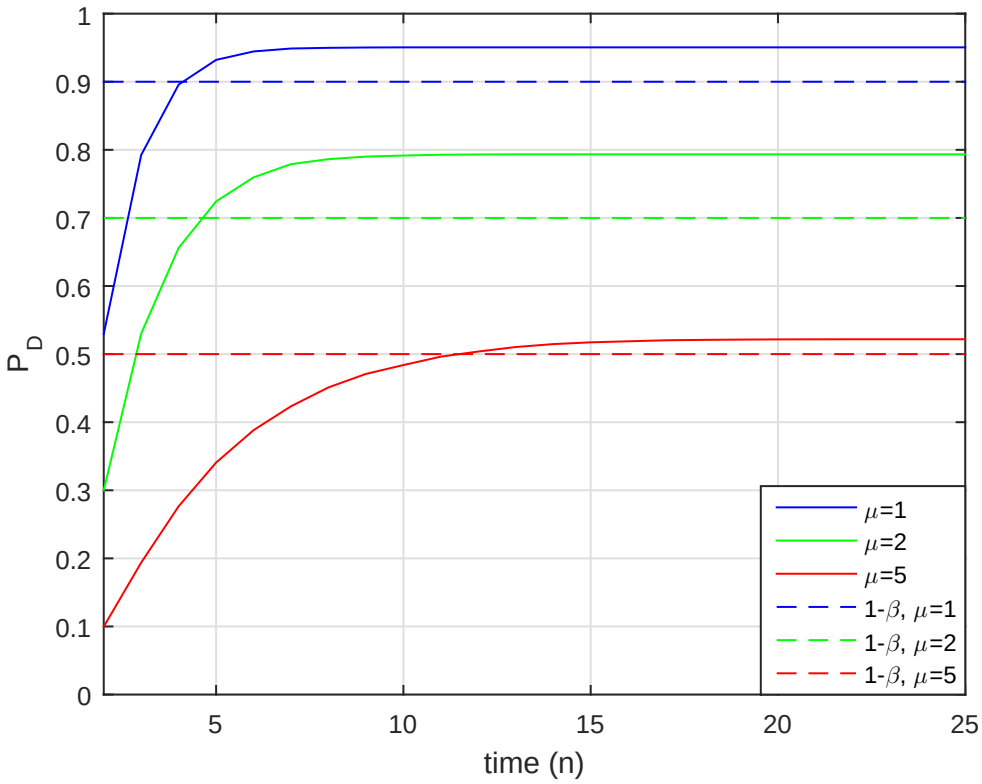


(b) Posterior likelihood ratio

Figure 4.4: Difference P_{FA} curves between predicted and posterior likelihood ratios

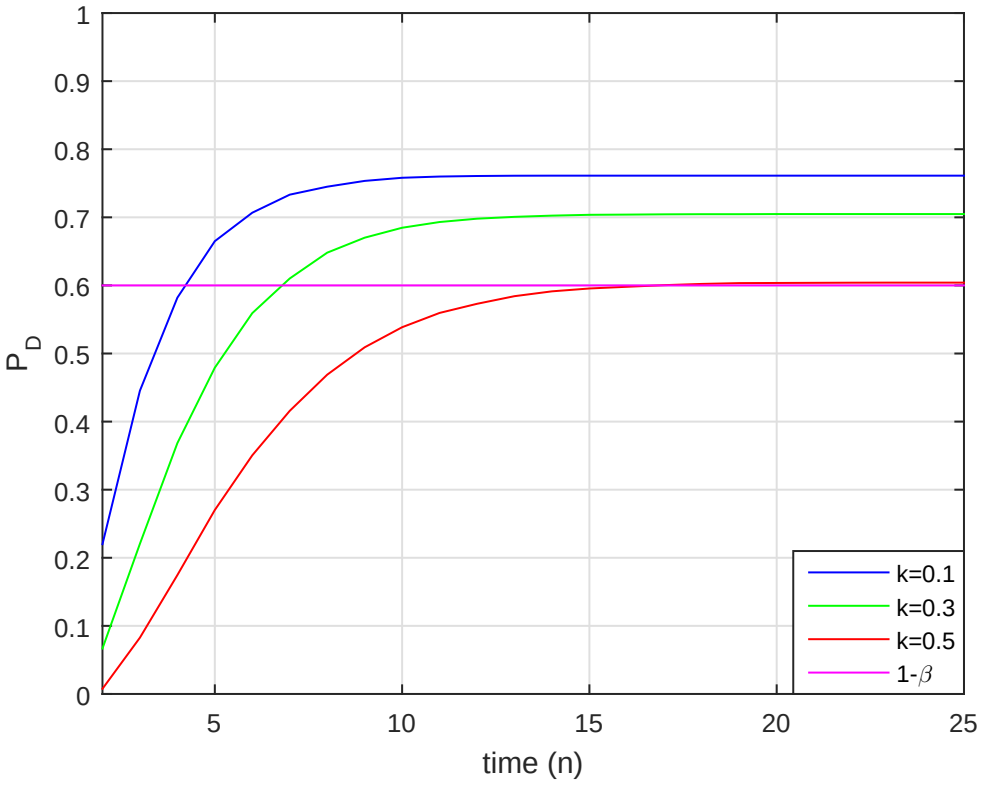


(a) Predicted likelihood ratio

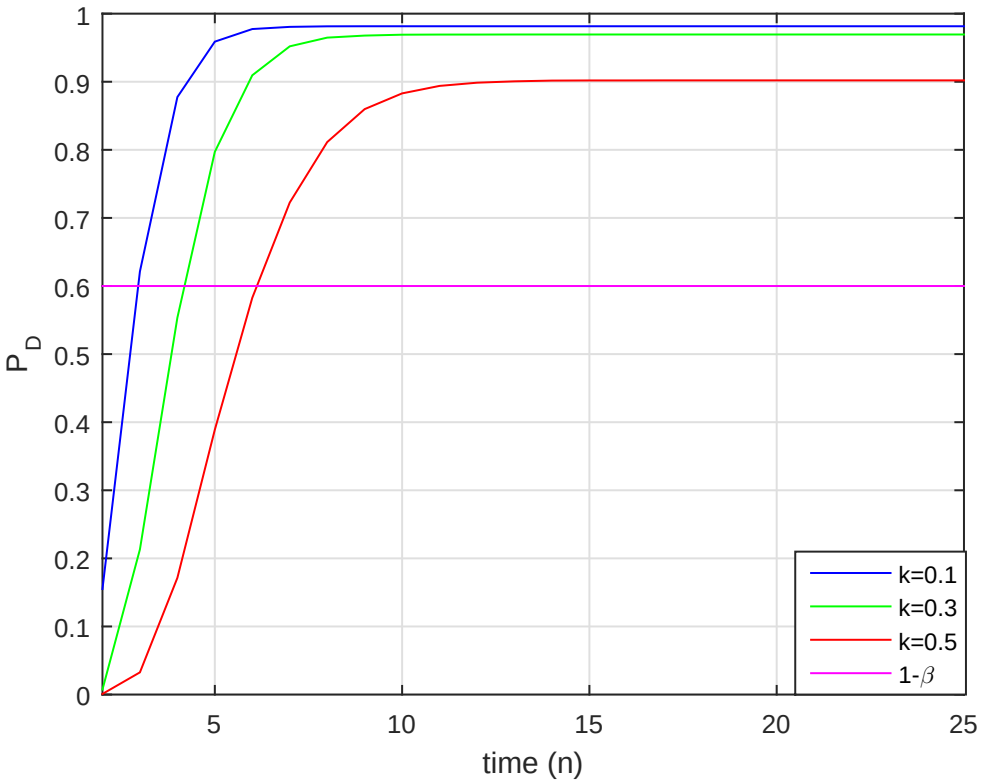


(b) Posterior likelihood ratio

Figure 4.5: Difference P_D curves between predicted and posterior likelihood ratios



(a) Swerling I target



(b) Swerling 0 target

Figure 4.6: Difference in P_D between Swerling I and Swerling 0 target.

4.2 Models as solution

In this section, models and methods will be proposed to improve the performance of detection schemes. First, the fluctuations in the RCS estimations are discussed. Second, a method is proposed to model the spread of the particle cloud. Finally, a method to model the MCMC step is derived. From Remark 15, the Swerling models will not be discussed in this section.

4.2.1 Handling the RCS fluctuations

Section 4.1.1 explained that parameter estimates of target and background change during the filtering. Despite simplifying the filtering by fixing the background power, the fluctuations in target power estimates heavily influence the P_{FA} curves when the likelihood ratio is tested against a threshold that is calculated without considering these fluctuations. In Sections 3.2 and 3.3 the thresholds are calculated according to a priori knowledge of the parameters. When the parameters are constant over time, the results of Wald's theory and Neyman-Pearson theory coincide with theoretically expected results as shown in Section 3.2. Unfortunately when the parameter estimates change over time, as in Section 3.3, the P_{FA} curves show undesired behavior as the P_{FA} values do not satisfy the set significance level. Note that even if the true parameter values are known a priori and the threshold is properly calculated according to this known quantities, using this threshold with filtered data will not give proper results.

To overcome this problem, the filtered parameter estimates should be used in the threshold calculations. These parameter estimates could be stored for all time to calculate a proper threshold afterwards, but ideally the threshold is updated iteratively according to new filtered estimates per time such that the most recent estimates are directly used.

4.2.2 Model spreading particles

In Section 4.1, it is discussed that the input sample of the particle cloud in $\lambda(x_n)$ seems to be not identically distributed due to the spreading particle cloud. Let us denote this input sample in what follows by the total measurement sample. The fact that the total measurement sample is not identically distributed causes problems in applying the hypothesis testing theories. It makes it impossible to guarantee Wald's criteria and it also makes adjustments necessary in applying Neyman-Pearson Theory.

The use of a grid based particle filter could be a solution. In these grid based filters, the position of the particles is forced to be constant and consequently the spread of the particle cloud over the cells is constant. For further details, let us refer to [8, 34, 35]. Two naive ways of constructing a filter that gives i.i.d. total measurement samples are

- The use of an infinite number of particles
- Predict the particle cloud according to the state model without process noise (error) and skip the MCMC step.

The first solution does not only solve the problem of the total measurement samples being not identically distributed, but it solves many more irregularities in the particle filter by covering the whole measurement space at any time by an infinite number of particles. However, it will be computationally impossible to process an infinite number of particles. Hence in practice this idea is not useful. The second solution could be used in practice. However, the P_D will

be horrible when the dynamic target model is not absolutely accurate, because in this case the target will evolve different than predicted and the particles will eventually lose sight of the target.

Since these two naive solutions will not be used in practice, it is important to look at the particle cloud and to model this in a way that is useful, at least for applying Neyman-Pearson theory. Following Example 14, the distribution of a sample x_n in $\lambda(x_n)$ of a particle cloud spread over one measurement cell is different from the distribution of a sample x_n in $\lambda(x_n)$ of a particle cloud spread over a number (c) cells, with $c > 1$. For now it is assumed that the particles are equally spread over the measurement cells, as in Example 14. As a result, x_n in $\lambda(x_n)$ is now distributed according to the sample mean of c samples. In case of Gaussian noise it will be shown how to integrate the spread of the particle cloud in the threshold calculations. In case of Pareto clutter only an expression is derived for the probability density function of the sample mean.

Gaussian Noise

In case of Gaussian noise, the power of noise samples is exponentially distributed. The sample mean of c exponential distributed random variables with mean μ is known to be Gamma distributed with parameters $(k = c, \theta = \mu/c)$ [31]. The probability density function of the sample mean of c exponential distributed r.v.'s is given by

$$\begin{aligned} p(x,k,\theta) &= \frac{1}{\theta^k \Gamma(k)} x^{k-1} \mathrm{e}^{-x/\theta}, \qquad k,\theta > 0, \quad x \in \mathbb{R}^+ \\ &= \frac{c^c}{\mu^c \Gamma(c)} x^{c-1} \mathrm{e}^{-xc/\mu}, \qquad c,\mu > 0, \quad x \in \mathbb{R}^+. \end{aligned} \tag{4.5}$$

Ultimately, the distribution of the sample mean should be used in the threshold calculation, given by (3.18). However, it seems impossible to integrate the distribution of the sample mean (4.5) together with the distribution of the time integrated sample (3.20) simultaneously. Therefore, the threshold calculations for the integrated time sample in Example 10 are changed using the following observation: the threshold found by eqs.(3.21)-(3.22) can also be found with the use of the pdf of the sample mean of n of i.i.d. exponential distributed r.v.'s, instead of the pdf of the sum of n i.i.d. exponential distributed r.v.'s. In exchange, also the critical region has to be changed. The critical region becomes:

$$\begin{aligned} \{\bar{\mathbf{x}}_n : \lambda(\bar{\mathbf{x}}_n) > \tau_n\} &= \left\{ \bar{\mathbf{x}}_n : \frac{p(\bar{\mathbf{x}}_n|\mathcal{H}_1)}{p(\bar{\mathbf{x}}_n|\mathcal{H}_0)} > \tau_n \right\} \\ &= \left\{ \bar{\mathbf{x}}_n : \left(\frac{\mu_b}{\mu_b + \mu_t} \right)^n \mathrm{e}^{\mu_t \left(\frac{1}{n} \sum_{i=1}^n x_i \right) n/(\mu_b(\mu_b+\mu_t))} > \tau_n \right\} \\ &= \left\{ \bar{\mathbf{x}}_n : \frac{1}{n} \sum_{i=1}^n x_i > \left(\ln(\tau_n) - n \ln \left(\frac{\mu_b}{\mu_b + \mu_t} \right) \right) \frac{\mu_b(\mu_b + \mu_t)}{n\mu_t} = \tilde{\tau}_n \right\}, \end{aligned} \tag{4.6}$$

where $\bar{\mathbf{x}}_n$ is the mean of the elements in the vector $\mathbf{x}_n$. Now, the pdf of the sample mean of n i.i.d. exponential random variables (4.5) is used to find $\tilde{\tau}_n$ by solving

$$
\begin{aligned}
P_{FA} &= \int_{\tilde{\tau}_n}^{\infty} \frac{n^n}{\Gamma(n)\mu_b^n} x^{n-1} \mathrm{e}^{-xn/\mu_b} \mathrm{d}x = \alpha \\
&= -\left[\frac{n^{n-1}}{\Gamma(n)\mu_b^{n-1}} x^{n-1} \mathrm{e}^{-xn/\mu_b}\right]_{\tilde{\tau}_n}^{\infty} + \int_{\tilde{\tau}_n}^{\infty} \frac{n^{n-1}}{\Gamma(n-1)\mu_b^{n-1}} x^{n-2} \mathrm{e}^{-xn/\mu_b} \mathrm{d}x \\
&= \frac{n^{n-1}\mathrm{e}^{-n\tilde{\tau}_n/\mu_b}}{\Gamma(n)\mu_b^{n-1}} - \left[\frac{n^{n-2}}{\Gamma(n-1)\mu_b^{n-2}} x^{n-2} \mathrm{e}^{-xn/\mu_b}\right]_{\tilde{\tau}_n}^{\infty} + \int_{\tilde{\tau}_n}^{\infty} \frac{n^{n-2}}{\Gamma(n-2)\mu_b^{n-2}} x^{n-3} \mathrm{e}^{-xn/\mu_b} \mathrm{d}x
\end{aligned}
$$

After performing these calculations n times, P_{FA} will be given by

$$
P_{FA} = \sum_{i=1}^{n} \left(\frac{(n\tilde{\tau}_n/\mu_b)^{i-1}}{\Gamma(i)}\right) \mathrm{e}^{-n\tilde{\tau}_n/\mu_b}. \tag{4.7}
$$

Once $\tilde{\tau}_n$ is derived, τ_n can be found from

$$
\tau_n = \left(\frac{\mu_b}{\mu_b + \mu_t}\right)^n \mathrm{e}^{n\tilde{\tau}_n\mu_t/(\mu_b(\mu_b+\mu_t))}.
$$

Now the threshold calculations for the integrated time samples are rewritten as a sample mean calculation, the integrated time sample and the sample mean of the c measurement cells can easily be integrated in the threshold calculations simultaneously. Suppose that the number of measurement cells that is covered at time n is c_n, then $\tilde{\tau}_n$ can be calculated by solving

$$
P_{FA} = \int_{\tilde{\tau}_n}^{\infty} \frac{\tilde{c_n}^{\tilde{c_n}}}{\Gamma(\tilde{c_n})\mu_b^{\tilde{c_n}}} x^{\tilde{c_n}-1} \mathrm{e}^{-x\tilde{c_n}/\mu_b} \mathrm{d}x = \sum_{i=1}^{\tilde{c_n}} \left(\frac{(\tilde{c_n}\tilde{\tau}_n/\mu_b)^{i-1}}{\Gamma(i)}\right) \mathrm{e}^{-\tilde{c_n}\tilde{\tau}_n/\mu_b}, \tag{4.8}
$$

where $\tilde{c_n}$ is given by

$$
\tilde{c_n} = \sum_{i=1}^{n} c_i. \tag{4.9}
$$

Finally, τ_n is given by

$$
\tau_n = \left(\frac{\mu_b}{\mu_b + \mu_t}\right)^{\tilde{c_n}} \mathrm{e}^{\tilde{c_n}\tilde{\tau}_n\mu_t/(\mu_b(\mu_b+\mu_t))}. \tag{4.10}
$$

Remark 16. In these threshold calculations, it is assumed that the particle cloud is equally spread over the measurement cells. Also the influence of the point spread function on the measurement samples is neglected. Hence in practical application, the P_{FA} values using eqs. (4.8)-(4.10) might not satisfy the exact level of significance α, but the performance in P_{FA} should at least improve compared to the threshold calculation in Example 10.

Remark 17. It is not possible to predict the spread of the particle cloud beforehand. Hence for accurate threshold calculations, it is important to update the information about the particle cloud iteratively.

Generalized Pareto Clutter

In case of generalized Pareto clutter, the pdf of the power is given by

$$p_X(x) = \begin{cases} \frac{1}{s}\left(1 + k\left(\frac{x-\theta}{s}\right)\right)^{-1-1/k}, & x \geq \theta, k > 0, s > 0, \theta \in \mathbb{R}, \\ \frac{1}{s}\mathrm{e}^{-\left(\frac{x-\theta}{s}\right)}, & k = 0, \end{cases} \tag{4.11}$$

with θ location parameter, s scale parameter, k shape parameter respectively. The cdf of the power is given by

$$F_X(x) = \begin{cases} 1 - \left[1 + k\left(\frac{x-\theta}{s}\right)\right]^{-1-1/k}, & k > 0, \\ 1 - \mathrm{e}^{-\left(\frac{x-\theta}{s}\right)}, & k = 0, \end{cases} \tag{4.12}$$

The distribution of the sample mean of c Pareto distributed r.v.'s is not clear at first sight. The method of characteristic functions, is applied to find an expression for this distribution of the sample mean. First, change of variables is applied to find the weighted distribution of a single measurement. $y = x/c \rightarrow x = c \cdot y$ and the Jacobian is $\frac{\mathrm{d}x}{\mathrm{d}y} = c$. Hence, the pdf of a weighted sample is given by

$$p_Y(y) = \begin{cases} \frac{c}{s}\left(1 + k\left(\frac{cy-\theta}{s}\right)\right)^{-1-1/k}, & y \geq \theta, k > 0, s > 0, \theta \in \mathbb{R}, \\ \frac{1}{s}\mathrm{e}^{-\left(\frac{cy-\theta}{s}\right)}, & k = 0, \end{cases} \tag{4.13}$$

The characteristic function $\varphi_Y(t)$ for the weighted sample is given by (see Appendix C)

$$\varphi_y(t) = \begin{cases} \mathrm{e}^{(it\theta/c)} \sum_{m=0}^{\infty} \frac{(its/c)^m}{\prod_{b=0}^{m}(1-bk)}, & k > 0, \\ \mathrm{e}^{(it\theta/c)} \sum_{m=0}^{\infty} (its/c)^m, & k = 0. \end{cases} \tag{4.14}$$

The Cauchy product [36] is applied to find

$$\varphi_{\bar{Y}}(t) = \varphi_{\sum_\ell} y_\ell = \prod_\ell \varphi_Y(y_\ell)$$

In Appendix C it is derived that $\varphi_{\bar{Y}}(t)$ is given by

$$\varphi_{\bar{Y}}(t) = \mathrm{e}^{it\theta} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{m_1} \cdots \sum_{m_c=0}^{m_{c-1}} \frac{(its/c)^{m_1}}{\prod_{b_1=0}^{m_c}(1-b_1k)\prod_{b_2=0}^{m_{c-1}-m_c}(1-b_2k)\ldots\prod_{b_c=0}^{m_1-m_2}(1-b_ck)}.$$

The pdf, $p_{\bar{Y}}(y)$, is found by using the inverse Fourier transform [37] of $\varphi_{\bar{Y}}(t)$:

$$p_{\bar{Y}}(y) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iny}\varphi_{\bar{Y}}(n)\mathrm{d}n$$

In Appendix C the final expression for the pdf of the sample mean of c i.i.d. Pareto r.v.'s $p_{\bar{Y}}(y)$ is found to be

$$p_{\bar{Y}}(y) = \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{m_1}\frac{\left[(-1)^{m_1+1} + (-1)^{m_1+m_2}(2\pi i(\theta-y))^{m_2}\mathrm{e}^{2\pi i(\theta-y)}\right](\bigstar)}{2\pi i(\theta-y)^{m_1+1}}, \tag{4.15}$$

with

$$(\bigstar) = \sum_{m_3=0}^{m_2}\sum_{m_4=0}^{m_3}\cdots\sum_{m_c=0}^{m_c-1}\frac{(s/c)^{m_1}}{\prod_{b_1=0}^{m_c}(1-b_1k)\prod_{b_2=0}^{m_{c-1}-m_c}(1-b_2k)\ldots\prod_{b_c=0}^{m_1-m_2}(1-b_ck)}.$$

Although this expression for the sample mean of c i.i.d. Pareto distributed r.v.'s is analytically correct, this expression is impossible to use in threshold calculations for now. The Lindeberg-Levý Central Limit Theorem (CLT) [38] states that the series converges in distribution (for $k < 0.5$):

Theorem 18 (Lindeberg-Levý Central Limit Theorem)**.**
Suppose $\{x_1, x_2, \ldots, x_n\}$ is a sequence of i.i.d. random variables with $\mathbb{E}[x_i] = \mu$ and $Var[x_i] = \sigma^2 < \infty$. Then as n approaches infinity, the random variables $\sqrt{n}\,(S_n - \mu)$, with $S_n = \frac{x_1+x_2+\ldots+x_n}{n}$ converge in distribution to a normal distribution $\mathcal{N}(0,\sigma^2)$:

$$\sqrt{(}n)\left(\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right) - \mu\right) \xrightarrow{d} \mathcal{N}(0,\sigma^2).$$

Remark 19. Although it is clear from the CLT that the series will converge in distribution for $k < 0.5$, the rate of convergence should be investigated first to see how many iterations are needed to approximate the distribution well enough. When this number of iterations is too large, expression (4.15) might not be useful in sequential threshold calculations because of the computational time it will take.

4.2.3 Proposed model MCMC step

In Section 4.1.3 it is explained that the MCMC step has a huge impact on the P_{FA} and P_D. To control the P_{FA} values, a good insight in the MCMC step is necessary. In this section the details of the MCMC step are discussed and by making some assumptions, a method is derived to model the impact of the MCMC step to the posterior likelihood ratio.

As discussed in Section 2.2.2, the particles only approximate the distribution under hypothesis $\mathcal{H}_1$, which is the distribution under the assumption that a target is present. In Section 2.2.3 it is discussed that the Metropolis-Hastings (MH) algorithm is used to perform the MCMC step. The MH algorithm is summarized in Algorithm 3.

Now the working of the algorithm is used to model the influence of the MCMC step on the posterior likelihood ratio. Suppose a set of predicted likelihoods, $\{p(z|x^j_{pred})\}_{j=1}^{NoP}$, is available. For simplicity it is assumed that the likelihood functions are calculated conditioned on a state vector x^j_{pred} with dimension 1. Suppose furthermore that $I = [0, b]$ is a closed interval contained in $[0, \infty)$, where b is the first integer greater than the highest (scalar) state value x^j_{pred}, in our case range value, i.e.

$$b = \lceil \max\{x^j_{pred}\}_{j=1}^{NoP} \rceil. \tag{4.16}$$

Using this value for b, the whole domain of $p(z|x^j_{pred})$ is contained in I. Now I is divided into S parts (segments) of equal length. Let us denote segment i by s^i and let us define the value of segment i, denoted by S^i, as

$$S^i = \frac{((i-1)+i)}{2} = \frac{2i-1}{2}, \quad i = 1, 2, \ldots, S$$

where the counting of the segments starts at the origin. Note that the value of the segment is monotonically increasing as the number of the segment is increasing. Instead of using the set of likelihoods $\{p(z|x^j_{pred})\}_{j=1}^{NoP}$, the set of segment values $\{S^i\}_{i=1}^{S}$ can now be used in the acceptance/rejection step in MH algorithm, which allows us to the derive a model for the MCMC step, independent of the particles. Similar to the set of particles with corresponding likelihood ratios, the set of segments as input and output of the MCMC are referred to as the set of predicted and posterior segments, denoted respectively by $\{s^i_{pred}\}_{i=1}^{S}$ and $\{s^i_{post}\}_{i=1}^{S}$. The MH algorithm with input the set of predicted segments with corresponding values, is summarized in Algorithm 5.
The set of predicted segments, approximates the target distribution as follows:

$$p(x|z) \approx \sum_{j=1}^{NoP} \delta_{x^j_{pred}}(x_{pred}) \approx \sum_{i=1}^{S} w(s^i_{pred}) \delta_{S^i}(S), \tag{4.17}$$

where

$$w(s^i_{pred}) = \frac{\tilde{w}(s^i_{pred})}{\sum_{i=1}^{S} \tilde{w}(s^i_{pred})}, \tag{4.18}$$

where

$$\tilde{w}(s^i) = \left|\{j | s^{i-1} \leq \{x^j\}_{j=1}^{NoP} < s^i\}\right|, i = 1, 2, \ldots, S, \tag{4.19}$$

where $|\cdot|$ is the Cardinality of the set. The model for the MCMC has to be derived to find the approximation for the target distribution of the posterior set of segments.

Algorithm 5: Metropolis-Hastings algorithm with segments

Input :$\{s^i_{pred}, S^i\}_{i=1}^S$, number of segments S, burn in period B
Output :$\{s^i_{post}\}_{i=1}^S$

```
for m:=1 to S+B do
    if m=1 then
        Randomly select segment s^i_pred out {s^i_pred}_{i=1}^S;
        Store s^i_pred: s^m_post = s^i_pred,   S^m = S^i;
    else
        Randomly draw proposed segment s^{i*}_pred out {s^i_pred}_{i=1}^S with corresponding value
        S^{i*};
        Calculate acceptance ratio: α = min(1, S^{i*}/S^i);
        Draw u ~ U[0,1];
        if α ≥ u then
            Accept s^{i*}_pred: s^m_post = s^{i*}_pred,   S^m = S^{i*};
        else
            Reject s^{i*}_pred: s^m_post = s^m_post,   S^m = S^m;
        end
    end
end
Discard first B samples, store other S segments: {s^m_post}_{m=B+1}^{S+B} → {s^m_post}_{m=1}^{NoP}
```

Since the following things are known,

- The value of segment S^i
- The acceptance probability of a proposed segment (which is equal to the acceptance ratio α, since $u \sim \mathcal{U}[0,1]$)
- Each segment is drawn with probability $\frac{1}{S}$,

the idea is to set up a transition matrix P over all segments, where element $P_{i,j}$ represents the probability to go from segment i to segment j and where the probability vector π represents the stationary distribution of the segments. Assuming that interval I is divided in a sufficiently large number of segments and that Algorithm 5, converges to the stationary distribution within the first B burn in iterations, the stationary state vector π will represent the resulting model for the MCMC step. Vector π can be found by solving

$$\pi P = \pi.$$

In P, all transition probabilities $P_{i,j\neq i}$ can be filled in directly from the acceptance probabilities. The diagonal elements $P_{i,i}$ are given by the sum of probabilities of rejecting segment s^{j*} in favor of segment $s^{(i)}$ multiplied by the probability of drawing proposal segment s^{j*}, i.e.

$$P_{i,i} = \sum_{(j\leq i)} (1 - P_{i,j}) P(s^{j*}),$$

where $P(s^j)$ is the probability of drawing proposal segment s^{j*} (equal to $\frac{1}{S}$). Making use of the fact that

$$\sum_j P_{i,j} = 1,$$

these probabilities can be found easily by

$$P_{i,i} = 1 - \sum_{\forall j \neq i} P_{i,j}. \tag{4.20}$$

For illustration, Example 20 is used to show how to determine the transition probabilities.

Example 20. In this example, $I = [0, b]$ with $b = S = 10$. Now, all probabilities $P_{i,j>i}$ will be

$$P_{i,j>i} = P(\alpha > u) \cdot P(s^{j*}) = P(1 \geq u) \cdot P(s^{j*}) = 1 \cdot \frac{1}{10} = \frac{1}{10}.$$

All probabilities $P_{i,j<i}$ will be

$$P_{i,j<i} = P(\alpha > u) \cdot P(s^{j*}) = P\left(\frac{2j-1}{2i-1} \geq u\right) \cdot P(s^{j*}) = \frac{2j-1}{10(2i-1)}.$$

So $P_{5,2}$, the probability of going from segment 5 to segment 2 is

$$P_{5,2} = \frac{2 \cdot 2 - 1}{10(2 \cdot 5 - 1)} = \frac{3}{90} = \frac{1}{30}.$$

The diagonal elements $P_{i,i}$ are calculated with (4.20). For example $P_{1,1}$ is given by

$$P_{1,1} = 1 - \sum_{j=2}^{10} (P_{1,j}) \cdot P(s^{j*}) = 1 - \frac{9}{10} = \frac{1}{10}$$

and $P_{10,10}$ is given by

$$P_{10,10} = 1 - \sum_{j=1}^{9} (P_{10,j}) \cdot P(s^{j*}) = 1 - \sum_{j=1}^{9} P\left(\frac{2j-1}{10(2 \cdot 10 - 1)} \geq u\right) \cdot P(s^{j*}) = 1 - \frac{\sum_{j=1}^{9} 2j-1}{190} = \frac{109}{190}.$$

Eventually, matrix P will be given by

$$P = \begin{bmatrix}
1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\
1/30 & 1/6 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\
1/50 & 3/50 & 11/50 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\
1/70 & 3/70 & 1/14 & 19/70 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\
1/90 & 1/30 & 1/18 & 7/90 & 29/90 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\
1/110 & 3/110 & 1/22 & 7/110 & 9/110 & 41/110 & 1/10 & 1/10 & 1/10 & 1/10 \\
1/130 & 3/130 & 1/26 & 7/130 & 9/130 & 11/130 & 11/26 & 1/10 & 1/10 & 1/10 \\
1/150 & 1/50 & 1/30 & 7/150 & 3/50 & 11/150 & 13/150 & 71/150 & 1/10 & 1/10 \\
1/170 & 3/170 & 1/34 & 7/170 & 9/170 & 11/170 & 13/170 & 3/34 & 89/170 & 1/10 \\
1/190 & 3/190 & 1/38 & 7/190 & 9/190 & 11/190 & 13/190 & 3/38 & 17/190 & 109/190
\end{bmatrix}$$

Now, the stationary state π can be found by solving

$$\pi P = \pi$$

$$\pi P - \pi = 0$$

$$\pi(P - I) = 0$$

$$\pi \in \mathcal{N}ull(P - I),$$

where $\mathcal{N}ull(P - I)$ is the Null space of matrix $(P - I)$. Since π represents a density it is scaled:

$$\pi = \frac{\pi}{||\pi||_1}. \tag{4.21}$$

For Example 20, π is given by

$$\pi = \begin{bmatrix} 0.01 & 0.03 & 0.05 & 0.07 & 0.09 & 0.11 & 0.13 & 0.15 & 0.17 & 0.19 \end{bmatrix}$$

In Figure 4.7, $\pi(i)$ is plotted against i, where i means the ith element in the vector π. The figure shows the model for the MCMC step graphically. Extrapolating towards the origin, a good analytical model for the MCMC step seems eq.(4.21), where the elements of π increase linearly with the number of elements. As a result, the posterior set of segments will approximate the target distribution by:

$$p(x|z) \approx \sum_{i=1}^{S} w(s^i_{pred})\delta_{S^i}(S)\pi(i), \tag{4.22}$$

where $w(s^i_{pred})$ is given by eq.(4.18).
In Figure 4.8, three different predicted probability density distributions are given. Figure 4.9 shows the corresponding posterior densities when the model from eq.(4.21) is used. Let us refer to Appendix D for proof by simulations to show the correctness of model (4.21).
Following Algorithm 4 in Section 2.3.5, instead of the likelihoods of the particles, the MCMC step is performed on the likelihood ratio of the particles. Note that the scaling by p_0, the distribution under the hypothesis that no target is present, does not influence the model for the MCMC. The scaling by p_0 can be seen as a factorization of the length of the interval I, and the length of the segments simultaneously. Therefore, the scaling does neither affect the transition probabilities nor the resulting model for the MCMC.
Since an accurate model for the MCMC step is available now, let us try to determine the influence of this model on the likelihood ratio per time, $\lambda(x_n)$, given by eq(4.3) in Section 4.1. Let us skip dependency of time and use the result of eq.(2.5) to model $\lambda(x_{pred})$ by the expected value of the density approximated by the predicted particles, i.e.

$$\lambda(x_{pred}) = \lim_{NoP \to \infty} \frac{1}{NoP} \sum_{j=1}^{NoP} \lambda(x^j_{pred}) \tag{4.23a}$$

$$= \frac{\frac{1}{NoP}\sum_{j=1}^{NoP} p_1(x^j_{pred})}{p_0(x)} \tag{4.23b}$$

$$= \frac{\mathbb{E}_1[X]}{p_0(x)}, \tag{4.23c}$$

where $\mathbb{E}_1[X]$ is the expected value of the density $p_1(x)$. Now the model for the MCMC step us used to model $\lambda(x_{post)}$ by

$$\lambda(x_{post}) = \lim_{NoP\to\infty} \frac{1}{NoP} \sum_{j=1}^{NoP} \lambda(x_{post}^j) \tag{4.24a}$$

$$= \frac{\lim_{NoP\to\infty} \frac{1}{NoP} \sum_{j=1}^{NoP} p_1(x_{post}^j)}{p_0(x)} \tag{4.24b}$$

$$= \frac{\lim_{NoP\to\infty} \frac{1}{NoP} \sum_{j=1}^{NoP} \pi(x) p_1(x_{pred}^j)}{p_0(x)} \tag{4.24c}$$

$$= \frac{\lim_{NoP\to\infty} \frac{1}{NoP} \sum_{j=1}^{NoP} x p_1(x_{pred}^j)}{p_0(x)} \tag{4.24d}$$

$$= \frac{\mathbb{E}_1[X^2]}{p_0(x)}. \tag{4.24e}$$

By (4.23a) and (4.24a), the influence of the on posterior likelihood ratio per time is given by a constant, denoted by C, satisfying

$$C = \frac{\lambda(x_{post})}{\lambda(x_{pred})} = \frac{\mathbb{E}_1[X^2]/p_0(x)}{\mathbb{E}_1[X]/p_0(x)} = \frac{\mathbb{E}_1[X^2]}{\mathbb{E}_1[X]} \tag{4.25}$$

As a result of this section, C can be calculated and implemented in the threshold calculations when the target pdf is a known density and approximated well enough by the particles. In Table 4.1, the values of C are given for the probability density functions of interest in this thesis.

Distribution	Uniform(a,b)	Exponential(μ)	Gamma(k,θ)	Generalized Pareto$(k,s,0)$
$\mathbb{E}_1[X]$	$\frac{1}{2}(a+b)$	μ	$k\theta = n\frac{\mu}{n} = n$	$\frac{s}{1-k}$
$\mathbb{E}_1[X^2]$	$\frac{a^2+ab+b^2}{3}$	$2\mu^2$	$\theta^2(k^2+k) = \frac{\mu(n^2+n)}{n}$	$\frac{2s^2}{(k-1)(2k-1)}$
C	$\frac{2(a^2+ab+b^2)}{3(a+b)}$	2μ	$\theta(k+1) = \frac{\mu(n+1)}{n}$	$\frac{2s}{(1-2k)}$

Table 4.1: Constant C for several distributions

Remark 21. For a good calculation of C, the (predicted) particle cloud should approximate a known probability density function. In practical application, it is often impossible to assign a good pdf to the particle cloud as the plots in Appendix D show. This makes it hard to use the factor C in threshold calculations.

Remark 22. As explained in Section 2.2.3, the MCMC step is performed to get a better approximation of the target distribution and to prevent the particles from degeneration. Remark 21 explained that it is very hard to model the influence of the MCMC step in practice. This makes it almost impossible to calculate a proper threshold for the posterior likelihood ratio. Hence, it might be better to use the predicted likelihood ratio instead.

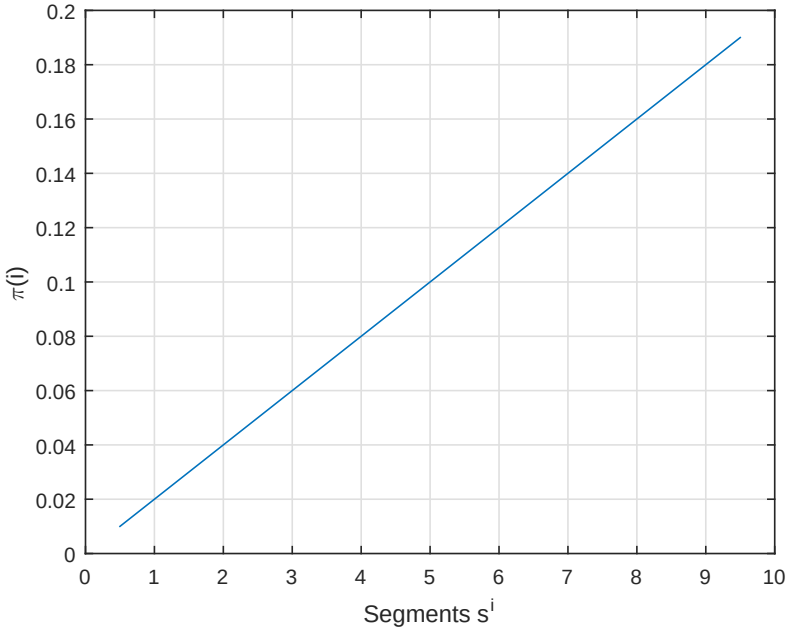
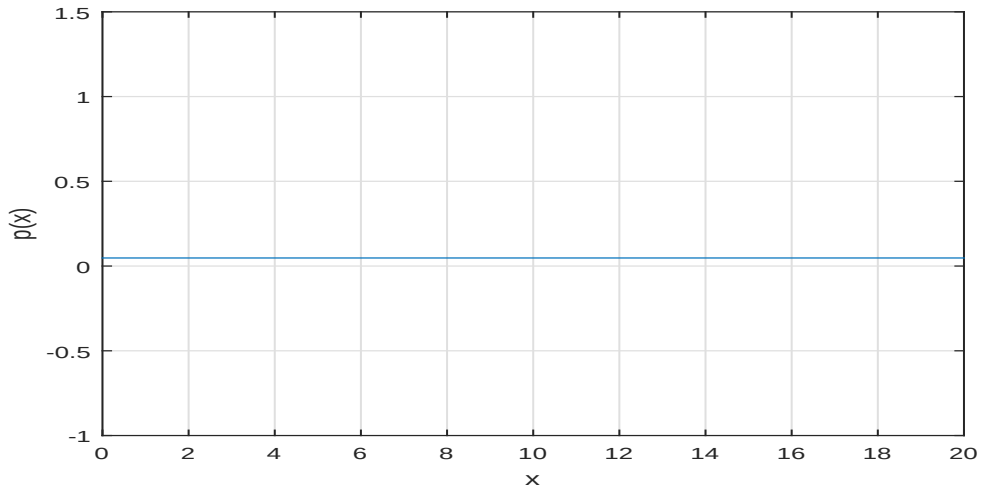
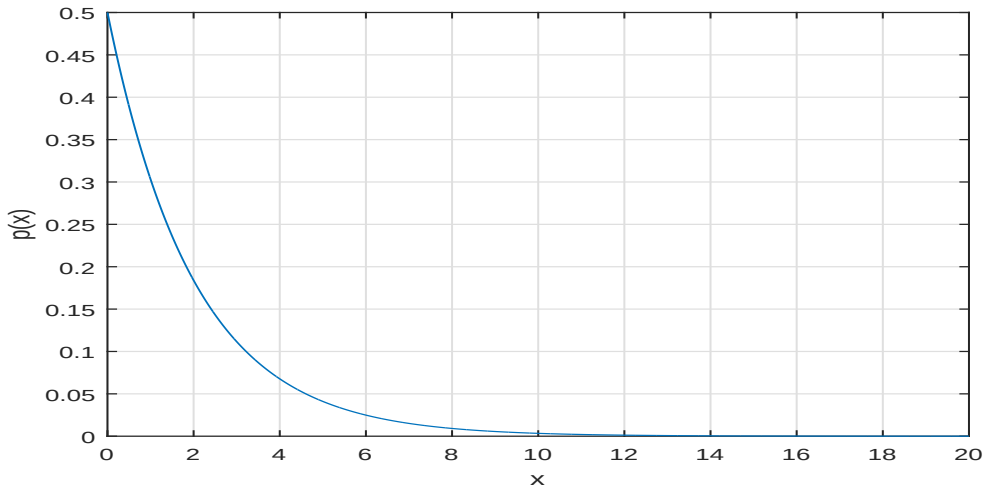


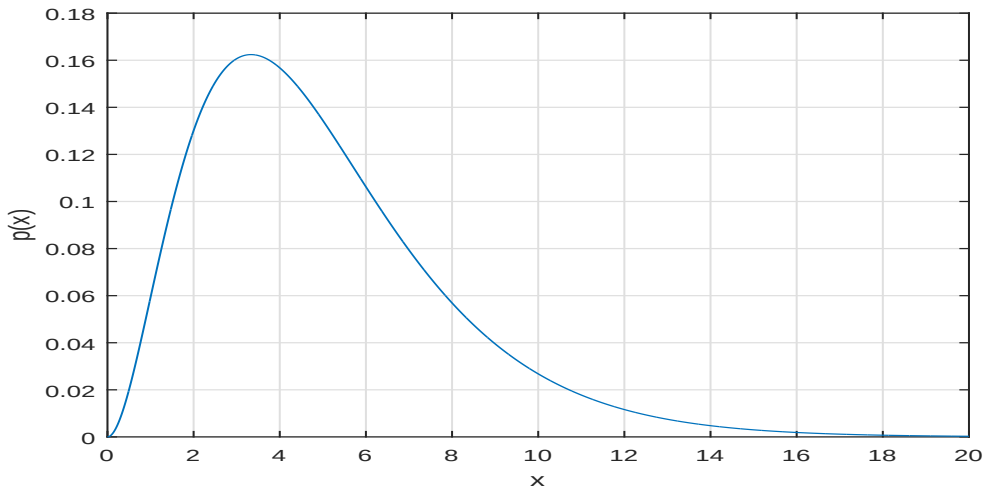
Figure 4.7: Graphical representation for model MCMC step



(a) Uniform distribution

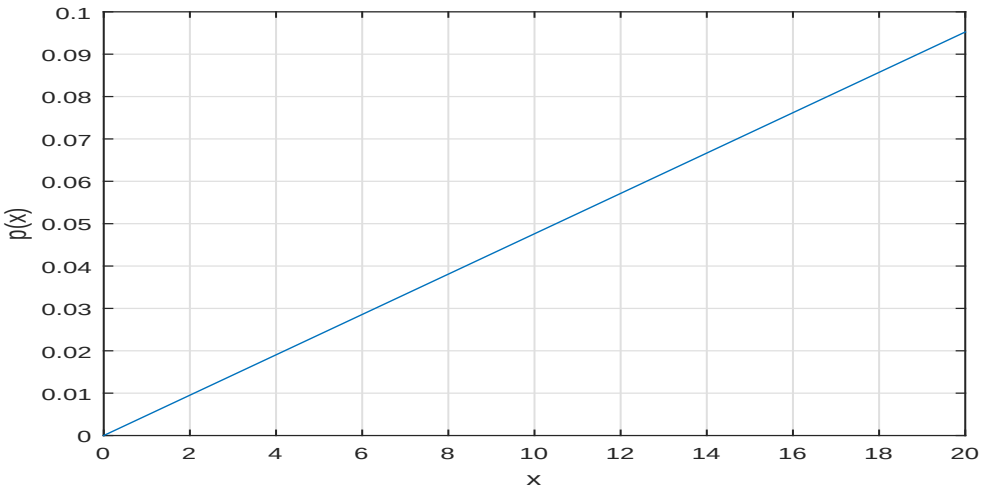


(b) Exponential distribution

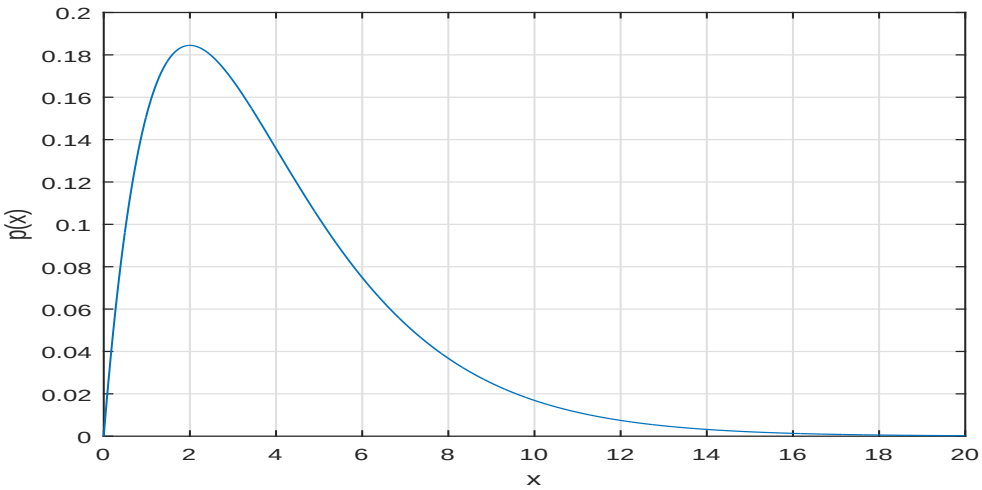


(c) Gamma distribution

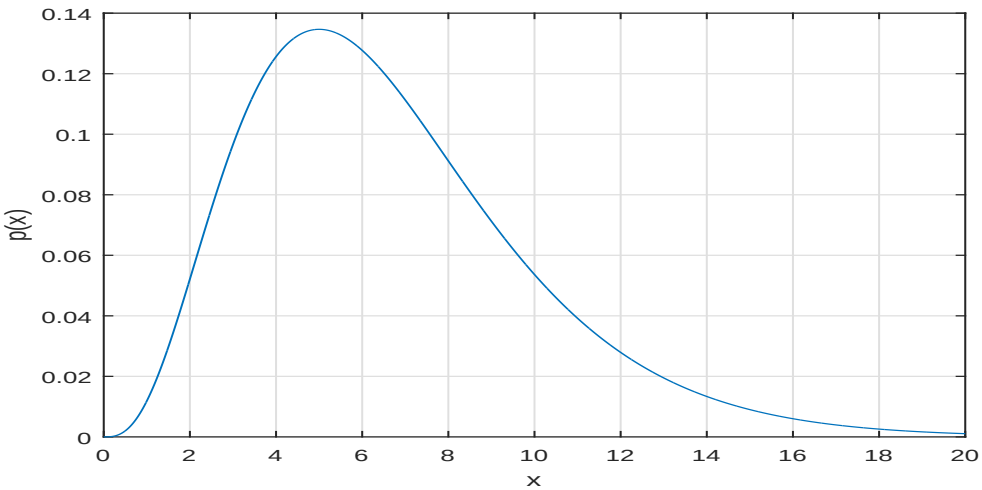
Figure 4.8: Predicted distributions



(a) Uniform distribution



(b) Exponential distribution



(c) Gamma distribution

Figure 4.9: Posterior distributions

4.3 Results proposed models

In this section, the results of the ideas in Section 4.2 will be presented. Section 4.2.3 concluded with the remark that in practice it is almost impossible to integrate the derived model for the MCMC step in threshold calculations because the particle cloud will not show a proper and known pdf in general. Therefore, rather than giving the results of the posterior likelihood ratios, as in Section 3.2 and Section 3.3 the results of the predicted likelihood ratios are presented. The results for the ideas in Section 4.2 will be compared to the results of predicted likelihood ratios for the same tracks used in Section 3.3. The results are obtained by gathering data from the filter and doing the hypothesis testing afterwards. The likelihood ratios obtained from the filter are tested against a threshold, where the mean of the estimated target power per time and a model for the spread of the particles are integrated in the threshold calculations.

In Figures 4.10 and 4.11, the P_{FA} and P_D curves for different scenarios are given. The curves correspond to data from the particle filter with Gaussian background noise for three different power levels, $\mu = 1$, $\mu = 2$ and $\mu = 5$. The results are obtained by testing the likelihood ratios against the thresholds given in Figure 4.12. These threshold are calculated by applying Neyman-Pearson theory. First of all, Figure 4.10a shows the P_{FA} curves of predicted likelihood ratios (for the SLRT) with data from the filter without using the parameter estimate fluctuations or the method to model the particle cloud. Despite the conclusion in Section 4.2.3 that the predicted likelihood ratios can be controlled much better than the posterior likelihood ratios, the P_{FA} values in Figure 4.10a show similar bad behavior to the curves of the posterior likelihood ratios in Figure 3.7a. Furthermore, Figure 4.12a shows that the thresholds for the different power levels are very different. This is a problem in achieving a useful CFAR detection scheme, because for CFAR detectors to be useful in practice, the threshold should be the same for varying background parameters.

In Figures 4.10b and 4.10c the P_{FA} curves for both the SLRT as the LRT are given. The LRT values are plotted for comparison with Figure 3.4a. Figure 4.10b shows that when the estimates of the RCS , σ, and consequently the target power, μ_t are integrated in the threshold calculations, the P_{FA} curves for both the SLRT and the LRT immediately fulfill the desired significance level, $P_{FA} \leq \alpha$. Unfortunately, the P_{FA} values are still not exactly the same and all three curves converge to values far beneath the time-integrated α. This indicates that the calculated thresholds may be too high. These thresholds are shown in Figure 4.12b. The first notable is that the threshold values for all three different power levels are exactly the same. This shows that when integrating the particle cloud estimate of σ and hence μ_t in the threshold calculations, already the same threshold can be applied for varying noise power levels. The second notable thing is that the threshold is no longer (almost) monotonically decreasing in time as in Figure 4.12a. This observation can be explained as follows. In Section 4.1.1, it is discussed that the estimate of the RCS and consequently the estimation of the mean target power drop heavily in the first few time steps when incoming measurements do not match with a high initialization (shown in Figures 4.1b and 4.2b). As a result, the threshold given by eq.(3.22) also drops during the first few time steps of the track. After some time, the estimate of μ_t stabilized and the thresholds slightly increases, due to the increasing variance of the time integrated measurement samples.

In Figures 4.10c, 4.11c and 4.12c also the spread of the particle cloud is integrated in the threshold calculations. The spread of the particle cloud can not be determined in advance because it depends on for example the target power, background power levels and the performance of the

filter. Let us refer back to Section 4.1 for figures and more comments about the variation in the spreading. To integrate the spread in the threshold calculations, a general model is needed to represent the "effective" number of cells covered by the cloud. This number is significantly lower than the total number of covered measurement cells in Figures 4.1a and 4.2a, because the number of particles in a cell should be sufficiently high to influence the likelihood ratio. Therefore the cells where very few particles are positioned, should not be counted. The model for the effective number of cells is chosen to be a linear model, starting at with 1 cell at time 1. By some quick simulations, where the gradient is interpolated between 0.2 and 0.5, a gradient of 0.33 seemed to be sufficient. As a results, the function used to model the (effective) number of cells covered by the particle cloud is

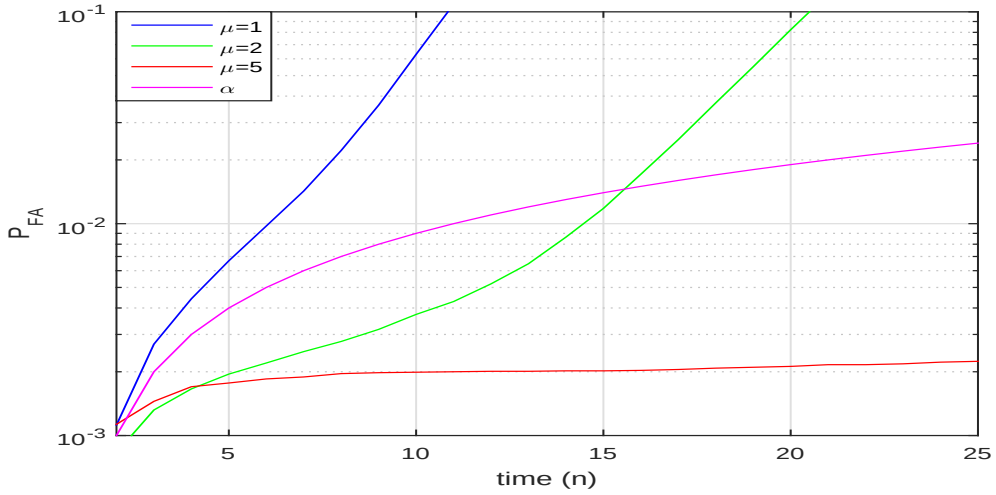
$$c_n = 1 + 0.33(n-1), \quad n = 1, 2, \ldots \tag{4.26}$$

Since, time integration is performed on the likelihood ratios, also time integration on c_n is used in the threshold calculations. Let us refer back to Section 4.2.2 for the exact calculations.

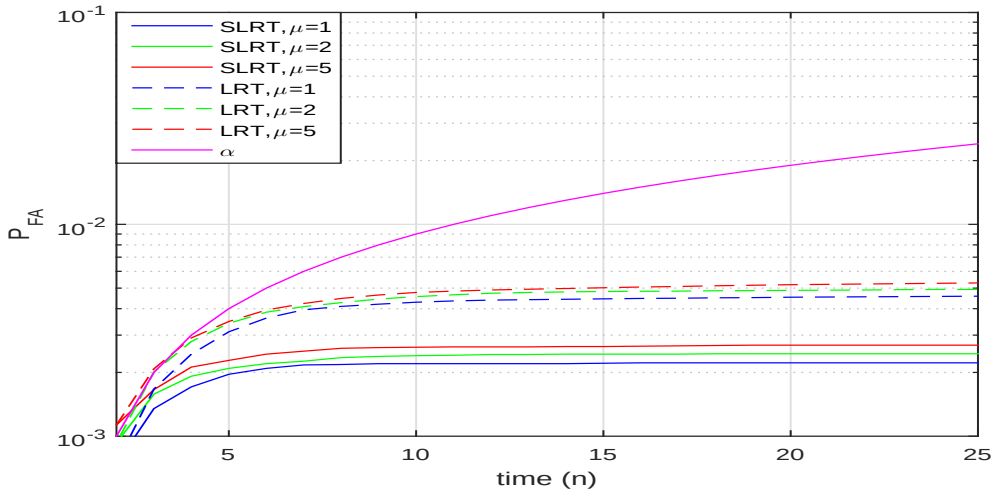
Figures 4.10c, 4.11c and 4.12c show promising results. In Figure 4.10c the P_{FA} values for both the SLRT and the LRT are approximately the same. Also the calculated threshold, given in Figure 4.12c, is the same. The small deviations in the threshold starting around $n = 12$ are caused by numerical approximation errors. The Gamma distribution is used to find $\tilde{\tau}$ by solving eq.(4.8). As a result, the exact value of $\tilde{\tau}$ can not be found and is therefore approximated with use of the cumulative Gamma distribution in Matlab. This causes the small deviations in the threshold. Compared to the thresholds in Figure 4.12b, the thresholds in Figure 4.12c are significantly lower. This can be explained by the fact that by integrating the spreading particle cloud using the sample mean, as discussed in Section 4.2.2, the variance of the sample mean is lower than the variance of the time integrated samples. Consequently, the thresholds are lower.

The combined results from Figures 4.10c and 4.12c fulfill the CFAR property. It seems that the particle cloud spread for different power levels in noise do not show a large variation. This suggests that the function in eq.(4.26) is an accurate model for all three power levels. The P_{FA} values for the LRT at $n = 25$ are extremely close to the desired significance level α. The P_{FA} values for the SLRT show approximately the same Sequential loss as for the i.i.d. generated data in Figure 3.4a. Also the P_D in Figure 4.11c show good results as even for $\mu = 5$ (SNR = 3dB) the probability of detection is nearly one at $n = 25$. Furthermore, it seems that using the thresholds Figure 4.12c instead of the threshold in Figure 4.12a, does not affect the P_D.

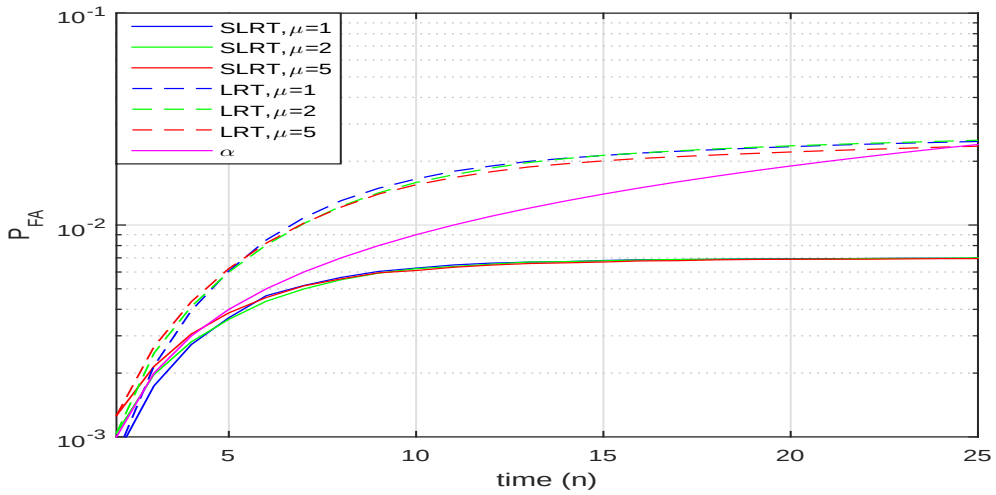
In Figure 4.13a it is shown that it is even possible to empirically tune the effective number of cells covered by the particle cloud in such a way, that the P_{FA} values for the LRT are exactly the same as the desired α. This suggests that if c_n can be modeled very accurately, the Neyman-Pearson theory defines the most powerful LRT (even without satisfying the condition of i.i.d. samples) when the ideas and methods described in Sections 4.1 and 4.2 are integrated in the threshold calculations.



(a) No μ_t and particle spread implemented

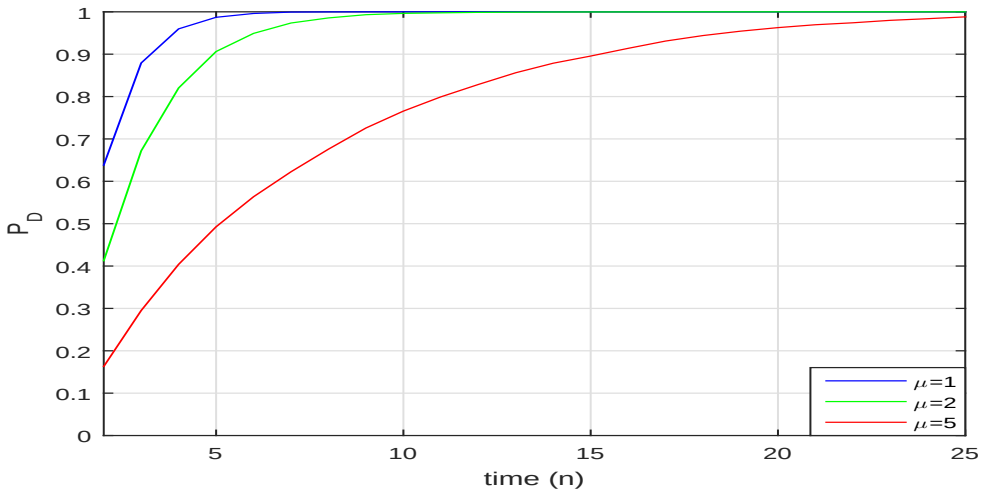


(b) μ_t implemented

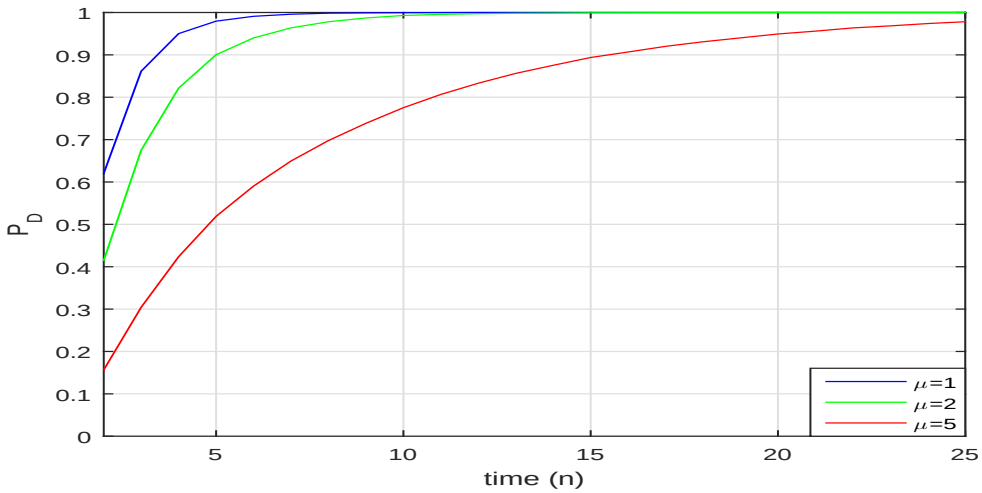


(c) μ_t and particle spread implemented

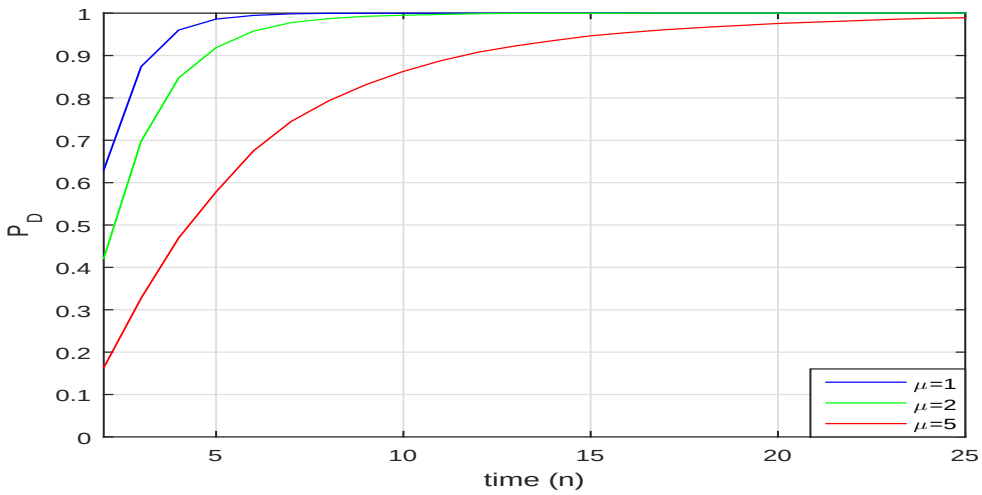
Figure 4.10: P_{FA} curves for predicted likelihood ratios with and without models implemented



(a) No μ_t and particle spread implemented

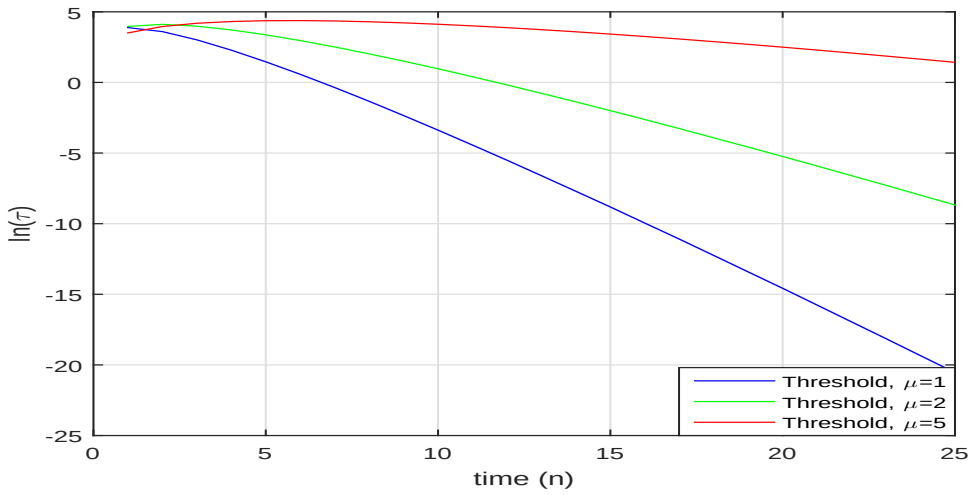


(b) μ_t implemented

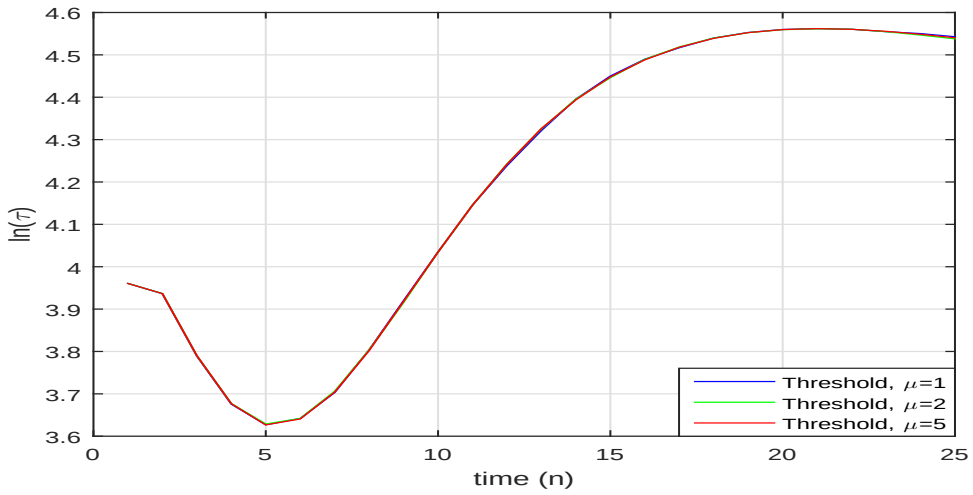


(c) μ_t and particle spread implemented

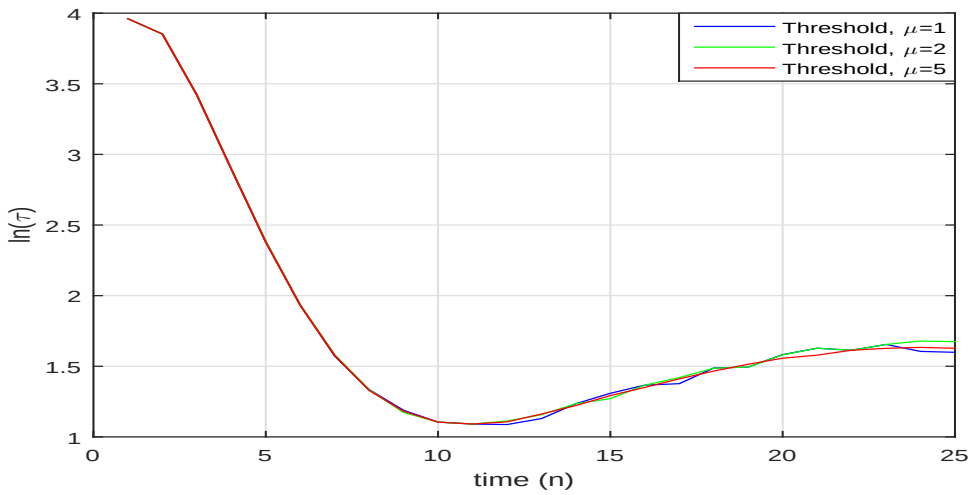
Figure 4.11: P_D curves for predicted likelihood ratios with and without models implemented



(a) No μ_t and particle spread implemented

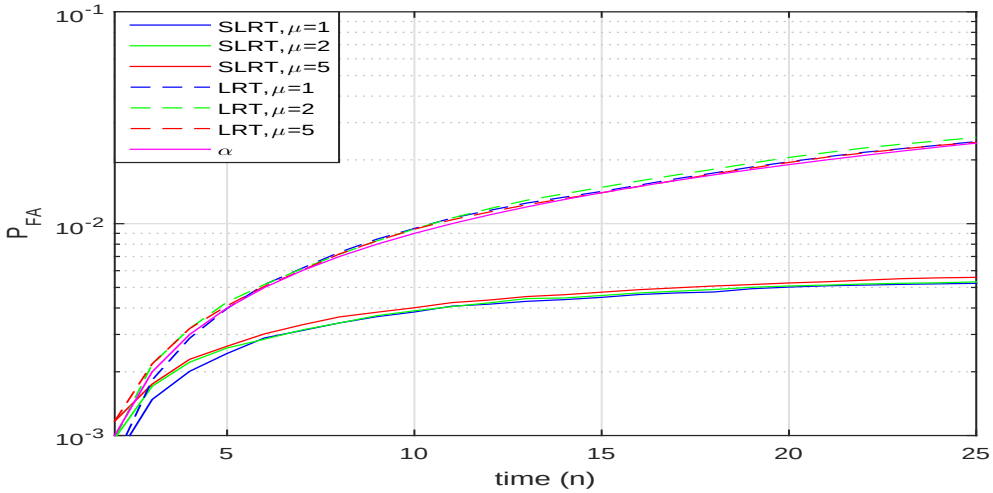


(b) μ_t implemented

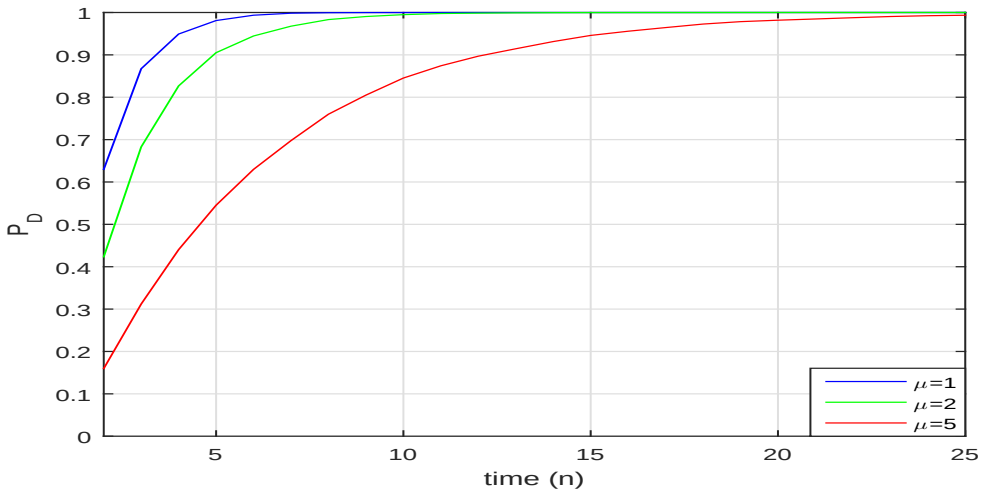


(c) μ_t and particle spread implemented

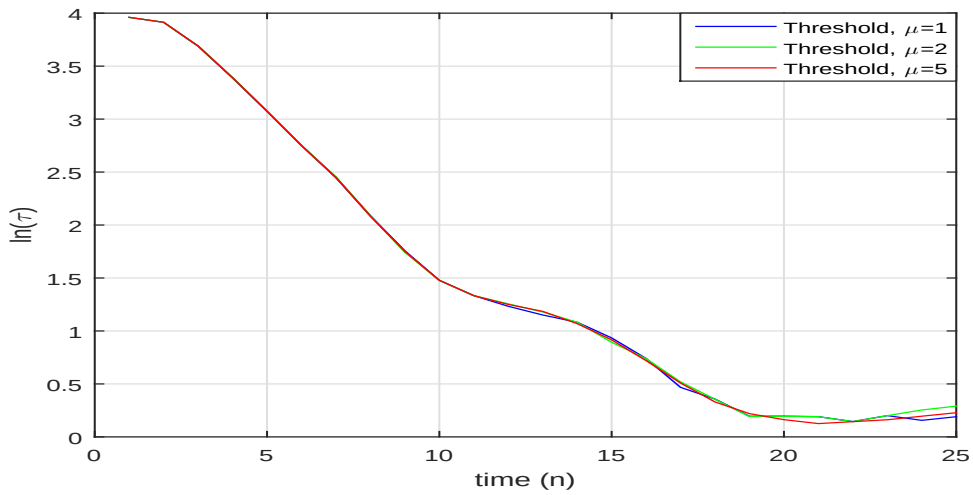
Figure 4.12: Thresholds $\ln(\tau)$ using Neyman-Pearson theory with and without models implemented



(a) P_{FA} curves



(b) P_D curves



(c) Threshold $\ln(\tau)$

Figure 4.13: Results for empirically tuned values of c_n

4.4 Proposed CFAR detection scheme

In this section, a sequential detection scheme is proposed that could be applied during the tracking. This detection scheme uses the ideas from Section 4.2, such that the results should fulfill the CFAR property, likewise the results of Section 4.3.

In Section 4.1 it is discussed that the integrated likelihood ratio at time n, $\lambda(\mathbf{x}_n)$, is used for the hypothesis testing, where $\lambda(\mathbf{x}_n)$ is given by

$$\lambda(\mathbf{x}_n) = \prod_{i=1}^{n} \lambda(x_i), \tag{4.27}$$

and $\lambda(x_i)$ is given by the mean over the particles, i.e.

$$\lambda(x_i) = \frac{1}{NoP} \sum_{j=1}^{NoP} \lambda(x_i^j). \tag{4.28}$$

Instead of testing $\lambda(\mathbf{x}_n)$ against τ_n, the idea is to scale the likelihood ratio per particle at time n, $\lambda(x_n^j)$ in a way that the threshold is contained in the scalings factor such that the new likelihood ratio can be tested against a new threshold, denoted by τ_{new}, equal to 1. The advantage is that this new likelihood ratio can be tested against a constant threshold, and available parameter estimates and information about spread particle cloud can be directly used in the design of the scalings factor. Let us refer to the scalings factor as "threshold likelihood ratio", denoted at time n by

$$\lambda^T(\tau_n) = \frac{p(\tau_n|\mathcal{H}_1)}{p(x_n|\mathcal{H}_1)}, \tag{4.29}$$

where τ_n can be found from solving

$$P_{FA} = \int_{\{x_n:\lambda(x_n)>\tau_n\}} p(x_n|\mathcal{H}_0)\mathrm{d}x_n = \alpha. \tag{4.30}$$

Note that $\lambda^T(\tau_n)$ varies per particle because of different estimates of μ_t. Let us define $\lambda^D(x_n^j)$ as the detection likelihood ratio per particle, denoted by

$$\lambda^D(x_n^j) = \frac{\lambda(x_n^j)}{\lambda^T(\tau_n)}.$$

Since the desired significance level α is included in $\lambda^T(\tau_n)$, each particle should satisfy $P_{FA} \leq \alpha$ when testing $\lambda^D(x_n^j) \gtrless 1$. As a result, the mean over the particles of the detection likelihood ratio, given by

$$\lambda^D(x_n) = \frac{1}{NoP} \sum_{j=1}^{NoP} \lambda^D(x_n^j) \tag{4.31}$$

should approximately satisfy $P_{FA} \leq \alpha$ as well. Unfortunately, time integration over $\lambda^D(x_n)$ by

$$\lambda^D(\mathbf{x}_n) = \prod_{i=1}^{n} \lambda^D(x_n),$$

would results in $\lambda^D(\mathbf{x}_n)$ being scaled n times by a threshold likelihood ratio in a way that testing $\lambda^D(\mathbf{x}_n) \gtrless 1$ satisfies $P_{FA} \leq \alpha^n$. This would ruin the probability of detection. To overcome this problem, $\lambda^D(\mathbf{x}_n)$ is defined as

$$\lambda^D(\mathbf{x}_n) = \left(\prod_{i=1}^{n-1} \lambda(x_i)\right) \lambda^D(x_n), \tag{4.32}$$

where $\lambda^D(x_n)$ is given by eq.(4.31) and $\lambda(x_i)$ by eq.(4.28).

Testing $\lambda^D(\mathbf{x}_n) \gtrless 1$ is a bit different than $\lambda(\mathbf{x}_n) \gtrless \tau_n$. When testing $\lambda(\mathbf{x}_n) \gtrless \tau_n$, the threshold is calculated according to the distribution of time integrated measurements, where the mean of RCS estimates is used as explained in Section 4.3 and possibly with $\tilde{c}_n$ from eq.(4.9) implemented. By design of $\lambda^D(\mathbf{x}_n)$, $\lambda^D(\mathbf{x}_n) \gtrless 1$ is tested based on the threshold implemented in $\lambda^D(x_n)$, calculated according to only the RCS estimates and possibly the number of cells at time n. Furthermore, a proper implementation of c_n as in Section 4.2.2 is not clear. In Section 4.2.2 the implementation of c_n, and later $\tilde{c}_n$, is performed by the sample mean, based on taking the mean of all particles to find $\lambda(x_n)$ as in eq.(4.3). Since in the implementation of $\lambda^D(\mathbf{x}_n)$ the threshold is already implemented in the scaling of particles, using the idea of the sample mean might not seem valid. However, because $\lambda^D(\mathbf{x}_n)$ is the mean over the particles as in eq.(4.31), the sample mean is used nonetheless.

Two methods for testing the idea of the sample mean are tested. In both methods τ_n is found by solving eq.(4.30), where the function for c_n (eq.(4.26)) is integrated in $p(x_n|\mathcal{H}_0)$. In method 1, $\lambda^{T_1}(\tau_n)$ is given by eq.(4.29). This scaling might not be strong enough, because c_n is only used to find τ_n but not again in constructing $\lambda^{T_1}(\tau_n)$. Therefore in method 2, $\lambda^{T_2}(\tau_n)$ is defined as

$$\lambda^{T_2}(\tau_n) = \left(\lambda^T(\tau_n)\right)^{c_n} \tag{4.33}$$

such that c_n is also used in the likelihood ratio.

The results for the proposed detection scheme in Gaussian noise are presented in Figures 4.14 and 4.15. Results for Pareto clutter are presented in Figures 4.16 and 4.17. The results of noise are compared to the results in Section 4.3 to discuss performance of the proposed detection scheme. Let us first compare Figure 4.14 with Figure 4.10b. These figures both show the P_{FA} curves when the RCS estimates are used in the threshold. The P_{FA} curves in both figures for the LRT show slightly different values as the curves in Figure 4.14 keep increasing in time. The curves for the SLRT seem to be exactly the same. This shows that the proposed detection scheme works approximately the same as the method in Section 4.3, at least when the spread of the particle cloud is not modeled.

In Figure 4.14b and Figure 4.14c, the P_{FA} results are given when the particle cloud is modeled with method 1 and 2 respectively. As expected, the P_{FA} values for method 1 are higher than for method 2. Compared to Section 4.3, the values of method 1 are higher and the values for method 2 are lower than the values in Figure 4.10c. Assuming that the threshold used to obtain Figure 4.10c is properly derived, the proper scaling of $\lambda^T(\tau_n)$ seems therefore to be somewhere in between the scalings of method 1 and 2. Important is that the P_{FA} values for method 1 and 2 for the three power levels are approximately the same, just as in Figure 4.10c. This means that the proposed detection satisfies the CFAR property as a constant false alarm rate is acquired by testing $\lambda^D(\mathbf{x}_n) \gtrless 1$ at all n. The P_D curves in Figure 4.15 also show promising results. The P_D curves for method 1 and 2 are slightly different as expected, but even for $\mu = 5$ (SNR = 3dB),

the probability of detection for the SLRT converges to 1.

In Figures 4.16 and 4.17 the P_{FA} and P_D curves of the SLRT for clutter are given for 6 different values of shape parameter k. Scale parameter s is fixed at 1 and θ is assumed to be 0. In Figures 4.16a and 4.17a, the particle cloud is not modeled and therefore these figures can be compared with Figures 4.14a and 4.15a, because the scaling with the threshold parameter is properly applied (assuming that eq.(2.32) approximates the pdf of the convolution of target+clutter close enough). In Figures 4.16b and 4.17b, the particle cloud is integrated using the scaling in method 1 and the Gamma distribution, knowing that this Gamma distribution does not coincide with the samples mean of clutter samples. Nonetheless, it is tested to see if the P_{FA} curves are merged compared to curves in Figure 4.16a.

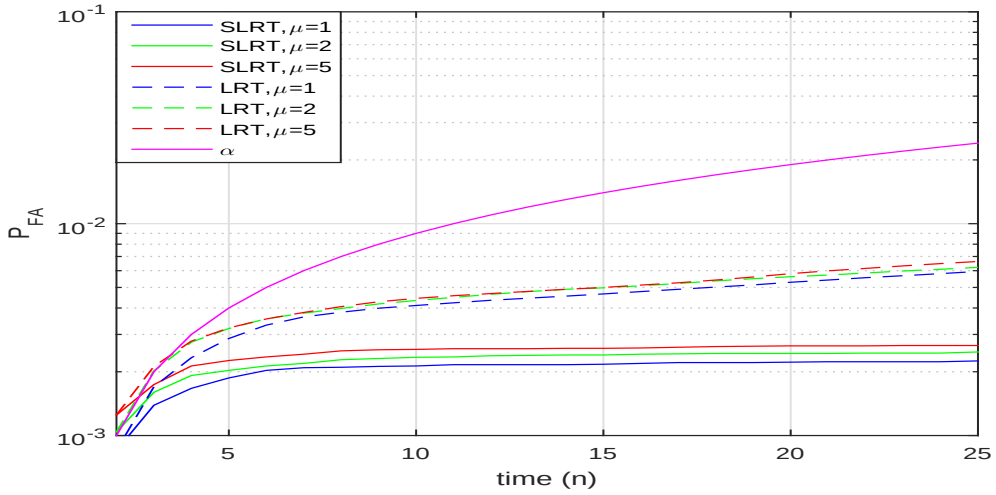
In Figures 4.16 and 4.17, the P_{FA} and P_D values are given for k varying from 10^{-5} to 0.5. $k = 10^{-5}$ is chosen such that eq.(2.16) is still valid, but 10^{-5} will now be referred to as 0. The values of k correspond to mean background powers of respectively 2, 2.1, 2.22, 2.35, 2.5 and 2.67. Despite these small changes in the mean clutter power, the P_{FA} curves significantly more spread than the curves in Figure 4.14a. This shows the difficulty of controlling the P_{FA} values in clutter. Compared to Figure 4.16a, the lines in Figure 4.16b are merged a bit. This shows again that integrating the spread of the particle cloud helps towards achieving the CFAR property.
In Figure 4.15, the values for k lead to a Signal to Clutter Ratio (SCR) of respectively 7dB, 6.8dB, 6.5dB, 6.3dB, 6dB and 5.7dB. The P_D curves show expected behavior as the P_D increases when SCR increases. Also, the P_D values in Figure 4.17b are higher than in Figure 4.17a similar to the results for noise. Furthermore, the P_D for all SCR's converges to 1 eventually.

Overall, for better performances of the P_{FA} values in clutter, the expression of the sample mean in Equation (4.15) could be investigated. Also investigation of the variation of the particle cloud spread is needed to integrate c_n in the threshold properly. The complete SMCMC filter with integrated detection scheme is summarized in Algorithm 6.

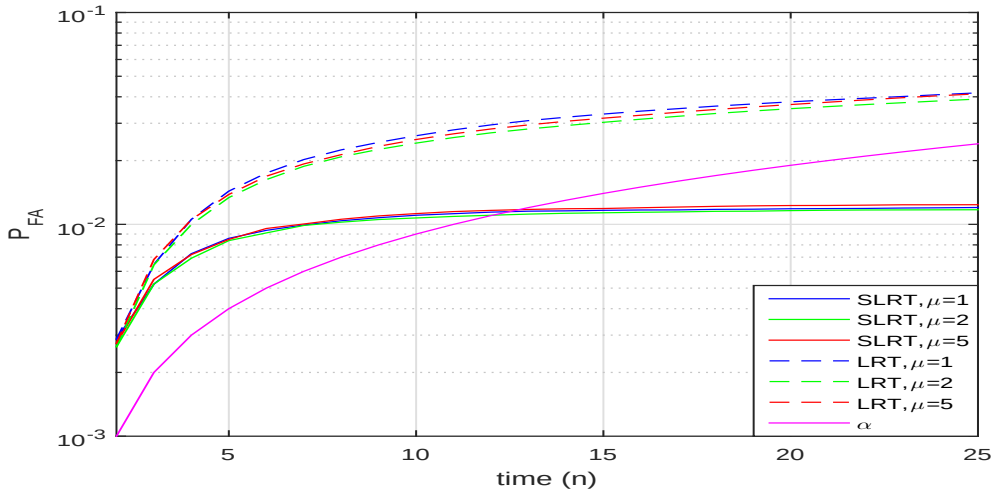
Algorithm 6: SMCMC with integrated CFAR detection scheme.

Input: Measurement z_0, Number of particles NoP, Significance level α, Maximum time horizon N.
Output: Either $\mathcal{H}_1$ or $\mathcal{H}_0$.

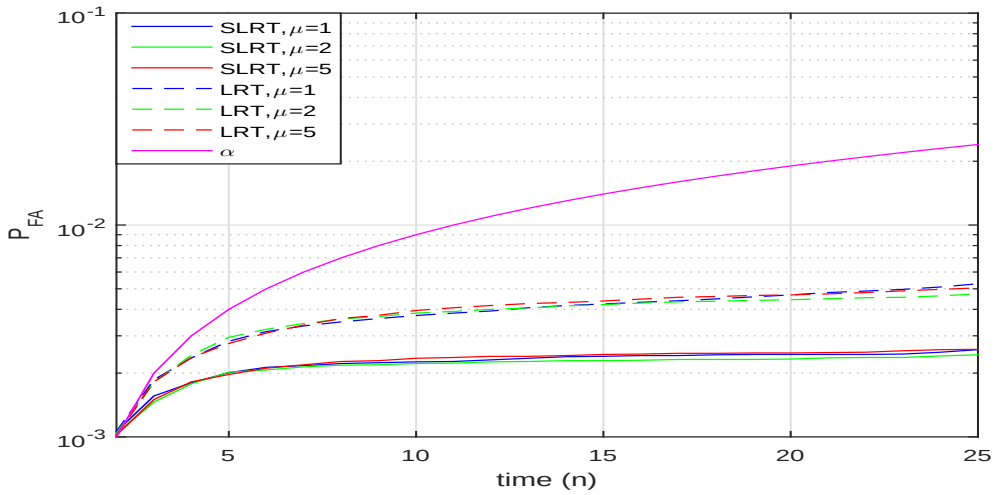
```
1 - Initialize track;
Start track at measurement cell ∈ z_0 with highest power;
Generate initial particles {x_0^j}_{j=1}^{NoP} from p(x_0);
for n:=1 to N do
    if n=1 then
        2 - Prediction;
        for j:=1 to NoP do
            Draw particle x_{n_pred}^j ~ q_n(x_n|x_{n-1}^j, z_n);
        end
        3 - Update;
        for j:=1 to NoP do
            Compute predicted likelihood ratios λ(x_{n_pred}^j) = p(z_n|x_{n_pred}^j) / p(z_n);
        end
        7 - Perform MCMC step: Apply MH algorithm;
        {x_{post}^m, λ(x_{post}^m)}_{m=1}^{NoP} = MH({x_{pred}^j; λ(x_{pred}^j)}_{j=1}^{NoP});
        8 - Copy set of posterior particles: {x_n^j}_{j=1}^{NoP} = {x_{n_post}^j}_{j=1}^{NoP};
    else
        2 - Prediction;
        for j:=1 to NoP do
            Draw particle x_{n_pred}^j ~ q_n(x_n|x_{n-1}^j, z_n);
        end
        3 - Update;
        for j:=1 to NoP do
            Compute predicted likelihood ratios λ(x_{n_pred}^j) = p(z_n|x_{n_pred}^j) / p(z_n);
        end
        4 - Calculate threshold likelihood ratio;
        for j:=1 to NoP do
            Solve P_FA = ∫_{{x_n:λ(x_n)>τ_n}} p(x_n|H_0)dx_n = α.;
        end
        5 - Calculate detection likelihood ratio λ^D(x_n);
        6 - Perform hypothesis testing;
        if λ^D(x_n) ≥ 1 then
            Conclude H_1, break;
        else
            No decision, go to next step;
        end
        7 - Perform MCMC step: Apply MH algorithm;
        {x_{post}^m, λ(x_{post}^m)}_{m=1}^{NoP} = MH({x_{pred}^j; λ(x_{pred}^j)}_{j=1}^{NoP});
        8 - Copy set of posterior particles: {x_n^j}_{j=1}^{NoP} = {x_{n_post}^j}_{j=1}^{NoP};
    end
end
if n=N then
    Conclude H_0;
else
    Conclude H_1;
end
```



(a) Particle cloud not modeled

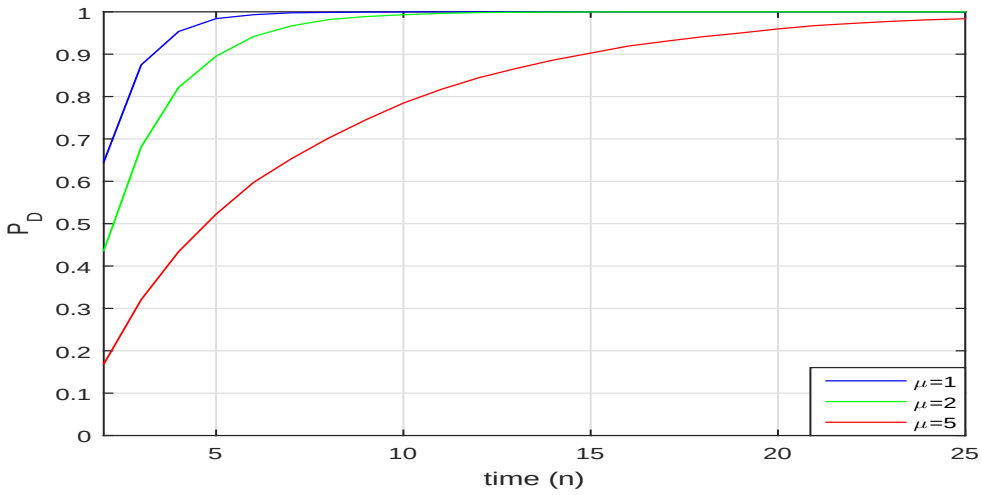


(b) Particle cloud modeled, method 1

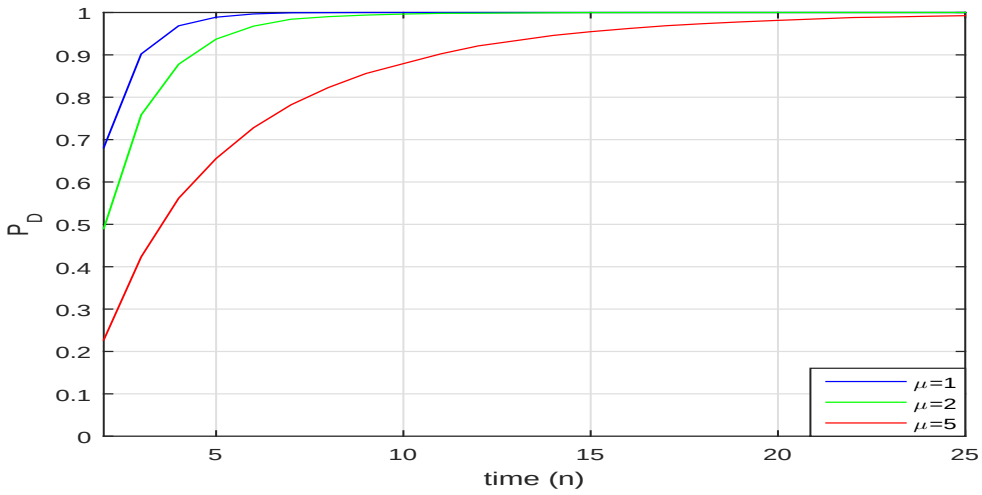


(c) Particle cloud modeled, method 2

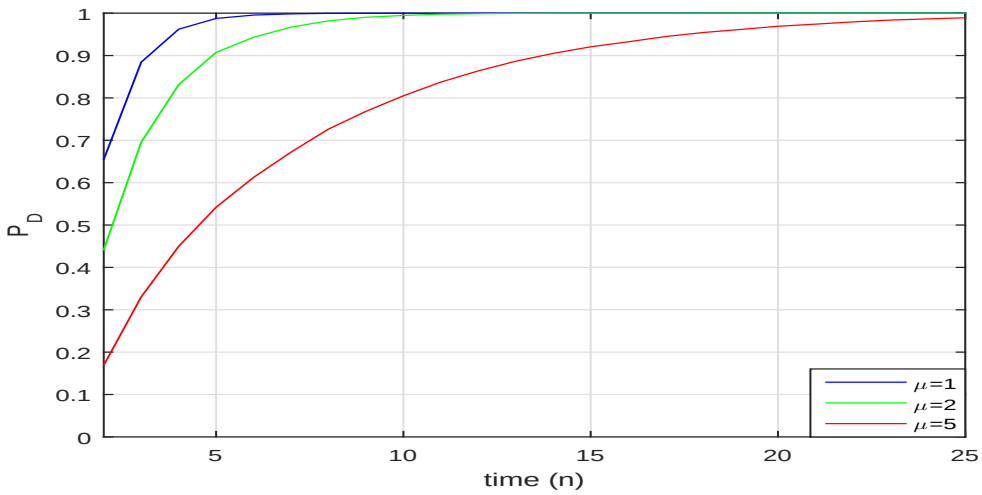
Figure 4.14: P_{FA} curves proposed detection scheme



(a) Particle cloud not modeled

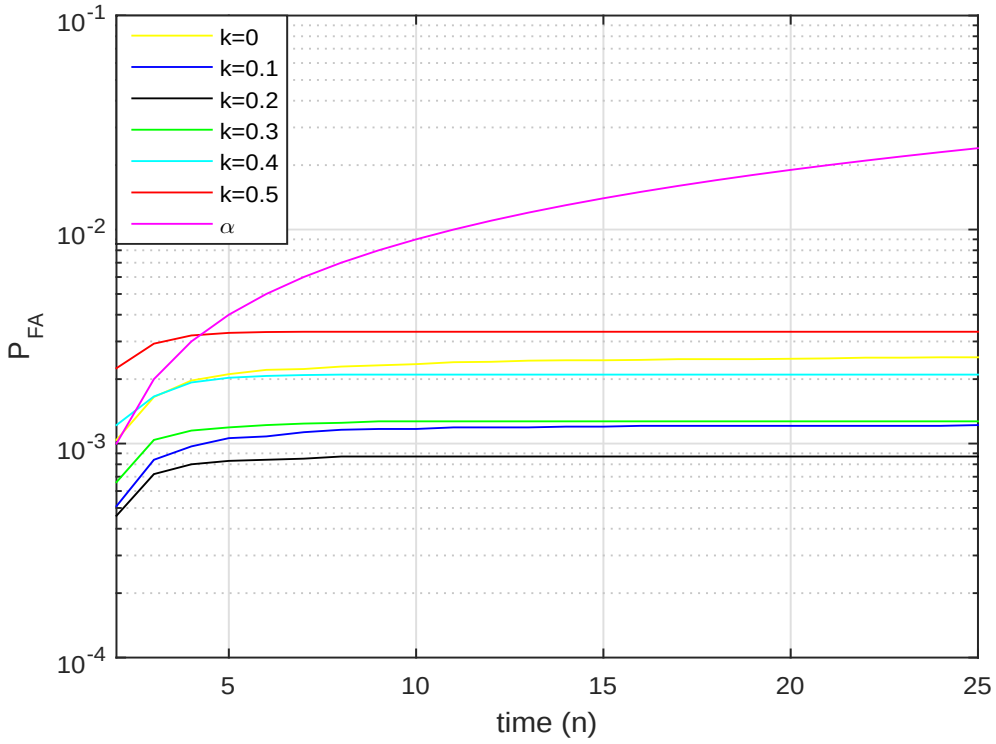


(b) Particle cloud modeled, method 1

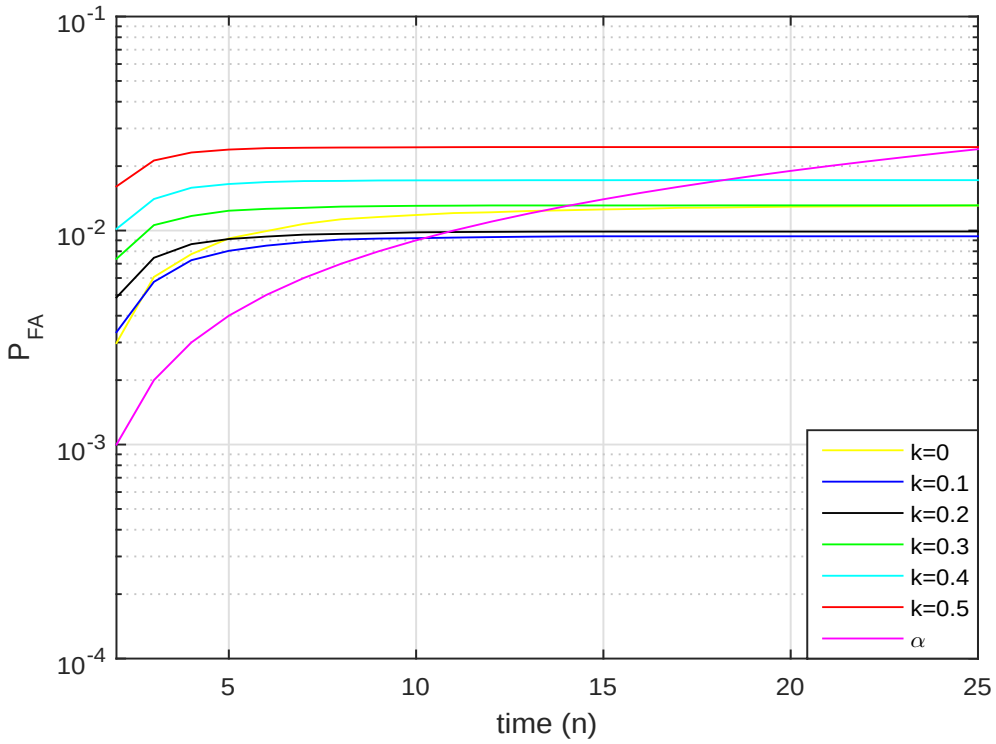


(c) Particle cloud modeled, method 2

Figure 4.15: P_D curves proposed detection scheme

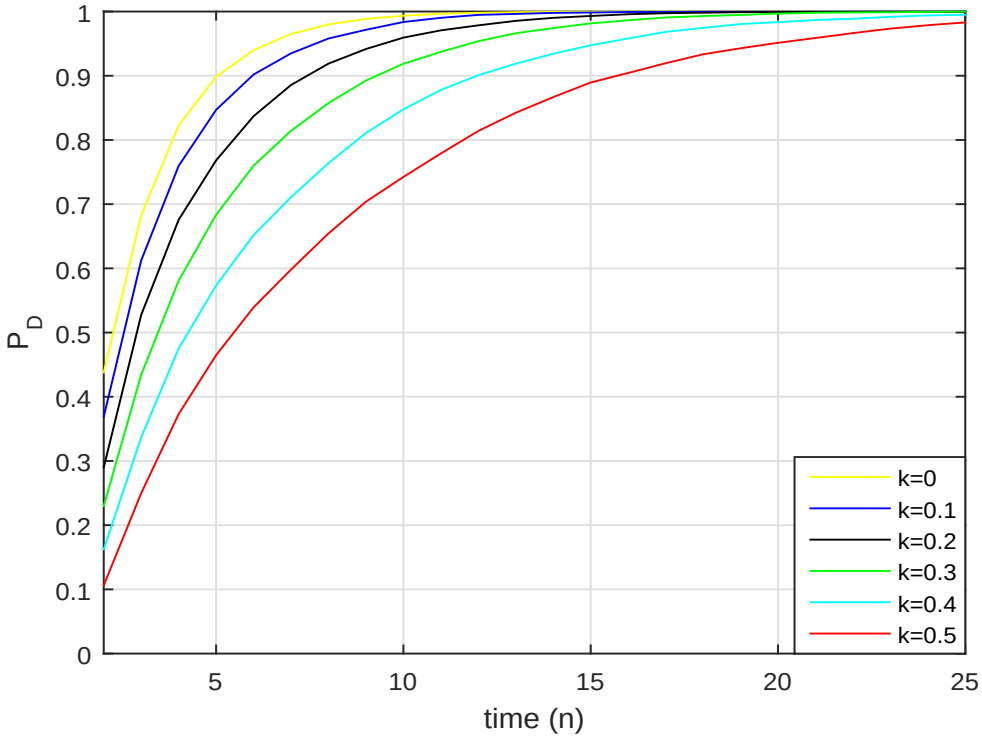


(a) Particle cloud not modeled

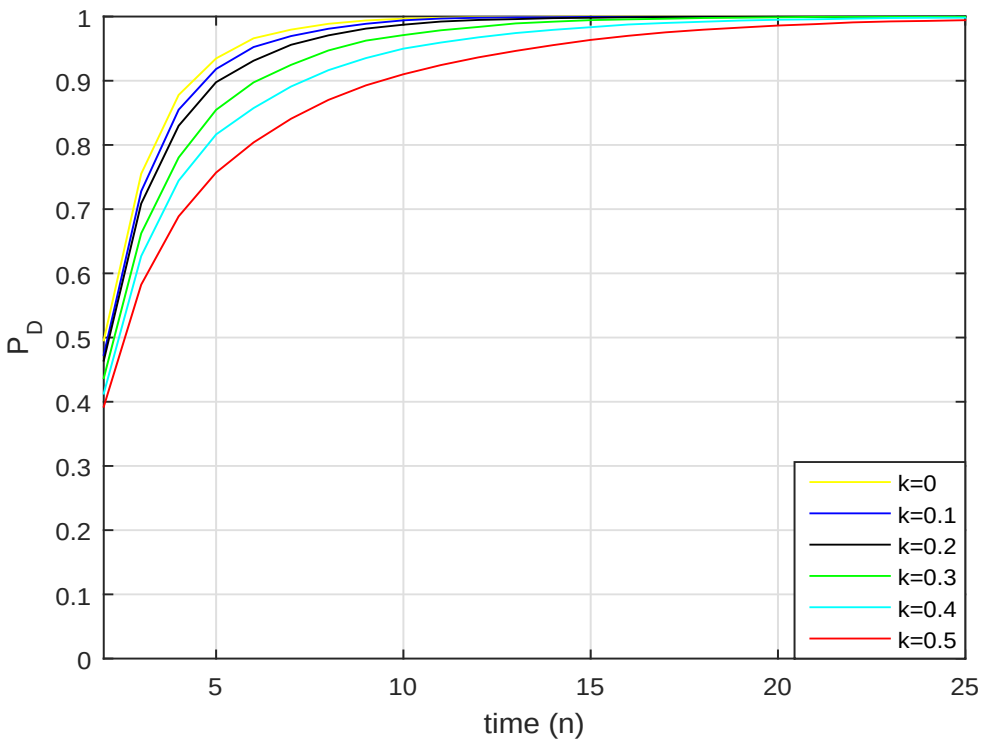


(b) Particle cloud modeled, (method 1)

Figure 4.16: P_{FA} curves with scaling of $\lambda(\mathbf{x}_n)$



(a) Particle cloud not modeled



(b) Particle cloud modeled, (method 1)

Figure 4.17: P_D curves with scaling of $\lambda(\mathbf{x}_n)$

Chapter 5

Conclusions and recommendations

In this chapter the conclusions of this research will be presented and recommendations for further research are given. First, the results in Sections 4.3 and 4.4 are discussed together with simplifications to the models. Based on the results and the discussion, the conclusions about the research are given in Section 5.2. The contributions for Thales are summarized in Section 5.3. Finally, several recommendations for further research are given

5.1 Discussion

In this section, the results of the research are discussed. Also, the most important properties, made assumptions and observations that play a role in this thesis are discussed. First, simplifications in the particle filter will be discussed. Second, comments about the assumptions of the hypothesis testing theories, i.e. Wald's theory and Neyman-Pearson Theory are given. Third, the reasoning behind the observations and the ideas for the models in Sections 4.1 and 4.2 are discussed. Finally, the results in Sections 4.3 and 4.4 are elaborated upon.

As explained in Chapter 2, the particle filter used in this research is a Sequential Markov Chain Monte Carlo (SMCMC) filter. Because of the complexity of the treated problem in this research, some simplifications in the filter are used compared to the filters in practical Radar application. The first simplification is that the background power is assumed to be known. Furthermore, filtered data is generated such that the used dynamic target model fits well with the data. Process noise is generated to simulate the uncertainty of the dynamic model in practice
The second simplification is the reduction in dimension of the state space $\mathcal{X}$. In this research the state space has dimension three (range, velocity and RCS), where only the range and RCS are used to compute the likelihood functions. The velocity associated to the single target is used to predict the target time evolution. In practice, the dimension of the state space $\mathcal{X}$ is often at least 6, such that the range, Doppler speed and azimuth of possible targets can be calculated. Furthermore, in this thesis is a single state space model. Hence it is assumed that a maximum of one target is present, while in practice multiple targets could be present simultaneously.
The third simplification is the assumption that a single scan of the Radar consist of only one burst, as briefly mentioned in Section 2.1.1. As a result, a Swerling I target model could be used where in case of multiple burst per scan a more complex Swerling II or Swerling IV should have been used, in which the RCS fluctuations are modeled from burst to burst.

A general remark is that during this research noise and clutter are modeled by respectively exponential and generalized Pareto power distributions. These distributions are generally accepted as good background models. The modeling of sea clutter by generalized Pareto distributions is primarily chosen because samples from clutter show a higher variance than the noise samples and the Pareto distribution allows us to model this higher variance accurately. The simplifications make the problem more tractable but the results not less applicable.
A second general remark is that a point spread function is used around the position of the particles to match the point spread function/impulse response of a reflected signal, as explained in Section 2.3.4. This point spread function plays a role in the likelihood ratio calculations, but in practice, the point spread depends on certain Radar properties that can vary per Radar. Despite the fact that the principle is the same for each Radar, this could lead to slightly different likelihood ratios.

In Chapter 3 it is explained that Wald's theory and Neyman-Pearson theory hold under certain conditions/assumptions. Wald's theory is widely used in sequential likelihood ratio tests, but in Chapter 4 it is discussed that the necessary condition of i.i.d. samples does not hold with filtered data and that therefore Wald's theory does not apply within a TBD approach. A newly designed sequential detection scheme is tested that uses Neyman-Pearson theory. Despite the fact that this theory uses a fixed time horizon as in an LRT, it is constructed such that it could work properly within the new sequential scheme.

In Sections 4.1 and 4.2 observations about the behavior of the particle filter are given together with some solutions to correct certain irregularities. Among others, a method is proposed to model the spread of the particles. This method is useful in the simplified filter in this research, but it needs some adjustments in the threshold calculations, when higher order state spaces are used to calculate likelihood ratios. Also the model for the MCMC step is designed with a scalar input in the likelihood functions and hence also this model needs adjustments in higher order state spaces. Nevertheless, the ideas behind the models do apply in higher order state spaces. Furthermore, the idea to use the predicted likelihood ratios instead of the posterior likelihood ratios (defined in Section 2.2.3) as proposed at the end of Section 4.2.3 is also immediately useful in higher order state spaces and multi target tracking filters.

The results for the proposed models and for the proposed detection scheme in Sections 4.3 and 4.4 show improved results compared to the results of detection schemes with current thresholds. Despite the simple and rough approximation of the linear model for the number of cells covered by the particle cloud in eq.(4.26), the P_{FA} curves for noise in Section 4.3 satisfy the CFAR property. Also the P_{FA} curves in noise for the proposed sequential detection scheme in Section 4.4 satisfy the CFAR property. However, the scaling of the particles with $\lambda^T(x_n)$ with both method 1 and 2 give other P_{FA} values than the values in Section 4.3; the values of scaling with method 1 are higher than the values in Section 4.3 and the values of scaling with method 2 are lower. Furthermore, in the proposed detection scheme, the parameter estimates of the particle cloud is immediately integrated in the calculation of $\lambda^T(x_n)$ where in Section 4.3 the threshold is calculated afterwards based on the mean parameters estimates of all Monte Carlo runs.

Unfortunately, the results of the P_{FA} curves for clutter in Section 4.4 do not satisfy the CFAR property. A number of reasons could cause this undesired behavior. First, as explained in Section 2.3.3, the used pdf for $\mathcal{H}_1$ is an empirical approximated density, because the analytical

expression is unknown/computationally inefficient. Furthermore, in the calculation of $\lambda^T(x_n)$ where also c_n is considered, the Gamma distribution is used to approximate the distribution of the sample mean of generalized Pareto distributed samples. This is not entirely proper as this sample mean is not Gamma distributed. However, the derived expression in Appendix C is computationally inefficient and first has to be investigated on convergence properties. Conspicuous is the fact that the P_{FA} curves in Figure 4.16a are more spread than the (SLRT) curves in Figure 4.14a. This is caused by the higher variance of clutter samples and this shows the difficulty of achieving the CFAR property in clutter. Also conspicuous in Figure 4.16a is the fact that the P_{FA} values for $k = 0.2$ (black) are below the values for $k = 0.1$ (dark blue). No reasonable explanation is found for this phenomenon, but when the spread of the particle cloud is modeled as well (Figure 4.16b), this irregularity is corrected. A last note on Figure 4.16 concerns the slightly different behavior of the P_{FA} curve for k=0 (yellow). This is explained by the fact that the generalized Pareto distribution with k approaching 0, transforms into the exponential distribution, which leads to a different scaling and hence to different P_{FA} values.

5.2 Conclusions

In this section, the conclusions of this research are presented. The conclusions follow from the observations and sub results during the research and the main results from Sections 4.3 and 4.4, earlier discussed in Section 5.1. In this thesis two sequential detection detection schemes, one based on Wald's theory and one newly designed based on the Neyman-Pearson theory, are tested. Both these detection schemes showed proper and expected results when tested with i.i.d. generated data. Both schemes showed undesired results when tested against filtered data, acquired from the simplified SMCMC filter. Thorough analysis of the SMCMC provided us with observations/explanations for these undesired results. These observations lead to the conclusion that Wald's theory is not suitable for applying a sequential likelihood ratio test with filtered data because the important condition for this theory, namely the condition of i.i.d. samples, does not hold.
To properly apply the new sequential detection scheme, methods have been designed to correct or improve the threshold, calculated recursively with Neyman-Pearson theory namely

- It is explained how to recursively update and integrate parameter estimates in threshold calculations.
- A method is designed to integrate the spread of particle cloud in the threshold calculations. To achieve this useful method, the sum of exponential distributed r.v.'s is transformed into a r.v. of the sample mean using properties of the Gamma distribution. Ultimately this method allows us to integrate over time and integrate the spread of the particle cloud simultaneously. Furthermore, an analytical expression for the sample mean of i.i.d. generalized Pareto clutter samples is derived, which is unfortunately not useful in threshold calculations yet.
- In Section 4.2.3 a method is designed to model the MCMC sampling step in the SMCMC, performed by the MH algorithm. It is also discussed that instead of using the designed model, it is advised to use the predicted likelihood ratio, as defined as in Section 2.2.3, rather than the posterior likelihood ratio.

The developed methods are tested separately and the conclusion is that each of them, improved the performance of the detection scheme. When these methods are used collectively, the P_{FA}

curves for Gaussian noise satisfy the CFAR property in Section 4.3. On top of that, it is shown in Section 4.3, that with an empirically tuned number of cells covered by the particle cloud, the P_D can be maximized by satisfying $P_{FA} = \alpha$. This means that the combined methods could lead to the most powerful sequential test, at least for Gaussian noise. The method improves the performance of the detection scheme when tested against filtered clutter samples as well. Unfortunately, the P_{FA} curves for clutter background samples still show variety for different power levels.
Eventually, the methods are integrated within the SMCMC which enables us to sequentially detect, while tracking a possible target. The results of this combined filter and detection scheme fulfill the CFAR property when used in Gaussian noise.

Overall, the conclusion of this research is that a desired sequential CFAR detection scheme is designed and successfully implemented in an SMCMC filter such that it can be use within a TBD approach. The detection scheme is CFAR when tested against Gaussian noise. Additional research is necessary to check the CFAR property when tested against sea clutter.

5.3 Contribution to Thales

After presenting the conclusions in Section 5.2, the contribution of this research to Thales can be summarized. Despite using several simplifications to the filter/models in this research compared to filters used in practical Radar application, there certainly are beneficial aspects for Thales. First, more insights in hypothesis testing theories is provided. It has become clear that Wald's theory, which is a widely used theory in applying sequential hypothesis testing, is not suitable for the purposes of Thales. Furthermore it is shown that the use of Neyman-Pearson theory can be useful in controlling a certain significance level of the detection scheme.

Second, ideas and methods are presented which improve the controllability of the probability of false alarm. Some ideas are immediately useful, such as the idea to use the predicted rather than the posterior likelihood ratio (when using an SMCMC filter). Other ideas need further research, such as a more accurate implementation of the spreading particle cloud.

Third, an expression is derived for the sample mean of c i.i.d. generalized Pareto distributed r.v.'s. This expression can be very useful in implementing the spreading particle cloud. Also research on the distribution of convoluted target+clutter samples has been done. Furthermore, it is shown that the empirical derived pdf for this convolution works well as the distribution for $\mathcal{H}_1$.

The final and most important contribution of this research is the implementation of the CFAR detection scheme within the TBD approach. An advantage is that not the exact threshold needs to be calculated at each time, but a so called threshold likelihood ratio can be used to scale the filtered likelihood ratio. As a result, the detection likelihood ratio $\lambda^D(\mathbf{x}_n)$ can always be tested against threshold $\tau_{new} = 1$ ($\mathrm{T}_{new} = 0$ when $\log(\lambda^D(\mathbf{x}_n))$ is used). The results in Section 4.4 show that the P_{FA} values can be easily controlled in satisfying a certain significance level. Furthermore, in Gaussian noise the P_{FA} values immediately fulfill the desired CFAR property.

5.4 Recommendations for further research

In this section several recommendations for further research are given, based on the experiences in this research. These recommendations are intended for researchers linked to Thales or new

students who continue to work in this subject as Thales.

The first recommendations are based on the different results for noise and clutter in Sections 4.3 and 4.4. Contrary to the CFAR results when applying newly designed sequential detection scheme in noise, applying this scheme in clutter still leads to varying P_{FA} values for different power levels. It is recommended to investigate and improve the ideas and models derived in this thesis. Improvements could be achieved by finding an analytical expression for the pdf of the convolution of target+clutter (if possible). Furthermore, the derived expression for the pdf of the sample mean of i.i.d. clutter samples should be investigated on convergence and computational efficiency. When the conclusion after further research is that efficient and analytically proper pdf's for the convolution and sample mean of clutter samples are not available, it recommended to use other distributions to model the power samples of clutter. The Weibull distribution is a distribution that could be used.

The second recommendation is based on the observations that modeling the spread of the particle cloud heavily improves the P_{FA} curves towards fulfilling the CFAR property. In this thesis a simple linear model is used to predict the particle spread but it is recommended to design a better model. Especially, because it is shown in Figure 4.13 that by empirically tuning the spread of the particle cloud it is possible that Neyman-Pearon theory constructs the most powerful test in an LRT application. A suggestion to start further research on the particle cloud spread, could be the use of a normal distribution over the measurement cells, centered on the cell with the highest particle population. This seems to be an accurate model in practice. It is also recommended to analyze the influence of the measurement model such that also the influence of the point spread function can be integrated in modeling the spreading particle cloud.

The third recommendation involves the MCMC step by the MH algorithm. The output of the MCMC could be a particle cloud with certain correlation, which contradicts the assumption of individually distributed samples. Therefore, it is recommended to investigate the correlation of samples in particle clouds. Strong correlation (relatively) between particles should be another reason the use the predicted rather than the posterior particle cloud for the hypothesis testing.

In Section 4.4 the ideas and models in Section 4.2 are successfully integrated within the SMCMC. However, it is explained that the influence of the scaling by $\lambda^T(x_n)$ is not totally clear, especially when c_n is included in the scaling. Further research is recommended to investigate proper scaling in the proposed detection scheme.

The last recommendation involves the time instants of false alarms. In applying Neyman-Pearson theory in the sequential detection scheme, it is assumed that the probability of a false alarm is constant for all time when calculating the thresholds. This assumption might be not correct in a sequential setting and also the P_{FA} curves of the LRT show that relatively more false alarms occur in the beginning. A better understanding of the time instants of false alarms could lead to better threshold setting.

Overall, further research is recommended to investigate the utility of the newly designed CFAR detection scheme and proposed models in practical Radar application by for example increasing the dimension of the state space, using multi target models and implementing multiple burst per scan, with the ultimate goal of designing a sequential CFAR detection scheme in the TBD approach to detect small targets in clutter with total control of the P_{FA} values.

Bibliography

[1] Samual Blackman and Robert Popoli, "Sensor and Source Chararcteristics", in *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[2] Jin Lu, Peng-Lang Shui, and Hong-Tao Su, "Track-before-detect method based on cost-reference particle filter in non-linear dynamic systems with unknown statistics", *Signal Pocessing, IET*, vol. 8, no. 1, pp. 85–94, Feb. 2014.

[3] K. Jishy and F. Lehmann, "A bayesian track-before-detect procedure for passive radars", *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, 2013.

[4] A. Doucet, C. Andrieu, and W. Fitzgerald, "Bayesian filtering for hidden Markov models via Monte Carlo methods", Aug. 1998, pp. 194–203.

[5] A. Bagchi, *Optimal Control of Stochastic Systems*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[6] Branko Ristic, Sanjeev Arulampalam, and Neil Gordon, "A tutorial on Particle Filters", in *Beyond the Kalman Filter*, Artech House Radar Library, 2004.

[7] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation", *Radar and Signal Processing, IEE Proceedings*, vol. 140, no. 2, pp. 107–113, Apr. 1993.

[8] M. Sanjeev Arulampalam, Simon Maskell, and Neil Gordon, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking", *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.

[9] K. L. Chung, "The Strong Law of Large Numbers", in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Calif.: University of California Press, 1951, pp. 341–352.

[10] Xiao-Li Hu, T. Schon, and L. Ljung, "A Basic Convergence Result for Particle Filtering", *Signal Processing, IEEE Transactions on*, vol. 56, no. 4, pp. 1337–1348, Apr. 2008.

[11] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practitioners", *Signal Processing, IEEE Transactions on*, vol. 50, no. 3, pp. 736–746, Mar. 2002.

[12] N. Bergman, "Recursive Bayesian Estimation: Navigation and Tracking Applications", Tech. Rep., 1999.

[13] O. Cappé, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in Sequential Monte Carlo", *Proceedings of the IEEE*, vol. 95, no. 5, pp. 899–924, May 2007.

[14] S. Brooks, A. Gelman, G. Jones, and X. Meng, *Handbook of Markov Chain Monte Carlo*, ser. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, 2011.

[15] C. Andrieu and E. Moulines, "On the ergodicity properties of some adaptive MCMC algorithms", *Annals of Applied Probability*, vol. 16, no. 3, 2006.

[16] M. Rosenblatt, "Remarks on ergodicity of stationary irreducible transient Markov chains", *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 6, no. 4, pp. 293–301, 1966.

[17] B. Walsh, *Markov Chain Monte Carlo and Gibbs Sampling*, 2004.

[18] Zinan Zhao and M. Kumar, "An MCMC-based particle filter for multiple target tracking", Jul. 2012, pp. 1676–1682.

[19] M. Rutten, N. Gordon, and S. Maskell, "Recursive track-before-detect with target amplitude fluctuations", *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 152, no. 5, pp. 345–352, Oct. 2005.

[20] P. Swerling, "Probability of detection for fluctuating targets", *Information Theory, IRE Transactions on*, vol. 6, no. 2, pp. 269–308, Apr. 1960.

[21] D. Shnidman, "Expanded Swerling target models", *Aerospace and Eectronic Systems, IEEE Transactions on*, vol. 39, no. 3, pp. 1059–1069, Jul. 2003.

[22] T. Sniekers, "Investigation of detection threshold setting", University of Twente, Tech. Rep., 2014.

[23] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses", *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, Feb. 1933.

[24] A. Wald, "Sequential tests of statistical hypotheses", *Ann. Math. Statist.*, vol. 16, no. 2, pp. 117–186, 1945.

[25] A. Wald, "On cumulative sums of random variables", *Ann. Math. Statist.*, vol. 15, no. 3, pp. 283–296, Sep. 1944.

[26] ——, "Some generalizations of the theory of cumulative sums of random variables", *Ann. Math. Statist.*, vol. 16, no. 3, pp. 287–293, Sep. 1945.

[27] S. Kullback and R. Leibler, "Letter to the editor: the Kullback-Leibler distance", *American Statistician*, vol. 41, no. 4, pp. 340–341, 1987.

[28] I. Wolfram Research, *Mathematica*, version 10.1, 2015. [Online]. Available: `https://www.wolfram.com`.

[29] P. J. Rao S. and A. Shukla, "Wright Type Hypergeometric Function and its Properties", *Advances in Pure Mathematics*, vol. 3, no. 3, pp. 335–342, 2013.

[30] R. T. Rockafellar, "Lagrange Multipliers and Optimality", *SIAM Rev.*, vol. 35, no. 2, pp. 183–238, 1993.

[31] W. DeCoursey, *Statistics and Probability for Engineering Applications*. Elsevier Science, 2003.

[32] C. Walck, *Hand-book on STATISTICAL DISTRIBUTIONS for experimentalists*. Dec. 1996.

[33] S. Nadarajah and S. Kotz, "The generalized Pareto sum", *Hydrological Processes*, vol. 22, no. 2, pp. 288–294, 2008.

[34] K. Kawamoto, "Grid-based Rao-Blackwellisation of particle filtering for switching observation models", in *Information Fusion, 2012 15th International Conference on*, Jul. 2012, pp. 143–148.

[35] G. Grisetti, C. Stachniss, and W. Burgard, "Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling", in *ICRA*, 2005, pp. 2443–2448.

[36] Tom M. Apostol, "Mathematical analysis: A modern approach to advanced Calculus", in. Addison-Wesley Publishing, 1964.

[37] Ronald Newbold Bracewell, *The Fourier Transform and Its Applications*. McGraw Hill, 2000., May 2009.

[38] L. L. Cam, "The Central Limit Theorem Around 1935", *Statist. Sci.*, vol. 1, no. 1, pp. 78–91, Feb. 1986.

[39] R. Courant, *Differential and Integral Calculus: Vol.: 1*, v. 1. Blackie & Son Limited, 1937.

Appendix A

Different domain calculations

In Section 2.3.1, the distributions for Gaussian noise and Pareto clutter are given for the I/Q domain, amplitude/phase domain as well as the power domain. In this section the results are recalled and the calculations are presented, first for Gaussian noise, later for Pareto clutter.

A.1 Gaussian Noise

IQ to Amplitude

In Gaussian noise, the I and Q samples are assumed to be identically distributed according to a normal distribution with zero mean:

$$z_I \sim \mathcal{N}(0, \sigma^2), \quad z_I \in \mathbb{R}$$

$$z_Q \sim \mathcal{N}(0, \sigma^2), \quad z_Q \in \mathbb{R}$$

The distribution of the amplitude z_P is given by the Rayleigh distribution

$$p(z_A) = \frac{z_A}{\sigma^2} \mathrm{e}^{-z_A^2/2\sigma^2}, \quad z_A \in \mathbb{R}^+. \tag{A.1}$$

Proof.

Proposition 23. *Let us consider two random variables z_I and z_Q distributed according to a Normal distribution with parameters $\mu = 0$ and $\sigma = \sigma_n$. Then the random variable $z_A := \sqrt{z_I^2 + z_Q^2}$ is distributed according to a Rayleigh distribution.*

Lemma 24. *The general formula for integration by substitution of two independent variables is:*

$$\int\int_{R_x} p(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2 = \int\int_{R_y} f\left(x_1(y_1, y_2), x_2(y_1, y_2)\right) \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \mathrm{d}y_1 \mathrm{d}y_2.$$

The transformation functions, $y_1(x_1, x_2)$ and $y_2(x_1, x_2)$, and their inverses are $x_1(y_1, y_2)$ and $x_1(y_1, y_2)$ are used. The regions R_x and R_y are identical subject to the first being specified in the x_1x_2 plane and the second being specified in the y_1y_2 plane. The Jacobian $\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}$ is defined as:

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{bmatrix} = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1}.$$

Let us denote $x_1 := z_I$, $x_2 := z_Q$, $y_1 := z_A$ and $y_2 := \theta$, then the transformation functions are given by:

$$y_1 = \sqrt{x_1^2 + x_2^2}, \qquad y_2 = \arctan\left(\frac{x_2}{x_1}\right),$$

with inverses

$$x_1 = y_1 \cos(y_2), \qquad x_2 = y_1 \sin(y_2).$$

Hence

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{bmatrix} \cos(y_2) & -y_1 \sin(y_2) \\ \sin(y_2) & y_1 \cos(y_2) \end{bmatrix} = y_1 \cos^2(y_2) + y_1 \sin^2(y_2) = y_1.$$

As a result

$$\begin{aligned} \int\int_{R_x} p(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2 &= \int\int_{R_y} p\left(y_1 \cos(y_2), y_1 \sin(y_2)\right) y_1 \mathrm{d}y_1 \mathrm{d}y_2. \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{Z_I}(x_1) p_{Z_Q}(x_2) \mathrm{d}x_1 \mathrm{d}x_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right) \mathrm{d}x_1 \mathrm{d}x_2, \end{aligned}$$

equals

$$\begin{aligned} \int\int_{R_y} p(y_1, y_2) \mathrm{d}y_1 \mathrm{d}y_2 &= \int_0^{\infty}\int_0^{2\pi} p_{z_A}(y_1) p_\theta(y_2) \mathrm{d}y_1 \mathrm{d}y_2 \\ &= \int_0^{\infty}\int_0^{2\pi} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{y_1^2}{2\sigma^2}\right) y_1 \mathrm{d}y_1 \mathrm{d}y_2. \end{aligned}$$

Hence

$$\begin{aligned} p_{z_A}(y_1) &= \int_0^{2\pi} p(y_1, y_2) \mathrm{d}y_2 \\ &= \int_0^{2\pi} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\frac{y_1^2}{\sigma^2}\right) y_1 \mathrm{d}y_2 \\ &= \frac{y_1}{\sigma^2} \exp\left(-\frac{y_1^2}{2\sigma^2}\right), \quad y_1 \in \mathbb{R}^+, \end{aligned}$$

which is the Rayleigh distribution. □

Amplitude to Power

If the amplitude distribution is given by the Rayleigh distribution (A.1), then the power distribution is given by

$$p(z_P) = \frac{1}{2\sigma^2} \mathrm{e}^{-z_P^2/2\sigma^2}, \quad z_P \in \mathbb{R}^+. \tag{A.2}$$

Proof. Let us consider the Rayleigh distributed random variable X, i.e.

$$X \sim \frac{x\mathrm{e}^{-x^2/2\sigma^2}}{\sigma^2}\mathrm{d}x,$$

and let us apply change of variables $y := x^2$, then $x = \sqrt{y}$ and $\mathrm{d}x = \frac{1}{2\sqrt{y}}\mathrm{d}y$. Hence Y is distributed according to an exponential distribution:

$$Y \sim \frac{(\sqrt{y})\mathrm{e}^{-\left(\sqrt{y}\right)^2/2\sigma^2}}{\sigma^2}\frac{1}{2\sqrt{y}}\mathrm{d}y = \frac{\mathrm{e}^{-y/2\sigma^2}}{2\sigma^2}\mathrm{d}y.$$

□

Power to Amplitude

Now the reversed results are proven, i.e. when the power distribution is exponentially distributed, then the amplitude distribution is the Rayleigh distribution with scale parameter $\sigma = \sqrt{\mu/2}$:

$$p(z_A) = \frac{2z_A}{\mu}\mathrm{e}^{-z_A^2/\mu}. \tag{A.3}$$

Theorem 25. *If X has an exponential distribution with parameter $\lambda = 1/\sigma$, i.e. X has the pdf:*

$$p_X(x) = \lambda \exp\left(-\lambda x\right), \quad x \in \mathbb{R}^+,$$

then $Y := \sqrt{2X\sigma^2\lambda} = \sqrt{2X\sigma}$ is distributed according a Rayleigh distribution with scale parameter σ. The Rayleigh distribution has the pdf:

$$p_X(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x \in \mathbb{R}^+.$$

Proof. Let us consider an exponential random variable X, i.e.

$$X \sim \lambda \exp\left(-\lambda x\right)\mathrm{d}x,$$

and let apply the change of variable $y := \sqrt{2x\sigma^2\lambda}$, then $y^2 = 2x\sigma^2\lambda$ and thus $x = \frac{y^2}{2\sigma^2\lambda}$, and $\mathrm{d}x = \frac{y}{\sigma^2\lambda}\mathrm{d}y$. Hence, Y is distributed according the Rayleigh distribution:

$$Y \sim \lambda \exp\left(-\lambda\frac{y^2}{2\sigma^2\lambda}\right)\frac{y}{\sigma^2\lambda}\mathrm{d}y = \frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right)\mathrm{d}y.$$

□

Amplitude to IQ domain

Finally, when the amplitude is given by the Rayleigh distribution (A.1), then the distribution of the I/Q samples is given by the Normal distribution with zero mean and variance σ^2, i.e.

$$p(z_{I/Q}) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-(z_{I/Q}^2)/2\sigma^2}. \tag{A.4}$$

Proof. Let us use the controversy of Proposition 23:

Proposition 26. *Let us consider a random variable z_A distributed according a Rayleigh distribution with scale parameter $\sigma = \sigma_n$. Let us assume a random phase θ on $[0, 2\pi]$. Then the random variables $z_I := z_A \cos(\phi)$ and $z_Q := z_A \sin(\phi)$ are distributed according a Normal distribution with parameters $\mu = 0$, $\sigma = \sigma_n$.*

Now, let us use Lemma 24 and let us denote by $x_1 := z_A$, $x_2 := \theta$, $y_1 := z_I$ and $y_2 := z_Q$, then the transformation functions are given by:

$$y_1 = x_1 \cos(x_2), \qquad y_2 = x_1 \sin(x_2),$$

with inverses

$$x_2 = \sqrt{y_1^2 + y_2^2}, \qquad x_2 = \arctan\left(\frac{y_2}{y_1}\right).$$

Now the Jacobian becomes

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{bmatrix} \frac{y_1}{\sqrt{y_1^2+y_2^2}} & \frac{y_2}{\sqrt{y_1^2+y_2^2}} \\ -\frac{y_2}{y_1^2+y_2^2} & \frac{y_1}{y_1^2+y_2^2} \end{bmatrix} = \frac{1}{\sqrt{y_1^2 + y_2^2}}.$$

Hence

$$\begin{aligned}
\int\int_{R_x} p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 &= \int\int_{R_y} p\left(\sqrt{y_1^2 + y_2^2}, \arctan\left(\frac{y_2}{y_1}\right)\right) \frac{1}{\sqrt{y_1^2 + y_2^2}}\mathrm{d}y_1\mathrm{d}y_2. \\
&= \int_0^\infty \int_0^{2\pi} p_{z_A}(x_1)p_\theta(x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_0^\infty \int_0^{2\pi} \frac{x_1}{\sigma^2} \exp\left(-\frac{x_1^2}{2\sigma^2}\right) \frac{1}{2\pi}\mathrm{d}x_1\mathrm{d}x_2
\end{aligned}$$

is equal to

$$\begin{aligned}
\int\int_{R_y} p(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 &= \int_{-\infty}^\infty \int_{-\infty}^\infty p_{z_I}(y_1)p_{z_Q}(y_2)\mathrm{d}y_1\mathrm{d}y_2 \\
&= \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{\sqrt{y_1^2 + y_2^2}}{2\pi\sigma^2} \exp\left(-\frac{y_1^2 + y_2^2}{2\sigma^2}\right) \frac{1}{\sqrt{y_1^2 + y_2^2}}\mathrm{d}y_1\mathrm{d}y_2 \\
&= \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{1}{2\pi\sigma^2} \exp\left(-\frac{y_1^2 + y_2^2}{2\sigma^2}\right) \mathrm{d}y_1\mathrm{d}y_2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
p_{z_I}(y_1) = \int_{-\infty}^\infty p(y_1, y_2)\mathrm{d}y_2 &= \int_{-\infty}^\infty \frac{1}{2\pi\sigma^2} \exp\left(-\frac{y_1^2 + y_2^2}{2\sigma^2}\right) \mathrm{d}y_2 \\
&= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{y_1^2}{2\sigma^2}\right) \int_{-\infty}^\infty \exp\left(-\frac{y_2^2}{2\sigma^2}\right) \mathrm{d}y_2 \\
&= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2\sigma^2}\right), \quad y_1 \in \mathbb{R}.
\end{aligned}$$

Similarly,

$$p_{z_Q}(y_2) = \int_{-\infty}^{\infty} p(y_1, y_2)\mathrm{d}y_1 = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{y_1^2 + y_2^2}{2\sigma^2}\right) \mathrm{d}y_1$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{y_2^2}{2\sigma^2}\right), \quad y_2 \in \mathbb{R}.$$

□

A.2 Generalized Pareto

In Section 2.3.1 it is discussed that the power distribution of clutter samples is given by

$$p(z_P) = \frac{1}{s}\left(1 + k\frac{z_P}{s}\right)^{-1-1/k}, \quad k \geq 0, s > 0, \quad z_P \in \mathbb{R}^+.$$

With the same change of variables as in the Gaussian noise case $z_p = z_A^2$ and Jacobian $J = 2z_A$ such that the distribution of the amplitude is given by

$$p(z_A) = \frac{2z_A}{s}\left(1 + k\frac{z_A^2}{s}\right)^{-1-1/k}, \quad k \geq 0, s > 0, \quad z_A \in \mathbb{R}^+$$

and the joint distribution of the amplitude and a random phase $p(z_A, \theta)$ is given by

$$p(z_A, \theta) = \frac{z_A}{s\pi}\left(1 + k\frac{z_A^2}{s}\right)^{-1-1/k}, \quad k \geq 0, s > 0, \quad z_A \in \mathbb{R}^+.$$

Again, the same change of variables is performed to find the joint distribution of the I,Q clutter samples, i.e

$$z_A = \sqrt{z_I^2 + z_Q^2}, \qquad \theta = \arctan(\frac{z_Q}{z_I}),$$

which gives $J = \frac{1}{z_A}$, such that $p(z_I, z_Q)$ is given by

$$p(z_I, z_Q) = \frac{1}{s\pi}\left(1 + k\frac{z_I^2 + z_Q^2}{s}\right)^{-1-1/k}, \quad k \geq 0, s > 0, \quad z_I, z_Q \in \mathbb{R}^+.$$

Now

$$p(z_I) = \int_{-\infty}^{\infty} \frac{1}{s\pi}\left(1 + k\frac{z_I^2 + z_Q^2}{s}\right)^{-1-1/k} \mathrm{d}z_Q \tag{A.5}$$

and

$$p(z_Q) = \int_{-\infty}^{\infty} \frac{1}{s\pi}\left(1 + k\frac{z_I^2 + z_Q^2}{s}\right)^{-1-1/k} \mathrm{d}z_I. \tag{A.6}$$

Let us denote $a_I = (1 + \frac{z_I^2}{s})$, $a_Q = (1 + \frac{z_Q^2}{s})$ to find

$$p(z_I) = \int_{-\infty}^{\infty} p(z_I, z_Q)\mathrm{d}z_Q \tag{A.7a}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\pi s}\left(1 + k\left(\frac{z_I^2 + z_Q^2}{s}\right)\right)^{-1-1/k} \mathrm{d}z_Q \tag{A.7b}$$

$$= \left[\frac{a_I^{-1-1/k}}{\pi s} {}_2F_1\left(\frac{1}{2}, 1 + \frac{1}{k}; \frac{3}{2}; -\frac{k z_Q^2}{s a_I}\right)\right]_{-\infty}^{\infty} \tag{A.7c}$$

$$= \frac{1}{\sqrt{\pi s}}\left(1 + k\left(\frac{z_I^2}{s}\right)\right)^{-1-1/k}, \quad z_I \in \mathbb{R}; \tag{A.7d}$$

and similarly

$$p(z_Q) = \int_{-\infty}^{\infty} p(z_I, z_Q)\mathrm{d}z_I \tag{A.8a}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\pi s}\left(1 + k\left(\frac{z_I^2 + z_Q^2}{s}\right)\right)^{-1-1/k} \mathrm{d}z_I \tag{A.8b}$$

$$= \left[\frac{a_Q^{-1-1/k}}{\pi s} {}_2F_1\left(\frac{1}{2}, 1 + \frac{1}{k}; \frac{3}{2}; -\frac{k z_I^2}{s a_Q}\right)\right]_{-\infty}^{\infty} \tag{A.8c}$$

$$= \frac{1}{\sqrt{\pi s}}\left(1 + k\left(\frac{z_Q^2}{s}\right)\right)^{-1-1/k}, \quad z_Q \in \mathbb{R}; \tag{A.8d}$$

Let us refer to [39] for more calculations regarding these distributions.

Appendix B

Distributions Target + Noise/Clutter

An echo signal does never consist only of target reflections. It also consists of background reflections. In this Appendix, the distributions of target+background reflections are derived. First, the distribution of a target+noise is given. Later, an expression for the distribution of a target+clutter is given together with a useful numerical approximation.

B.1 Target + Noise

Physically, the samples from target and noise are added in the I/Q domain. Hence, first the distribution of target and noise is derived in the I/Q domain. From Section 2.3.1, the target and noise distributions in the I/Q domain are

$$p(z) = \frac{1}{\sigma_i\sqrt{2\pi}}\mathrm{e}^{-z^2/2\sigma_i^2}, \quad i = t, b, \quad z \in \mathbb{R},$$

where σ_t and σ_b are the standard deviation for respectively the target and background. The characteristic function of a normal distributed random variable is given by

$$\varphi_x(t) = \mathbb{E}[\mathrm{e}^{itx}] = \mathrm{e}^{-it\mu-\sigma^2t^2/2}.$$

Since the I/Q samples from the target and noise are independent, $\varphi_{z_t+z_b}(t)$ is given by

$$\varphi_{z_t+z_b}(t) = \varphi_{z_t}(t)\varphi_{z_b}(t) = \mathrm{e}^{-it\mu_t-\sigma_t^2t^2/2} \cdot \mathrm{e}^{-it\mu_b-\sigma_b^2t^2/2} = \mathrm{e}^{-it(\mu_t+\mu_b)-(\sigma_t^2+\sigma_b^2)t^2/2},$$

which is the characteristic function of a normal distribution with mean $\mu_t + \mu_b$ and variance $\sigma_t^2 + \sigma_b^2$. Hence, the distribution of a I/Q sample from a target and noise is given by

$$p(z_{(I/Q)}) = \frac{1}{(\sigma_t + \sigma_b)\sqrt{2\pi}}\mathrm{e}^{-z_{(I/Q)}^2/2(\sigma_t^2+\sigma_b^2)}, \quad z_{(I/Q)} \in \mathbb{R}$$

and from Appendix A, the power distribution for target and noise results in

$$p(z_P) = \frac{1}{\mu_t + \mu_b}\mathrm{e}^{-z_P/\mu_t+\mu_b}, \quad z_P \in \mathbb{R}^+,$$

where $\mu_t + \mu_b$ is the mean of the exponential power distribution.

B.2 Target and Clutter

First the pdf for a target+clutter samples it tried to calculated analytically. Later empirical derived parameter estimates are shown.

B.2.1 Analytical

From Appendix A.2, the distribution of clutter I samples is given by

$$p(z_I) = \frac{1}{\sqrt{\pi s}} \left(1 + k\left(\frac{z_I^2}{s}\right)\right)^{-1-1/k}, \quad z_I \in \mathbb{R}; \tag{B.1}$$

To find the distribution of target+clutter, the characteristic functions are needed. $\varphi_{z_I}(t)$ for clutter is given by

$$\varphi_{z_I}(t) = \mathbb{E}[\mathrm{e}^{itz_I}] \tag{B.2a}$$

$$= \int_{\mathbb{R}} \mathrm{e}^{itz_I} \frac{1}{\sqrt{\pi s}} \left(1 + k\left(\frac{z_I^2}{s}\right)\right)^{-1-1/k} \mathrm{d}z_I \tag{B.2b}$$

$$= \left[\mathrm{e}^{itz_I} \frac{z_I \cdot {}_2F_1\left(\frac{1}{2}; 1+\frac{1}{k}; \frac{3}{2}; \frac{-kz_I^2}{s}\right)}{s}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} it\mathrm{e}^{itz_I} \frac{z_I \cdot {}_2F_1\left(\frac{1}{2}; 1+\frac{1}{k}; \frac{3}{2}; \frac{-kz_I^2}{s}\right)}{s} \mathrm{d}z_I, \tag{B.2c}$$

where ${}_2F_1$ is the Hypergeometric function [29]. This expression is not useful. Therefore the following transform is used to get the pdf for the complex signal $\bar{z}$.

$$\bar{z} = x + iy, \tag{B.3}$$

where

- $x = z_I$,
- $y = z_Q$.

The distribution $p(\bar{z})$ for clutter samples is given by

$$p(\bar{z}) = \frac{1}{\pi s} \left(1 + k\frac{|\bar{z}|^2}{s}\right)^{-1-1/k}, \quad \bar{z} \in \mathbb{C}. \tag{B.4}$$

By comparing eq.(B.4) with eq.(B.1), calculating $\varphi_{\bar{z}}(t)$ directly seems not useful. Instead, let us transform to polar coordinates. Then $\varphi_{\bar{z}}(t)$ is given by

$$\varphi_{\bar{z}}(t) = \mathbb{E}[\mathrm{e}^{it\bar{z}}] = \int_{\mathbb{R}} \mathrm{e}^{it\bar{z}} \frac{1}{\pi s} \left(1 + k\left(\frac{\bar{z}^2}{s}\right)\right)^{-1-1/k} \mathrm{d}\bar{z} \tag{B.5a}$$

$$= \mathbb{E}[\mathrm{e}^{itr}] = \int\int_{\mathbb{R}} \mathrm{e}^{itr} p(r,\theta) \mathrm{d}r\mathrm{d}\theta, \tag{B.5b}$$

where $p(r,\theta)$ for the clutter samples is given by

$$p(r,\theta) = \frac{r}{\pi s}\left(1 + k\frac{r^2}{s}\right)^{-1-1/k}, \quad \theta \in [0, 2\pi], \quad r \in \mathbb{R}^+.$$

Hence $\varphi_{\bar{z}}(t)$ is given by

$$\varphi_{\bar{z}}(t) = \mathbb{E}[\mathrm{e}^{itr}] = \int_0^\infty \int_0^{2\pi} \mathrm{e}^{itr} \frac{r}{\pi s}\left(1 + k\frac{r^2}{s}\right)^{-1-1/k} \mathrm{d}r\mathrm{d}\theta \tag{B.6a}$$

$$= \int_0^\infty \mathrm{e}^{itr} \frac{2r}{s}\left(1 + k\frac{r^2}{s}\right)^{-1-1/k} \mathrm{d}r \tag{B.6b}$$

$$= \left[-\mathrm{e}^{itr}\left(1 + k\frac{r^2}{s}\right)^{-1/k}\right]_0^\infty \tag{B.6c}$$

$$+ \int_0^\infty it\mathrm{e}^{itr}\left(1 + k\frac{r^2}{s}\right)^{-1/k} \mathrm{d}r, \tag{B.6d}$$

where eq.(B.6d) will again lead to the usefulness Hypergeometric function.

B.2.2 Parameter estimation

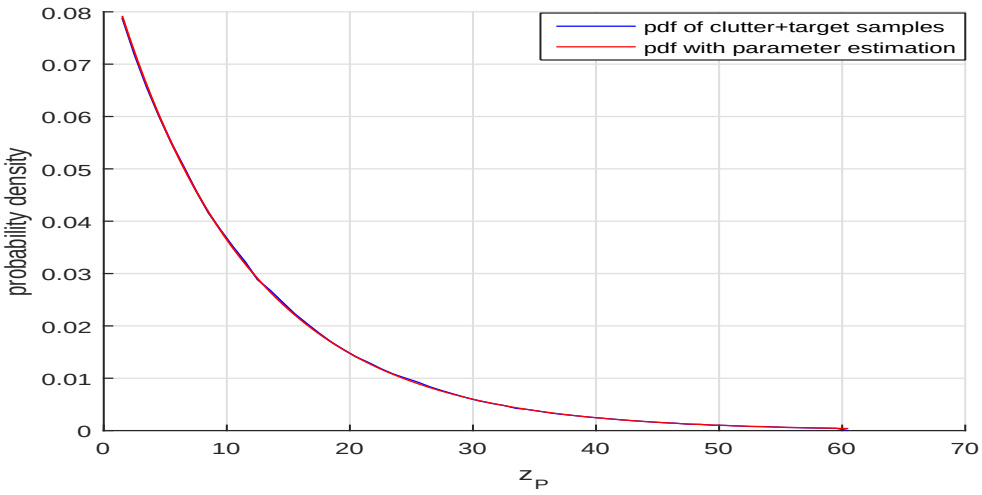
In this Appendix, the accuracy of the parameter estimation proposed in Section 2.3.3 is shown. The proposed pdf for the target+clutter power samples is given by

$$p(z_P) = \frac{1}{\tilde{s}}\left(1 + \tilde{k}\frac{z_P}{\tilde{s}}\right)^{-1-1/\tilde{k}}, \quad z_P \in \mathbb{R}^+ \tag{B.7}$$

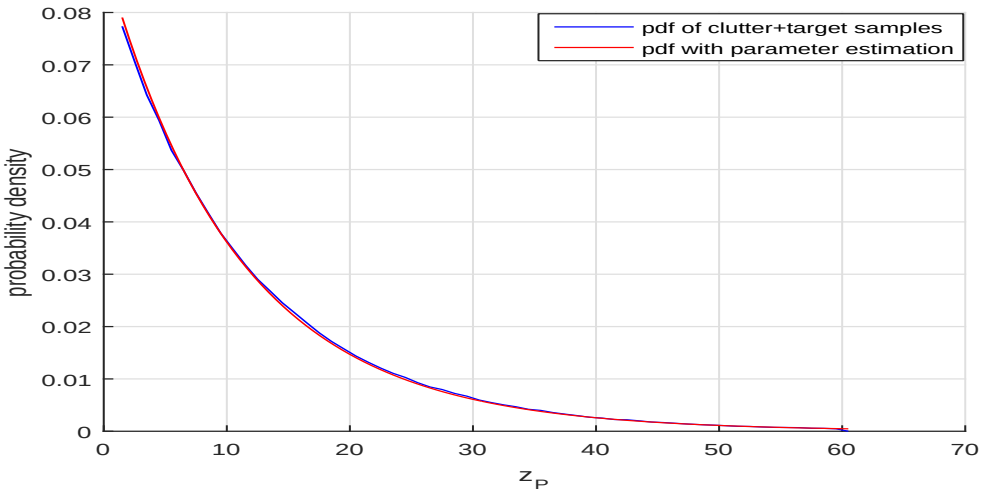
where

- $\tilde{s} = s + \mu_t$,
- $\tilde{k} = \frac{k \cdot s}{\tilde{s}} = \frac{k \cdot s}{s + \mu_t}$.

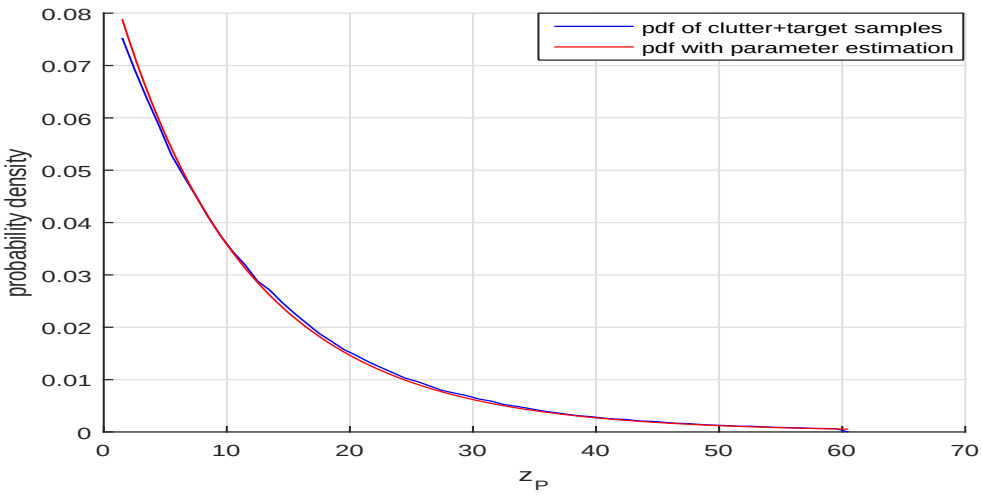
In Figure B.1 the proposed pdf is plotted together with pdf representation of generated target+clutter samples for three values of k,namely 0.1, 0.3, 0.5. The target power is fixed at $\mu_t = 10$. These are settings as used in this thesis. Clearly the proposed parameter estimation is very accurate as all red lines (from parameter estimation pdf) are very close to the blue lines (pdf of generated samples). Furthermore, when k increases the approximation seems to be slightly less accurate.



(a) $\mu_t = 10$, $k = 0.1$



(b) $\mu_t = 10$, $k = 0.3$



(c) $\mu_t = 10$, $k = 0.5$

Figure B.1: Accuracy parameter estimation of distribution target+clutter

Appendix C

Distribution of sample mean of i.i.d Pareto distributed r.v.'s

In this Appendix, the derivation is performed to find an expression for the probability density function of the sample mean of i.i.d. Pareto distributed r.v.'s. The pdf of a single measurement is given by

$$p_X(x) = \frac{1}{s}\left(1 + k\left(\frac{x-\theta}{s}\right)\right)^{-1-1/k} \tag{C.1}$$

First, the characteristic function $\varphi_X(x)$ is calculated.

$$\begin{aligned}
\varphi_X(t) &= \mathbb{E}[\mathrm{e}^{itx}] \\
&= \int_\theta^\infty \frac{\mathrm{e}^{itx}}{s}\left(1 + k\frac{x-\theta}{s}\right)^{-1-\frac{1}{k}} \mathrm{d}x \\
&= \left[-\left(1 + k\frac{x-\theta}{s}\right)^{-\frac{1}{k}} \cdot \mathrm{e}^{itx}\right]_\theta^\infty + \int_\theta^\infty it\left(1 + k\frac{x-\theta}{s}\right)^{-\frac{1}{k}} \cdot \mathrm{e}^{itx}\mathrm{d}x \\
&= \mathrm{e}^{it\theta} + \left[\frac{its\mathrm{e}^{itx}}{k(1-\frac{1}{k})}\left(1 + k\frac{x-\theta}{s}\right)^{1-\frac{1}{k}}\right]_\theta^\infty + \int_\theta^\infty \frac{(it)^2 s\mathrm{e}^{itx}}{(1-k)}\left(1 + k\frac{x-\theta}{s}\right)^{1-\frac{1}{k}} \mathrm{d}x \\
&= \mathrm{e}^{it\theta} + \frac{its\mathrm{e}^{it\theta}}{1-k} + \left[\frac{its^2\mathrm{e}^{itx}}{(1-k)(2k-1)}\left(1 + k\frac{x-\theta}{s}\right)^{2-\frac{1}{k}}\right]_\theta^\infty + \int_\theta^\infty \frac{(it)^3 s^2\mathrm{e}^{itx}}{(1-k)(1-2k)}\left(1 + k\frac{x-\theta}{s}\right)^{2-\frac{1}{k}} \mathrm{d}x
\end{aligned}$$

After expanding these calculations, $\varphi_X(t)$ is given by

$$\varphi_X(t) = \begin{cases} \mathrm{e}^{(it\theta)} \sum_{m=0}^\infty \frac{(its)^m}{\prod_{b=0}^m (1-bk)}, & k > 0, \\ \mathrm{e}^{(it\theta)} \sum_{m=0}^\infty (its)^m, & k = 0. \end{cases} \tag{C.2}$$

Let us use change of variable with $y = x/c \to x = c \cdot y$, $\frac{\mathrm{d}x}{\mathrm{d}y} = c$ to find that the pdf of a weighted sample is

$$p_Y(y) = \frac{c}{s}\left(1 + k\left(\frac{cy-\theta}{s}\right)\right)^{-1-1/k} \tag{C.3}$$

and that $\varphi_Y(t)$ is given by

$$\varphi_Y(t) = \begin{cases} \mathrm{e}^{(it\theta/c)} \sum_{m=0}^{\infty} \frac{(its/c)^m}{\prod_{b=0}^{m}(1-bk)}, & k > 0, \\ \mathrm{e}^{(it\theta/c)} \sum_{m=0}^{\infty} (its/c)^m, & k = 0. \end{cases} \tag{C.4}$$

Let use apply the following lemma to find $\varphi_{\bar{Y}}(t) = \varphi_{\sum_\ell y_\ell}(t) = \prod_\ell \varphi_Y(y_\ell)$

Lemma 27 (Cauchy product)**.** *Let a_n and b_n be terms of two series, then the Cauchy product is defined by*

$$\left(\sum_{n=0}^{\infty} a_n\right) \cdot \left(\sum_{m=0}^{\infty} b_m\right) = \sum_{j=0}^{\infty} c_j, \quad \textit{where} \quad c_j = \sum_{k=0}^{j} a_k b_{j-k}, \quad n = 0, 1, 2, \ldots$$

For $c = 2$, this results in

$$\begin{aligned} \varphi_{\bar{Y}}(t) &= \left(\mathrm{e}^{(it\theta/2)} \sum_{m=0}^{\infty} \frac{(its/2)^m}{\prod_{b=0}^{m}(1-bk)}\right)^2 \\ &= \mathrm{e}^{it\theta} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{m_1} \frac{(its/2)^{m_2}(its/2)^{m_1-m_2}}{\prod_{b_1=0}^{m_2}(1-b_1k)\prod_{b_2=0}^{m_1-m_2}(1-b_2k)} \\ &= \mathrm{e}^{it\theta} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{m_1} \frac{(its/2)^{m_1}}{\prod_{b_1=0}^{m_2}(1-b_1k)\prod_{b_2=0}^{m_1-m_2}(1-b_2k)}. \end{aligned} \tag{C.5}$$

For $c = 3$, this results in

$$\begin{aligned} \varphi_{\bar{Y}}(t) &= \left(\mathrm{e}^{(it\theta/3)} \sum_{m=0}^{\infty} \frac{(its/3)^m}{\prod_{b=0}^{m}(1-bk)}\right)^3 \\ &= \mathrm{e}^{it\theta} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{m_1} \sum_{m_3=0}^{m_2} \frac{(its/3)^{m_1}}{\prod_{b_1=0}^{m_3}(1-b_1k)\prod_{b_2=0}^{m_2-m_3}(1-b_2k)\prod_{b_3=0}^{m_1-m_2}(1-b_3k)}. \end{aligned} \tag{C.6}$$

Expanding these results for general c yields

$$\varphi_{\bar{Y}}(t) = \mathrm{e}^{it\theta} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{m_1} \cdots \sum_{m_c=0}^{m_{c-1}} \frac{(its/c)^{m_1}}{\prod_{b_1=0}^{m_c}(1-b_1k)\prod_{b_2=0}^{m_{c-1}-m_c}(1-b_2k)\ldots\prod_{b_c=0}^{m_1-m_2}(1-b_ck)}. \tag{C.7}$$

The inverse Fourier transform is used to find the pdf of the mean sample c i.i.d. clutter samples $p_{\bar{Y}}(y)$ by

$$p_{\bar{Y}}(y) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-iny}\varphi_{\bar{Y}}(n)\mathrm{d}n.$$

The derivation is performed by calculating the terms of the expansion eq.(C.5) individually, starting with $(m_1 = 0)$. Hence, for $c = 2$, the first term yields

$$\begin{aligned}
p_{(m_1=0)}(y) &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-iny}\mathrm{e}^{in\theta}\mathrm{d}n \\
&= \frac{1}{2\pi}\left[\frac{1}{i(\theta-y)}\mathrm{e}^{in(\theta-y)}\right]_0^{2\pi} \\
&= \frac{1}{2\pi}\left[\frac{\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)} - \frac{1}{i(\theta-y)}\right] \\
&= \frac{\mathrm{e}^{2\pi i(\theta-y)} - 1}{2\pi i(\theta-y)}
\end{aligned}$$

The second term $(m_1 = 1)$ is

$$\begin{aligned}
p_{(m_1=1)}(y) &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-iny}\mathrm{e}^{in\theta}\left(\frac{1}{(1-k)} + \frac{ins}{2(1-k)}\right)\mathrm{d}n \\
&= \frac{1}{2\pi}\left(\int_0^{2\pi}\frac{\mathrm{e}^{in(\theta-y)}}{(1-k)}\mathrm{d}n + \int_0^{2\pi}\frac{ins\mathrm{e}^{in(\theta-y)}}{2(1-k)}\mathrm{d}n\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{in(\theta-y)}}{i(\theta-y)(1-k)}\right]_0^{2\pi} + \left[\frac{ns\mathrm{e}^{in(\theta-y)}}{2(\theta-y)(1-k)}\right]_0^{2\pi} - \int_0^{2\pi}\frac{s\mathrm{e}^{in(\theta-x)}}{2(\theta-y)(1-k)}\mathrm{d}n\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)(1-k)}\right] + \left[\frac{2\pi s\mathrm{e}^{2\pi i(\theta-y)}}{2(\theta-y)(1-k)}\right] - \left[\frac{s\mathrm{e}^{in(\theta-x)}}{2i(\theta-x)^2(1-k)}\right]_0^{2\pi}\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)(1-k)}\right] + \left[\frac{2\pi s\mathrm{e}^{2\pi i(\theta-y)}}{2(\theta-y)(1-k)}\right] - \left[\frac{s\mathrm{e}^{2\pi i(\theta-y)}}{2i(\theta-y)^2(1-k)}\right] + \left[\frac{s}{2i(\theta-x)^2(1-k)}\right]\right) \\
&= \sum_{m_2=0}^{1}\frac{\left[(-1)^{1+1} + (-1)^{1+m_2}(2\pi)^{m_2}(\theta-y)^{m_2}\mathrm{e}^{2\pi i(\theta-y)}\right]\sum_{m_2=0}^{1}\frac{(s/2)}{\prod_{b_1=0}^{m_2}(1-b_1k)\prod_{b_2=0}^{1-m_2}(1-b_2k)}}{2\pi i(\theta-y)^{1+1}}
\end{aligned}$$

The third term ($m_1 = 2$) is

$$
\begin{aligned}
p_{(m_1=2)}(y) &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-iny}\mathrm{e}^{in\theta}\left(\frac{1}{(1-k)(1-2k)} + \frac{(ins/2)}{(1-k)^2} + \frac{(ins/2)^2}{(1-k)(1-2k)}\right)\mathrm{d}n \\
&= \frac{1}{2\pi}\left(\int_0^{2\pi}\frac{\mathrm{e}^{in(\theta-y)}}{(1-k)(1-2k)}\mathrm{d}n + \int_0^{2\pi}\frac{(ins/2)\mathrm{e}^{in(\theta-y)}}{(1-k)^2}\mathrm{d}n + \int_0^{2\pi}\frac{(ins/2)^2\mathrm{e}^{in(\theta-y)}}{(1-k)(1-2k)}\mathrm{d}n\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{in(\theta-y)}}{i(\theta-y)(1-k)(1-2k)}\right]_0^{2\pi} + \left[\frac{(ns/2)\mathrm{e}^{in(\theta-y)}}{(\theta-y)(1-k)^2}\right]_0^{2\pi} + \left[\frac{i(ns/2)^2\mathrm{e}^{in(\theta-y)}}{(\theta-y)(1-k)(1-2k)}\right]_0^{2\pi}\right) \\
&- \frac{1}{2\pi}\left(\int_0^{2\pi}\frac{(s/2)\mathrm{e}^{in(\theta-y)}}{(\theta-y)(1-k)^2}\mathrm{d}n + \int_0^{2\pi}\frac{in(s/2)^2\mathrm{e}^{in(\theta-y)}}{(\theta-y)(1-k)(1-2k)}\mathrm{d}n\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)(1-k)(1-2k)}\right] + \left[\frac{\pi s\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)(1-k)^2}\right] + \left[\frac{i(2\pi s/2)^2\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)(1-k)(1-2k)}\right]\right) \\
&- \frac{1}{2\pi}\left(\left[\frac{(s/2)\mathrm{e}^{in(\theta-y)}}{i(\theta-y)^2(1-k)^2}\right]_0^{2\pi} + \left[\frac{n(s/2)^2\mathrm{e}^{in(\theta-y)}}{(\theta-y)^2(1-k)(1-2k)}\right]_0^{2\pi} - \int_0^{2\pi}\frac{(s/2)^2\mathrm{e}^{in(\theta-y)}}{(\theta-y)^2(1-k)(1-2k)}\mathrm{d}n\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)(1-k)(1-2k)}\right] + \left[\frac{\pi s\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)(1-k)^2}\right] + \left[\frac{i(2\pi s/2)^2\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)(1-k)(1-2k)}\right]\right) \\
&- \frac{1}{2\pi}\left(\left[\frac{(s/2)\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)^2(1-k)^2}\right] + \left[\frac{2\pi(s/2)^2\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)^2(1-k)(1-2k)}\right] - \left[\frac{(s/2)^2\mathrm{e}^{in(\theta-y)}}{i(\theta-y)^3(1-k)(1-2k)}\right]_0^{2\pi}\right) \\
&= \frac{1}{2\pi}\left(\left[\frac{\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)(1-k)(1-2k)}\right] + \left[\frac{\pi s\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)(1-k)^2}\right] + \left[\frac{i(2\pi s/2)^2\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)(1-k)(1-2k)}\right]\right) \\
&- \frac{1}{2\pi}\left(\left[\frac{(s/2)\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)^2(1-k)^2}\right] + \left[\frac{2\pi(s/2)^2\mathrm{e}^{2\pi i(\theta-y)}}{(\theta-y)^2(1-k)(1-2k)}\right]\right) \\
&+ \left(\left[\frac{(s/2)^2\mathrm{e}^{2\pi i(\theta-y)}}{i(\theta-y)^3(1-k)(1-2k)}\right] - \left[\frac{(s/2)^2}{i(\theta-y)^3(1-k)(1-2k)}\right]\right) \\
&= \sum_{m_3=0}^{2}\frac{\left[(-1)^{2+1} + (-1)^{2+m_3}(2\pi i)^{m_3}(\theta-y)^{m_3}\mathrm{e}^{2\pi i(\theta-x)}\right]\sum_{m_3=0}^{2}\frac{(s/2)^2}{\prod_{b_1=0}^{m_3}(1-b_1k)\prod_{b_2=0}^{2-m_3}(1-b_2k)}}{2\pi i(\theta-y)^{2+1}}
\end{aligned}
$$

Expanding the calculations over all terms, the expression for the pdf for 2 i.i.d. Pareto r.v.'s can be written as

$$
p_{\bar{Y}}(y) = \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{m_1}\frac{\left[(-1)^{m_1+1} + (-1)^{(m_1+m_2)}(2\pi i(\theta-y))^{m_2}\mathrm{e}^{2\pi i(\theta-y)}\right]\cdot\sum_{m_2=0}^{m_1}\frac{(s/2)^{m_1}}{\prod_{b_1=0}^{m_2}(1-b_1k)\prod_{b_2=0}^{m_1-m_2}(1-b_2k)}}{2\pi i(\theta-y)^{m_1+1}}.
$$

Using that for general c, $\varphi_{\bar{Y}}(t)$ is given by eq.(C.7), he final expression for the pdf of c i.i.d. samples is

$$p_{\bar{Y}}(y) = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{m_1} \frac{\left[(-1)^{m_1+1} + (-1)^{m_1+m_2}(2\pi i)^{m_2}(\theta - y)^{m_2} \mathrm{e}^{2\pi i(\theta - y)}\right](\bigstar)}{2\pi i(\theta - y)^{m_1+1}},$$

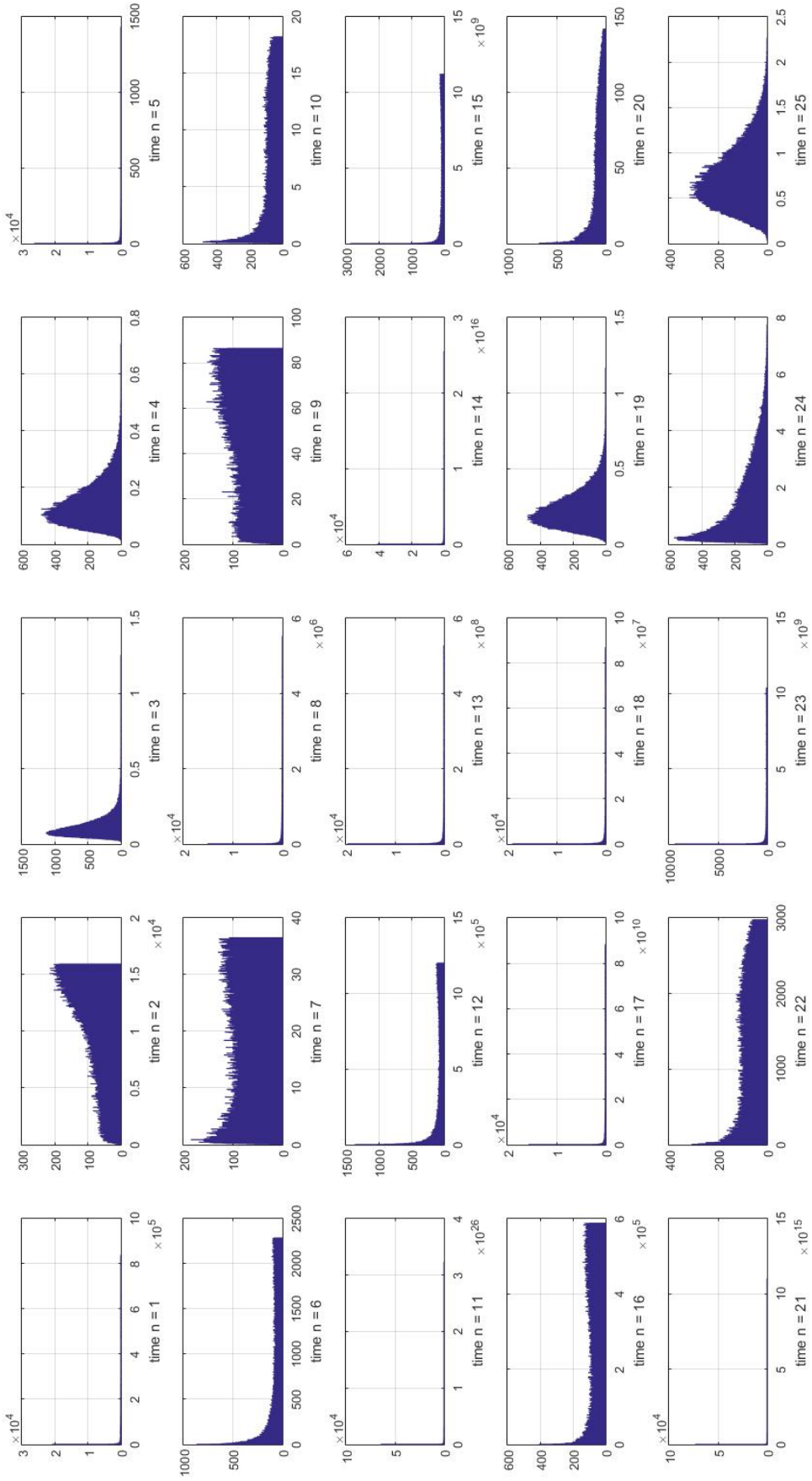
with

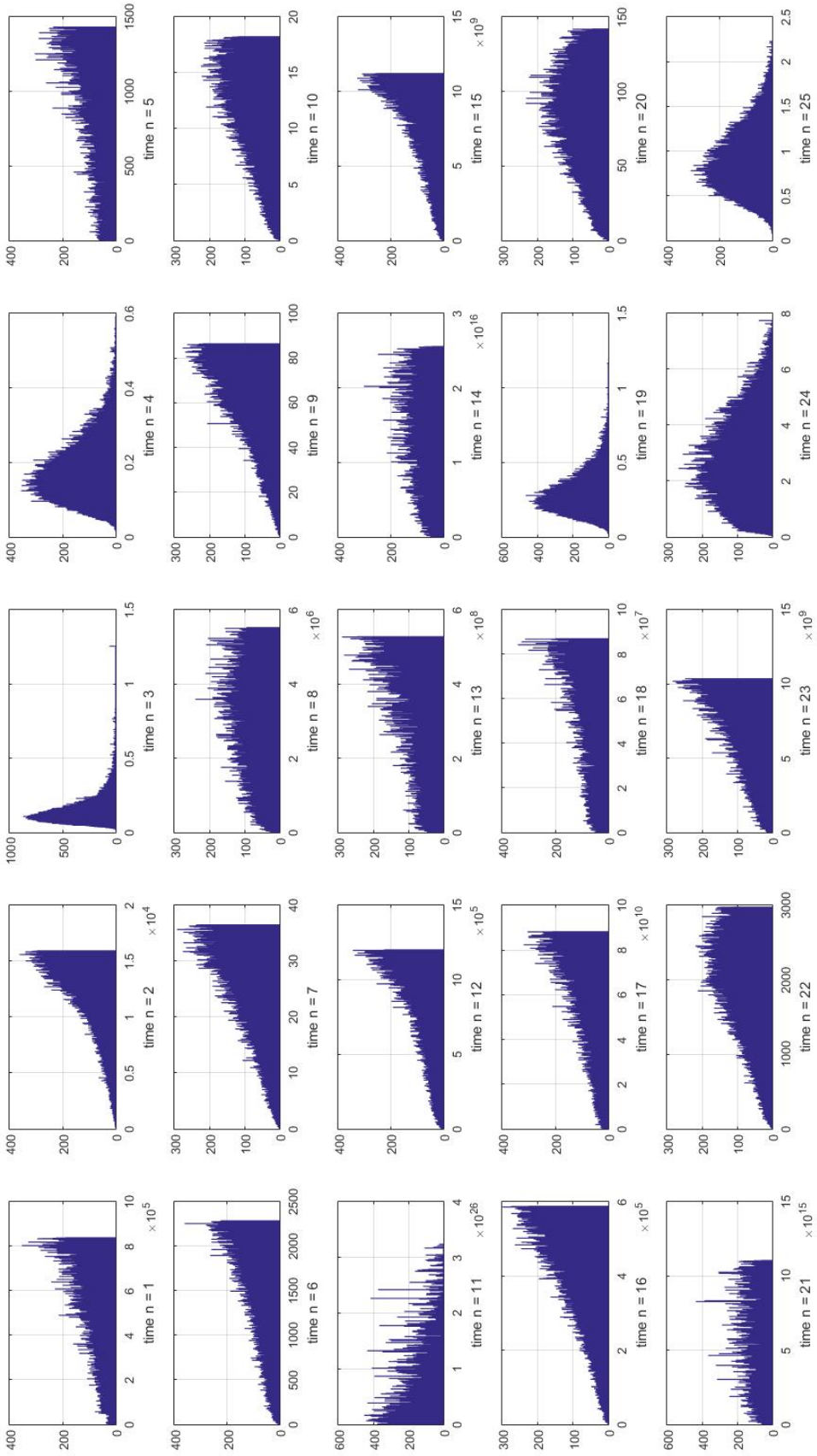
$$(\bigstar) = \sum_{m_3=0}^{m_2} \sum_{m_4=0}^{m_3} \cdots \sum_{m_c=0}^{m_c-1} \frac{(s/c)^{m_1}}{\prod_{b_1=0}^{m_c}(1 - b_1 k) \prod_{b_2=0}^{m_{c-1}-m_c}(1 - b_2 k) \ldots \prod_{b_c=0}^{m_1-m_2}(1 - b_c k)}.$$

Appendix D

Histograms from the Filter

In Figures D.1 and D.2, the likelihood ratio distribution of the predicted and posterior particle cloud is shown during one Monte Carlo run of the SMCMC with 25 time steps. A few plots will be discussed to make the model in Section 4.2.3 more rationalizable. We pick the figures from the predicted distribution with the most similarities to one of the distributions of Figure 4.8 and check whether or not the corresponding posterior distribution also shows similarities with the corresponding distribution in Figure 4.9.
First of all, some plots for the predicted distribution look empty (time 1,5,8 etc.). At these times, a few particles had very high measurements (possibly from a target), but most particles had not. In this case the histogram is almost empty because of the equal length of the segments s^i.
At some time steps, the predicted distribution show similarities with the uniform distribution in Figure 4.8a, for example at time 6,10,16 etc. Also the posterior distribution at these times show similarities with the distribution in Figure 4.9a. Furthermore, at time $n = 24$, the predicted distribution looks similar to the exponential distribution in Figure 4.8b. Also the posterior distribution has some similarities to the corresponding posterior distribution in Figure 4.9b, although the posterior looks not as similar to Figure 4.9b as in the previous case. This is because the histogram at $n = 24$ shows more high samples than the density in Figure 4.8b. At last, the plots for times $n = 4, 19$ and 25 have similarities to Figure 4.8c and also the corresponding posterior distribution looks similar to Figure 4.9c.


time n = 1
time n = 2
time n = 3
time n = 4
time n = 5
time n = 6
time n = 7
time n = 8
time n = 9
time n = 10
time n = 11
time n = 12
time n = 13
time n = 14
time n = 15
time n = 16
time n = 17
time n = 18
time n = 19
time n = 20
time n = 21
time n = 22
time n = 23
time n = 24
time n = 25


time n = 1
time n = 2
time n = 3
time n = 4
time n = 5
time n = 6
time n = 7
time n = 8
time n = 9
time n = 10
time n = 11
time n = 12
time n = 13
time n = 14
time n = 15
time n = 16
time n = 17
time n = 18
time n = 19
time n = 20
time n = 21
time n = 22
time n = 23
time n = 24
time n = 25