UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DIGITAL-AT-EVERY-ELEMENT RADAR RESOURCE ALLOCATION FOR MULTI-TARGET TRACKING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

DAVID JOSEPH LUCKING

Norman, Oklahoma

2019

DIGITAL-AT-EVERY-ELEMENT RADAR RESOURCE ALLOCATION
FOR MULTI-TARGET TRACKING

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

Dr. Nathan A. Goodman, Chair

Dr. Hjalti H. Sigmarsson

Dr. Caleb J. Fulton

Dr. Mark B. Yeary

Dr. John K. Antonio

This dissertation is dedicated to my family.

Acknowledgments

Due to limitations on identifying personal information, I am not able to directly acknowledge everyone that has helped me. I sincerely apologize for not being able to directly thank everyone that has affected me and helped me reach this point, but I am extremely thankful to have such great family, friends, and colleagues to support me.

First, I would like to sincerely thank my research advisor Professor Nathan A. Goodman for your mentorship throughout my time as a graduate student, and for continuing to answer my phone calls while working remotely. Your guidance and support was instrumental in making this dissertation a success.

I would also like to extend my gratitude to my remaining committee members Professor Hjalti H. Sigmarsson, Professor Caleb J. Fulton, Professor Mark B. Yeary, and Professor John K. Antonio for your support, recommendations, and advice.

I would also like to thank my family for their continued support. Your love and care has developed me into the person I am today. All of my accomplishments through life, including this work, is a reflection of the support and love that they have given me.

Finally, I would like to thank all my colleagues, both in Oklahoma and Ohio. Your support, friendship, and advice has helped me stay sane during this process.

Table of Contents

List of Tables

List of Figures

Abstract

A sensor's performance is constrained by the amount of resources at its disposal and the utilization of those resources. A radar system, for example, has a limited amount of transmit power-aperture per unit time to track a multitude of targets. A typical approach when tracking multiple dynamic targets is to time interleave the update intervals until all the radar tasks are performed. The advent of more agile sensors, such as digital-at-every-element apertures, opens the possibility for dynamic sensor resource allocation strategies to achieve better tracking performance in target-dense, resource-constrained scenarios. With proper research into aperture allocation, such as the analysis provided in this dissertation, an all-digital radar can intelligently exploit the degrees of freedom offered by all-digital radars to increase tracking performance. In this dissertation, we investigate adaptive aperture allocation for tracking a large number of targets. The strategies are first introduced with a parallel, linear channel model, then increased in realism with a non-linear measurement model, and finally applied to a full tracking system. We derive various strategies for allocating power and aperture, and compare their performance based on tracking related metrics. Finally, we investigate the relationship between the aperture allocation strategies and the target locations for multiple scenarios designed to represent the environment for a radar tracking system. This research provides groundbreaking strategies for optimal radar aperture

allocation using the digital-at-every-element architectures to reduce the overall system uncertainty and decrease the uncertainty on a per-target basis. Integrating aperture allocation with the management of other degrees of freedom will increase multi-target tracking performance well beyond the current state of the art.

Chapter 1

Introduction

1.1 Overview

Active sensors, such as radar systems, illuminate their surroundings and process the reflected signal to gain knowledge of the environment. Commonly, a radar system is capable of observing and estimating a moving target's location and radial velocity through signal processing techniques. By consistently observing these characteristics, the system can combine multiple measurements to track the target's characteristics, or state, over time. The accuracy of the system's track estimate depends on resolution, signal-to-noise ratio (SNR), and update rate. Unfortunately, these radar properties are related and cannot all be improved arbitrarily. For example, to maximize the update rate, the radar's transmit energy could be divided to illuminate every target during every update interval. By splitting the energy across targets, the individual target SNR will not achieve the maximum possible compared to illuminating the same target with the full antenna aperture. Therefore, a compromise is necessary to achieve the best results for the required application.

Fortunately, modern radar architectures are increasingly able to reconfigure and achieve this compromise dynamically based on the scenario. Arbitrary

waveform generators [1] allow waveform diversity for reconfigurable SNR or bandwidth while frequency variability [2] provides interference avoidance or circumvention of target radar cross section (RCS) fluctuations [3], [4]. In general, more agile radar systems are necessary to address the increasingly congested frequency spectrum and environments inundated with targets. For example, the next-generation, digital at every element, or all-digital, radar architectures [5] will extend system adaptability to include independent waveform synthesis and signal digitization at every antenna element. Contrary to conventional phased array radars, which are limited in the number and diversity of simultaneously formed beams [6], the independent elemental control of all-digital radar systems allows the aperture to be arbitrarily configured into sub-arrays, as well as the ability to perform even more sophisticated waveform synthesis techniques. Therefore, the all-digital architecture can control the gain and power allocated per target by configuring the aperture and forming beams using a variable number of elements.

Thus, the ability of radar systems to reconfigure is a promising attribute toward achieving the desired operating performance in highly congested and dense target environments. As these system architectures are emerging, the strategies and algorithms to configure these architectures for optimal sensing resource allocation are also still in their infancy. Techniques, such as particle filtering, and metrics, such as mutual information and quality-of-service, have been proposed [7]–[10] to facilitate dynamic resource allocation in multifunction radar systems. In [11], a radar network's target tracking performance was improved by providing feedback to the network's fusion center, and a multiple beam power distribution algorithm in [12] exploits prior knowledge to increase multi-target tracking performance. In [13], adaptive transmit wave-

form techniques based on previous observations aided in better focused beams to increase estimation accuracy.

Along with the traditional phased array mode, all-digital radars also have the potential to dynamically reconfigure for multiple input, multiple output (MIMO) or distributed radar modes. Many papers have researched adaptive allocation for MIMO and distributed systems [14]–[19]. In [20], the mutual information was maximized by measuring targets in an adaptive sequence. Similarly, in [21], QR decomposition, using a waterfilling approach to calculate the optimal weights in R and to enforce the total power constraint in Q, was used to determine the power allocation for a linear Gaussian measurement model, represented in matrix form.

Many previous efforts have researched waveform design using information theoretic optimization techniques [22]–[27]. In [22], an information theoretic approach was used to design the waveform for measuring targets that do not behave as simple point targets, which were modeled as Gaussian random processes. In [23], the mutual information metric and the enemy intercept performance were combined for low probability of intercept radar into a single function using the Kullback-Leibler divergence, which can be optimized using an interior point and sequential quadratic programming method. In [24], the mutual information criteria was used to optimally design the waveform for a joint radar/communication system without degrading the communication system performance. By maximizing the mutual information in [24], it was shown that increasing mutual information does not necessarily increase detection performance. Although information-based waveform optimization aims to exploit sensor reconfigurability for improved radar performance, this dissertation focuses on reconfiguring the aperture spatially and does not consider

the performance benefits of adapting the temporal waveform.

This dissertation is motivated by the potential of all-digital radars to adaptively reconfigure their apertures and form an arbitrary number of beams on transmit. To abstract this architecture into mathematical frameworks suitable for optimization, multiple simplified models are used to investigate the performance of adaptively allocated radar resources. Through these simplified models, we examine the trade-offs between different dynamic aperture allocation approaches for differing target properties and dynamics. The aperture allocation strategies proposed are based on the potential of the all-digital, reconfigurable radar architecture, and are derived using multiple information theoretic metrics. Optimized metrics include mutual information, mean-squared error (MSE), the number of lost targets, and the number of deleted tracks. These metrics compare the effect of dynamically allocating transmit elements over multiple targets on system error. This dissertation also examines the effect of allocating available sensing resources to counter the increase in parameter uncertainty due to target dynamics.

1.2 Research Objective

This dissertation introduces multiple useful measurement and estimation models for multi-target radar tracking systems. The measurement and estimation models are used to derive and compare multiple innovative aperture allocation strategies for reconfigurable tracking radars. Then these strategies are compared for contested and target-dense environments using an overall system uncertainty metric, the number of targets lost by the radar system metric, and the number of deleted tracks metric.

1.3 Executive Summary

This dissertation combines concepts, such as the parallel channel model [28], the Kalman filter [29], and the non-linear measurement behavior [30], from the literature with established mathematical concepts, such as Lagrangian multipliers [31] and Bayesian estimation [32], to derive novel approaches to allocate an all-digital aperture. These innovative approaches and the accompanying original analysis and results extend the current state-of-the-art in radar resource allocation to include the additional spatial aperture degree of freedom. Many traditional metrics and terms, such as MSE, entropy, and mutual information [28], are included with the new lost targets (i.e. failed channels) and deleted tracks metrics to compare the approaches, and provide original results. The non-linear representation for angle-of-arrival and range with the given specifications in Chapters 6 and 2, and the results in Sections 4.3, 5.3, 6.3, and 7.2 are original contributions by this publication.

1.4 Outline of the Dissertation

Following the Introduction, Chapter 2 describes the operation of the Kalman Filter for tracking multiple targets in Gaussian distributed noise. Chapter 3 explains how information theory can be used to model the operation of active sensors, with an emphasis on radar systems for multi-target tracking. Chapter 4 details the first, and most simplistic, radar model, approaches for power allocation, and the results of each allocation method. Chapter 5 extends the first radar model to include the possibility of coordinated movement between the targets by including a covariance matrix in place of independent target dynamics. The same approaches for power allocation as in Chapter 4 are

extended to include the additional target correlation, and the results of each allocation method are provided.

Chapter 6 replaces the constant noise power model from Chapter 4 with a more realistic variable noise power model that calculates the measurement error variance for each measurement as a non-linear look up table (LUT) dependent on SNR. The aperture allocation approaches and the results from each strategy are also explained in detail. Chapter 7 includes the non-linear measurement error variance LUT from Chapter 6 as the input data for a Kalman Filter tracker to demonstrate actual tracker performance in the Cartesian coordinate system for the proposed allocation approaches. Finally, Chapter 8 concludes the dissertation.

Chapter 2

Kalman Filter

2.1 Introduction

A common goal for a sensor such as a radar is to determine the location of a target. These targets move over time requiring the sensor to periodically update the location estimate. To overcome the noise included in each observation and achieve a more accurate estimate, a sensor can combine multiple measurements together to track the target over time. This amalgamation of measurements is referred to as tracking due to its ability to keep track of the target, or filtering due to its smoothing effect over time [33]. More generally, the filtering, or tracking, technique can track any parameter state; for example, target location and velocity. The Kalman filter is the most widely used approach to combine these measurements for state tracking over time [33]. While this chapter focuses on tracking a single parameter's state over time, these techniques can be implemented in parallel to track multiple parameters simultaneously. The Kalman filter technique is used in many different fields, including biomedical signal processing, space vehicles reentry, aircraft, car, and missile tracking, navigation, guidance, altitude control, sonar, radar, and others [34].

The Kalman filter's popularity is due to its measurement model simplicity and its effectiveness on Gaussian processes with additive white Gaussian noise [33]. A process, or state variable, is Gaussian if the process' values, or the target's movement, vary according to a Gaussian distribution around a central point. For example, an airplane moving in a straight line would actually vary along the line depending on the steadiness of the pilot, the control's exactness, the environmental effects of the wind and air resistance, etc. All of these factors contribute to a Gaussian distributed movement along the desired trajectory. Along with the process' Gaussian distribution, the Kalman filter also accounts for Gaussian distributed noise included in each measurement. The measurement noise induces a Gaussian distributed estimate error around the correct process value. The Kalman filter assumes that these errors do not depend on the process' state.

While many adaptations exist to apply the Kalman filter technique to different models [35]–[39], we focus on the discrete-time, linear-Gaussian Kalman filter. The discrete-time Kalman filter performs measurements and subsequent estimation updates at distinct, separate points in time. For a radar system, these points in time correspond to each coherent processing interval (CPI). After each measurement is received, the Kalman filter calculates the new estimate using a recursive approach by predicting the process' state and then combining that prediction with the received measurement. The Kalman filter measurement model is

$$\mathbf{p}_k = \mathbf{M}\mathbf{q}_k + \mathbf{r}_k, \tag{2.1}$$

where $\mathbf{p}_k$ is the vector of measurements at time step k, $\mathbf{M}$ is the observation matrix, $\mathbf{q}_k$ is the actual state vector at time step k, and $\mathbf{r}_k$ is the vector of zero mean Gaussian noise with covariance matrix $\mathbf{R}$ at time step k. Since the

measurements are captured at discrete time steps, the change in state between each set of measurements must be modeled. The tracker predicts this state change by leveraging the intimate relationship between state variables and applying a state transition matrix. For the target movement tracker example, the state transition matrix exploits the interaction between the motion state variables (location, velocity, acceleration, etc.). This state transition is applied using

$$\tilde{\mathbf{q}}_k = \mathbf{\Phi}\hat{\mathbf{q}}_{k-1}, \tag{2.2}$$

where $\mathbf{\Phi}$ is the state transition matrix, $\hat{\mathbf{q}}_{k-1}$ is the state value estimate at time step $k-1$, and $\tilde{\mathbf{q}}_k$ is the predicted state at time step k.

In reality, the state transition between measurements also include a dynamic noise caused by environmental and external factors. Therefore, the tracker includes a dynamic noise when calculating the state covariance matrix. By estimating both the parameter state and parameter covariance matrix through the same iterative approach, the tracker is able to completely characterize a random process [40] at a single time step. The corresponding covariance prediction equation is

$$\tilde{\mathbf{S}}_k = \mathbf{\Phi}\mathbf{S}_{k-1}\mathbf{\Phi}^T + \mathbf{D}, \tag{2.3}$$

where $\tilde{\mathbf{S}}_k$ is the predicted parameter covariance matrix estimate at time step k. ^{T} is the transpose operator, $\mathbf{D}$ is the parameter dynamic noise matrix, and $\mathbf{S}_{k-1}$ is the parameter covariance matrix estimate at time step $k-1$.

$\mathbf{D}$ represents the covariance matrix due to the parameter state's variability. In tracking terminology, this increase is referred to as the process noise in sequential estimation. Each entry of $\mathbf{D}$ signifies the parameter's variability

in that dimension. For moving target tracking, a larger dynamic variance specifies the object has a better ability to maneuver in that dimension, while a larger amount of covariance specifies the change of one target state variable will affect the other state variable. The amount of uncertainty added by $\mathbf{D}$ can be determined with knowledge of the parameter [41].

After predicting the current parameter state estimate from the previous state estimate and measuring the current actual parameter state, the tracker must combine both values into one estimate. This integration is calculated by

$$\hat{\mathbf{q}}_k = \tilde{\mathbf{q}}_k + \mathbf{K}_k(\mathbf{p}_k - \mathbf{M}\tilde{\mathbf{q}}_k), \tag{2.4}$$

where $\mathbf{K}_k$ is the Kalman filter weight matrix defined as

$$\mathbf{K}_k = \tilde{\mathbf{S}}_k\mathbf{M}^T\left[\mathbf{M}\tilde{\mathbf{S}}_k\mathbf{M}^T + \mathbf{R}\right]^{-1}, \tag{2.5}$$

where $\mathbf{R}$ is the measurement error covariance matrix, which is inversely dependent on SNR. By allocating the aperture, the radar affects the signal's SNR for each target, and subsequently the measurement error covariance matrix for each measurement. By modifying the measurement error covariance matrix, the Kalman weights in (2.5) change and ultimately affect tracking performance. By intelligently allocating the aperture, the allocation strategies can positively affect tracking performance to achieve better results for the metrics in Chapter 3.

Although the parameter covariance matrix is not measured by the sensor, the covariance update equation in (2.6) includes the Kalman filter weights in (2.5), which depend on SNR, and the previously predicted covariance estimate from (2.3), which depends on the target's maneuverability. The relationship

between $\mathbf{D}$ in (2.3) and $\mathbf{R}$ in (2.5) determines whether the resulting estimated variance for each parameter is increased due to the parameter's maneuverability or decreased by the measurement strength. By combining (2.3) and (2.5), the parameter covariance matrix becomes

$$\mathbf{S}_k = [\mathbf{I} - \mathbf{K}_k\mathbf{M}]\,\tilde{\mathbf{S}}_k, \tag{2.6}$$

where $\mathbf{I}$ is the identity matrix. The updated estimates for both the state and covariance matrices are recursively applied as the previous estimate at the next time step.

2.2 Example Implementation

To implement a Kalman filter, both the tracker's coordinate system and the parameter's dynamic model are necessary. While many coordinate systems [42] and dynamic models [43] exist, we focus on the three dimensional Cartesian coordinate system and the constant velocity model. With these choices, the parameter state becomes

$$\mathbf{q}_k = \begin{bmatrix} x_k \\ y_k \\ z_k \\ \dot{x}_k \\ \dot{y}_k \\ \dot{z}_k \end{bmatrix}, \tag{2.7}$$

where x_k, y_k, and z_k are the actual state location in Cartesian coordinates at time step k, and $\dot{x_k}$, $\dot{y_k}$, and $\dot{z_k}$ are the actual velocity components in the Cartesian plane at time step k [44]. The corresponding covariance matrix

becomes

$$\mathbf{S} = \begin{bmatrix} \sigma_x^2 & \sigma_{x,y}^2 & \sigma_{x,z}^2 & \sigma_{x,\dot{x}}^2 & \sigma_{x,\dot{y}}^2 & \sigma_{x,\dot{z}}^2 \\ \sigma_{x,y}^2 & \sigma_y^2 & \sigma_{y,z}^2 & \sigma_{y,\dot{x}}^2 & \sigma_{y,\dot{y}}^2 & \sigma_{y,\dot{z}}^2 \\ \sigma_{x,z}^2 & \sigma_{y,z}^2 & \sigma_z^2 & \sigma_{z,\dot{x}}^2 & \sigma_{z,\dot{y}}^2 & \sigma_{z,\dot{z}}^2 \\ \sigma_{x,\dot{x}}^2 & \sigma_{y,\dot{x}}^2 & \sigma_{z,\dot{x}}^2 & \sigma_{\dot{x}}^2 & \sigma_{\dot{x},\dot{y}}^2 & \sigma_{\dot{x},\dot{z}}^2 \\ \sigma_{x,\dot{y}}^2 & \sigma_{y,\dot{y}}^2 & \sigma_{z,\dot{y}}^2 & \sigma_{\dot{x},\dot{y}}^2 & \sigma_{\dot{y}}^2 & \sigma_{\dot{y},\dot{z}}^2 \\ \sigma_{x,\dot{z}}^2 & \sigma_{y,\dot{z}}^2 & \sigma_{z,\dot{z}}^2 & \sigma_{\dot{x},\dot{z}}^2 & \sigma_{\dot{y},\dot{z}}^2 & \sigma_{\dot{z}}^2 \end{bmatrix}, \tag{2.8}$$

where $\sigma_{a,b}^2$ is the covariance defined as

$$\sigma_{a,b}^2 = E[(a - E[a])(b - E[b])], \tag{2.9}$$

where E[] is the expected value of the random variable [44].

Once the state vector is determined, the sensor's output is also necessary to define the measurement equation. In this example, the sensor will be a radar system sensing the target's range and angle of arrival in both azimuth and elevation. The radar system implements a coordinate system transformation to produce measured values in the tracker's Cartesian coordinate system. The resulting measurement equation is

$$\mathbf{p}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ z_k \\ \dot{x}_k \\ \dot{y}_k \\ \dot{z}_k \end{bmatrix} + \begin{bmatrix} \nu_x^2 \\ \nu_y^2 \\ \nu_z^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{2.10}$$

where ν_x^2, ν_y^2, and ν_z^2 are the measurement noise variances in the x, y, and z dimension.

The observation matrix, $\mathbf{M}$, is defined in (2.10) for a Kalman filter implementation, and does not necessarily match the observation matrix for the measurement sensor. For example, (2.10) can be implemented for a radar tracking system, even though the radar sensor measures a target's range, elevation, azimuth, and radial velocity. In this case, the range, elevation, and azimuth must first be transformed into the tracker's Cartesian coordinate system before being input into the Kalman tracker as a measurement. Additionally, the radial velocity can be ignored, and the implemented tracker is capable of tracking the targets with only measured location data.

The corresponding parameter state estimate is

$$\hat{\mathbf{q}}_k = \begin{bmatrix} \hat{x}_k \\ \hat{y}_k \\ \hat{z}_k \\ \hat{\dot{x}}_k \\ \hat{\dot{y}}_k \\ \hat{\dot{z}}_k \end{bmatrix}, \tag{2.11}$$

where $\hat{x}_k$, $\hat{y}_k$, and $\hat{z}_k$ are the state location estimates in Cartesian coordinates, and $\hat{\dot{x}}_k$, $\hat{\dot{y}}_k$, and $\hat{\dot{z}}_k$ are the velocity vector estimates in the Cartesian plane. Substituting (2.11), (2.2) becomes

$$\tilde{\mathbf{q}}_k = \begin{bmatrix} 1 & 0 & 0 & T & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 \\ 0 & 0 & 1 & 0 & 0 & T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{k-1} \\ \hat{y}_{k-1} \\ \hat{z}_{k-1} \\ \hat{\dot{x}}_{k-1} \\ \hat{\dot{y}}_{k-1} \\ \hat{\dot{z}}_{k-1} \end{bmatrix}, \tag{2.12}$$

where T is the amount of time between each CPI. Since the target model in (2.7) includes only location and velocity, the dynamic noise, $\mathbf{D}$, is determined based on the target's acceleration. Haug [41] provides a detailed explanation and calculation of the dynamic noise based on acceleration for a constant velocity tracker. Based on Haug's explanation, the dynamic noise is

$$\mathbf{D} = \begin{bmatrix} \delta_x^2 & 0 & 0 & \delta_{x,\dot{x}}^2 & 0 & 0 \\ 0 & \delta_y^2 & 0 & 0 & \delta_{y,\dot{y}}^2 & 0 \\ 0 & 0 & \delta_z^2 & 0 & 0 & \delta_{z,\dot{z}}^2 \\ \delta_{x,\dot{x}}^2 & 0 & 0 & \delta_{\dot{x}}^2 & 0 & 0 \\ 0 & \delta_{y,\dot{y}}^2 & 0 & 0 & \delta_{\dot{y}}^2 & 0 \\ 0 & 0 & \delta_{z,\dot{z}}^2 & 0 & 0 & \delta_{\dot{z}}^2 \end{bmatrix}, \tag{2.13}$$

where δ_x^2, δ_y^2, and δ_z^2 are the variances in the x, y, and z dimension, $\delta_{\dot{x}}^2$, $\delta_{\dot{y}}^2$, and $\delta_{\dot{z}}^2$ are the variances in the x, y, and z velocity vector, and $\delta_{x,\dot{x}}^2$, $\delta_{y,\dot{y}}^2$, and $\delta_{z,\dot{z}}^2$ are the covariances between the target's location and velocity states [41]. These dynamic variances and covariances are determined by the parameter's acceleration and the period between each measurement [41]. The number of properties and dimensions tracked by the parameter state determines the size of $\mathbf{D}$. In this implementation, a three dimensional parameter state including location and velocity causes $\mathbf{D}$ to be a six by six square matrix to include dynamic noise for each parameter state variable and their covariances.

Substituting (2.13), (2.3) becomes

$$\tilde{\mathbf{S}}_k = \begin{bmatrix} 1 & 0 & 0 & T & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 \\ 0 & 0 & 1 & 0 & 0 & T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{S}_{k-1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ T & 0 & 0 & 1 & 0 & 0 \\ 0 & T & 0 & 0 & 1 & 0 \\ 0 & 0 & T & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \delta_x^2 & 0 & 0 & \delta_{x,\dot{x}}^2 & 0 & 0 \\ 0 & \delta_y^2 & 0 & 0 & \delta_{y,\dot{y}}^2 & 0 \\ 0 & 0 & \delta_z^2 & 0 & 0 & \delta_{z,\dot{z}}^2 \\ \delta_{x,\dot{x}}^2 & 0 & 0 & \delta_{\dot{x}}^2 & 0 & 0 \\ 0 & \delta_{y,\dot{y}}^2 & 0 & 0 & \delta_{\dot{y}}^2 & 0 \\ 0 & 0 & \delta_{z,\dot{z}}^2 & 0 & 0 & \delta_{\dot{z}}^2 \end{bmatrix}. \quad (2.14)$$

Substituting these equations into (2.4), (2.5), and (2.6) implements the full Kalman filter.

This chapter introduced concepts and equations to recursively estimate a parameter's value using the Kalman filter. These techniques are a useful introduction to future concepts in this dissertation. The relationship between $\mathbf{D}$ and SNR will be investigated in this dissertation to achieve the best system results for specific metrics. The metrics used for comparison along with a parallel channel model to describe the interaction between multiple parameters will be described in Chapter 3. Chapter 7 will implement the Kalman tracker and provide results from the multiple approaches on the tracker's performance.

Chapter 3

Information Theory

A common information theory construct models the act of transmitting values through a channel and receiving the values corrupted by noise with independent parallel channels [28]. This model is shown in Figure 3.1, where $X_{n,k}$ is the input value for the n^{th} channel at the k^{th} time step, $\rho_{n,k}$ is the amount of transmit power allocated to the n^{th} channel at the k^{th} time step, $Z_{n,k}$ is the additive noise (commonly Gaussian [28]) for the n^{th} channel at the k^{th} time step, and $Y_{n,k}$ is the received data, corrupted by noise, for the n^{th} channel at the k^{th} time step.

Contrary to communication systems, which try to recover the transmitted value on the channel output, this effort applies the parallel Gaussian channel model to the parameter estimation problem with an active sensor. The parameters are modeled as the channel input, while the illumination strength of the active sensor toward each parameter is modeled as the power amplifier of the channel. For a radar system, the all-digital architecture enables the sensor to reconfigure and allocate resources dynamically across the multiple channels. For this dissertation, the channels are independent, including the independent Gaussian noise sources, with the only interaction between them being a total power constraint for a specific illumination or time step. Chapter 5 will in-

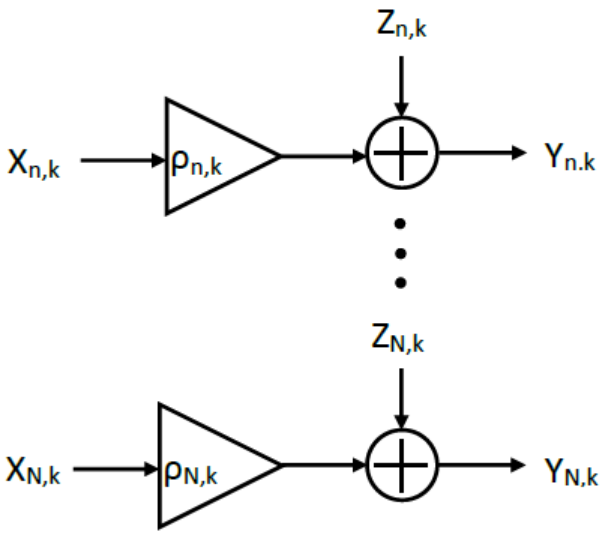

Figure 3.1: Parallel channel model for estimation of multiple parameters from [45]

vestigate the case when the input values are correlated, but still models the channel measurements as independent. Similar to (2.1), we mathematically define the model as

$$\mathbf{y}_k = \sqrt{\mathbf{P}_k}\mathbf{x}_k + \mathbf{z}_k \quad k = 1, 2, ..., K \tag{3.1}$$

where $\mathbf{x}_k$ is the vector of parameters to estimate at time step k, $\mathbf{z}_k$ is the vector of zero mean Gaussian noise with covariance matrix $\mathbf{N}$ at time k, $\mathbf{P}_k$ is the power matrix allocated at time step k, $\mathbf{y}_k$ is the measured data vector at time k, and K is the total number of time steps.

While (3.1) is strikingly similar to (2.1), there are distinct differences to warrant separate variables for each signal. Equation (2.1) defines the measurement of a single parameter in multiple dimensions, while (3.1) defines the measurement of multiple parameters. To integrate (2.1) and (3.1), the vectors of (2.1) must be combined to represent one element of the vectors in (3.1).

To model the overall power constraint of a typical radar, a total overall power, P, constrains the amount of allocated power by

$$\mathrm{tr}(\mathbf{P}_k) \leq P, \tag{3.2}$$

where tr() is the trace operator. Since the amount of power allocated to each channel is distinct and the measurement channels are independent, the power matrix is a diagonal matrix given by

$$\mathbf{P}_k = \begin{pmatrix} \rho_{1,k} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \rho_{2,k} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \rho_{n,k} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \rho_{N,k} \end{pmatrix}. \tag{3.3}$$

Substituting (3.3), (3.2) can simplify to

$$\sum_{n=1}^{N} \rho_{n,k} \leq P. \tag{3.4}$$

Entropy, mutual information, and mean-square error (MSE) are important quantitative mathematical metrics for sufficiently comparing a system's ability to estimate parameter values. The measure of a single random variable's uncertainty is referred to as entropy, while the measure of uncertainty for a set of random variables is joint entropy [28]. For a group of random parameters $\mathbf{X}_k = [X_{1,k}, \ldots, X_{N,k}]$, with instances defined as $\mathbf{x}_k = [x_{1,k}, \ldots, x_{N,k}]$ and joint

probability density function (pdf) defined by $p_k(\mathbf{x}_k)$, the joint entropy is

$$H(\mathbf{X}_k) = -\int_{\mathbf{x}_k \in \mathbf{X}_k} p_k(\mathbf{x}_k) \log p_k(\mathbf{x}_k) d\mathbf{x}_k \tag{3.5}$$

where the integration bounds are defined by every possible combination of the random parameters and log is the natural logarithm. For a multivariate, real-valued Gaussian random vector, the joint entropy becomes

$$H(\mathbf{X}_k) = \frac{1}{2} \log \left[\det \left(2\pi e \boldsymbol{\Sigma}_\mathbf{k} \right) \right] \tag{3.6}$$

where $\boldsymbol{\Sigma}_k$ is the multivariate covariance matrix and det is the determinant operator.

Conditional entropy is useful in ascertaining a parameter's entropy when the random parameter is not directly observed, but a separate yet related set of variables can be measured. In this case, the conditional entropy of a set of random variables, $\mathbf{X}_k = [X_{1,k}, \ldots, X_{N,k}]$, given another set of random variables, $\mathbf{Y}_k = [Y_{1,k}, \ldots, Y_{N,k}]$, is

$$H(\mathbf{X}_k | \mathbf{Y}_k) = -\int_{\mathbf{y}_k \in \mathbf{Y}_k} \int_{\mathbf{x}_k \in \mathbf{X}_k} p_k(\mathbf{y}_k, \mathbf{x}_k) \log p_k(\mathbf{x}_k | \mathbf{y}_k) d\mathbf{y}_k d\mathbf{x}_k. \tag{3.7}$$

Conditional entropy is related to mutual information because both metrics describe the relationship between the entropies of two random variables. While conditional entropy relates the value of both variables' entropies, mutual information determines the reduction in entropy of one variable when provided access to another separate but related variable. The mutual information, derived using both joint and conditional entropy, between two groups of random

variables is defined as

$$I(\mathbf{X}_k; \mathbf{Y}_k) = H(\mathbf{X}_k) - H(\mathbf{X}_k|\mathbf{Y}_k). \tag{3.8}$$

For Gaussian random variables and a linear measurement model with additve Gaussian noise, the resulting measurements and parameters are jointly Gaussian [46]. The mutual information between jointly Gaussian variables is [25], [47]

$$I(\mathbf{x}_k, \mathbf{y}_k) = \frac{1}{2} \log \left[\det \left(\mathbf{\Sigma}_k \mathbf{N}^{-1} \mathbf{P}_k + \mathbf{I} \right) \right], \tag{3.9}$$

where $\mathbf{I}$ is the identity matrix. The relationship between mutual information, entropy, and conditional entropy is depicted in Figure 3.2 [28].

To account for noise corruption, an estimator is useful to extract an estimate, $\hat{\mathbf{x}}_{n,k}$, of the actual signal, $\mathbf{x}_{n,k}$, for the n^{th} parameter at time k. To compare the accuracy of the estimate with the actual signal, the estimate's MSE is [48]

$$MSE\left(\hat{\mathbf{x}}_k\right) = \frac{1}{N} \sum_{n=1}^{N} \left(\hat{\mathbf{x}}_{n,k} - \mathbf{x}_{n,k}\right)^2. \tag{3.10}$$

For completeness, we note that Guo, Shamai, and Verdú [49] derived the intimate relationship between mutual information and minimum MSE given by

$$\frac{\partial}{\partial \mathbf{P}_k} I(\mathbf{x}_k, \mathbf{y}_k) = \frac{1}{2} MMSE\left(\hat{\mathbf{x}}\right) \tag{3.11}$$

where $MMSE$ is the minimum mean-squared error. Therefore, information metrics and MSE, the typical metric for tracking performance quantification, are very closely associated. These quantitative metrics will be used throughout the rest of this dissertation to quantify and derive multiple approaches to aperture allocation for multiple target tracking.

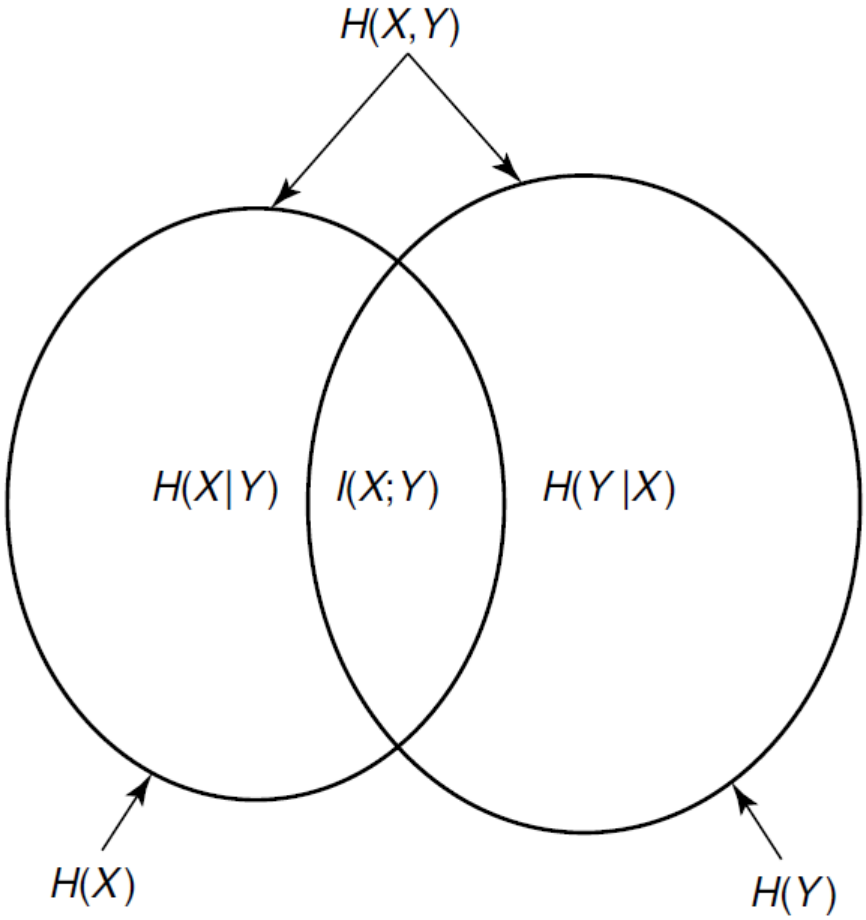

Figure 3.2: Relationship between entropy and mutual information from [28]

Chapter 4

Power Allocation

4.1 System Model

A simplified measurement model is crucial for initial investigations into aperture allocation without building a complete operating system. Applying the system model from Chapter 3, we assume the parameters in (3.1) are completely independent. For a target tracking application, the channels from the parallel channel model in Figure 3.1 represent each target for tracking. Therefore, channels and targets are used interchangeably throughout this chapter. The channel independence assumption simplifies the parameter covariance matrix, $\mathbf{\Sigma}_k$, to a diagonal matrix such that

$$\mathbf{\Sigma}_k = \begin{pmatrix} \sigma_{1,k}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{2,k}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{n,k}^2 \end{pmatrix} \tag{4.1}$$

where $\sigma_{n,k}^2$ is the n^{th} channel's variance at time k. For this simplification, we use a Bayesian estimation paradigm to determine the parameters' variance over time using the common model of Gaussian priors [28] with known variance.

To quantify the Bayesian approach, the belief state defined as the probability distribution of the parameter [8] is

$$b_{n,k} = p(x_{n,k}|y_{n,k}) \sim \mathcal{N}(\mu_{n,k}, \sigma^2_{n,k}) \tag{4.2}$$

where $x_{n,k}$ is the channel input, $y_{n,k}$ is the channel measurement, and $\mu_{n,k}$ is the estimated parameter value. Each measurement updates the estimated parameter value of the belief state with new information and decreases the uncertainty, or variance, of the belief state. When the estimated value of the parameter changes due to a measurement, the decrease in uncertainty, or variance, due to the measurement, is countered by an increase between measurements due to the parameter dynamics or variability. Before the subsequent measurement, the increase in variance, $\delta^2_{n,k}$ for the n^{th} channel at time k, causes the belief state to become

$$\hat{b}_{n,k+1} \sim \mathcal{N}(\mu_{n,k}, \sigma^2_{n,k} + \delta^2_{n,k}). \tag{4.3}$$

After the subsequent measurement, the posterior belief state for a prior Gaussian with a known variance becomes

$$b_{n,k+1} \sim \mathcal{N}\left(\frac{y_{n,k}\rho_{n,k+1}(\sigma^2_{n,k} + \delta^2_n) + \mu_{n,k}\nu^2_n}{\rho_{n,k+1}(\sigma^2_{n,k} + \delta^2_n) + \nu^2_n}, \frac{(\sigma^2_{n,k} + \delta^2_n)\nu^2_n}{\rho_{n,k+1}(\sigma^2_{n,k} + \delta^2_n) + \nu^2_n}\right) \tag{4.4}$$

where $\mu_{n,k}$ and $\sigma^2_{n,k}$ are the previous belief state's mean and variance, and $y_{n,k}$ is the current measurement [50].

Substituting the diagonal matrix in (4.1) and extracting a single channel's entropy, the single channel version of (3.6) becomes

$$h(X_{n,k}) = \frac{1}{2}\log\left(2\pi e\sigma^2_{n,k}\right) \tag{4.5}$$

Applying the same simplification to the definition of mutual information and accounting for the increase in uncertainty between measurements due to parameter variability produces the single channel version of (3.9) as

$$I(x_{n,k}, y_{n,k}) = \frac{1}{2} \log \left(1 + \frac{\rho_{n,k+1}(\sigma_{n,k} + \delta_{n,k}^2)}{\nu_{n,k}^2} \right). \tag{4.6}$$

4.2 Power Allocation Strategies

Using the model in Section 4.1, we develop a short-term objective function and three long-term objective functions to obtain four resource allocation strategies in this section. The approaches in this section derive solutions for $\sqrt{\mathbf{P}_k}$ in (3.1) to represent the power allocated to each channel. For a sensor with constant gain across channels, the power allocated for each channel determines the total resources allocated.

4.2.1 Mutual Information Maximization

By measuring a set of variables, the knowledge or information of each variable is increased based on the strength of each measurement. The mutual information metric is a quantitative measure of the information amount gained from each measurement. By maximizing this metric, the knowledge of each variable is subsequently increased and the amount of uncertainty is decreased. The first power allocation approach, described in this section, maximizes the summation of mutual information over all channels as the optimization function.

Mutual information is the knowledge gained for a specific measurement. Although increasing mutual information for the immediate measurement maximizes the benefit of that measurement, it doesn't correlate to optimal long-term performance [24]. To optimize the long-term performance, mutual infor-

mation must be maximized across all measurements and for every time step, which is computationally infeasible for a real time system. Instead, the mutual information maximization approach of this section increases the information gained for the immediate measurement, and must be recalculated before each measurement to determine the power allocation. By recalculating before each measurement, the mutual information maximization approach optimizes the power allocation for the immediate measurement on a short-term time line. Summing (4.6) over all channels results in the objective function to optimize the power allocation for the next measurement at the k^{th} time step being

$$R_{tot}^{(k)} = \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \rho_{n,k+1} \frac{\sigma_{n,k}^2 + \delta_n^2}{\nu_n^2} \right). \tag{4.7}$$

For many active sensors, such as a radar, the amount of power being radiated by the aperture at a specific time is constrained by a given total constraint. The power constraint is captured in (3.4) and must be applied to each measurement. To apply the constraint to the maximum mutual information approach, the Lagrangian multiplier technique is implemented to include the constraint in (3.4). By applying the Lagrangian multiplier technique, (4.7) becomes

$$R_{tot}^{(k)} = \sum_{n=1}^{N} \frac{1}{2} \log \left(1 + \rho_{n,k+1} \frac{\sigma_{n,k}^2 + \delta_n^2}{\nu_n^2} \right) - \lambda \left(\sum_{n=1}^{N} \rho_{n,k+1} - P \right) \tag{4.8}$$

where λ is the Lagrangian multiplier. The derivative of (4.8) with respect to the power allocated to the n^{th} channel, ρ_n, is

$$\frac{\partial R_{tot}^{(k)}}{\partial \rho_{n,k+1}} = \frac{\sigma_{n,k}^2 + \delta_n^2}{2\rho_{n,k+1} \left(\sigma_{n,k}^2 + \delta_n^2 \right) + 2\nu_n^2} - \lambda. \tag{4.9}$$

Setting (4.9) equal to zero to minimize and solving for $\rho_{n,k+1}$, (4.9) becomes

$$\rho_{n,k+1} = \frac{1}{2\lambda} - \frac{\nu_n^2}{\sigma_{n,k}^2 + \delta_n^2}. \tag{4.10}$$

The Langrangian multiplier can be removed from (4.10) by calculating the derivative of (4.8) with respect to the Langrangian multiplier. The derivative of (4.8) with respect to λ is

$$\frac{\partial R_{tot}^{(k)}}{\partial \lambda} = \sum_{n=1}^{N} \rho_{n,k+1} - P. \tag{4.11}$$

Minimizing (4.11) by setting it equal to zero, substituting (4.10), and solving for λ becomes

$$\lambda = \frac{N}{2\left(P + \sum_{n=1}^{N} \frac{\nu_n^2}{\sigma_{n,k}^2 + \delta_n^2}\right)}. \tag{4.12}$$

Since (4.12) has a summation across all channels, a change in indexes is required when substituting into any equations using n to symbolize the current channel. After changing the index in (4.12) to p, substituting (4.12) back into (4.10) will remove the Lagrangian multiplier from the set of equations. After the substitution, the maximum solution of (4.7) becomes

$$\rho_{n,k+1} = \frac{1}{N}\left(P + \sum_{p=1}^{N} \frac{\nu_p^2}{\sigma_{p,k}^2 + \delta_p^2}\right) - \frac{\nu_n^2}{\sigma_{n,k}^2 + \delta_n^2}. \tag{4.13}$$

Calculating the power allocation using (4.13) maximizes total mutual information across the parallel measurements at a time instance $k+1$.

For a constant measurement and dynamic variance, maximizing the mutual information across parallel channels at each time step will eventually cause the parameter variances to settle at a steady state. When this steady state

is reached, the posterior and prior variances should become equal, meaning the mutual information gained at each time step is the same as the dynamic variance injected between steps. We can exploit this behavior to derive the steady state variance by equating the prior and posterior variances. After substituting $\sigma^2_{n,\infty}$ for the prior and posterior variance in (4.4), the asymptotic variance is

$$\sigma^2_{n,\infty} = \frac{(\sigma^2_{n,\infty} + \delta^2_n)\nu^2_n}{\rho_{n,\infty}(\sigma^2_{n,\infty} + \delta^2_n) + \nu^2_n}. \tag{4.14}$$

By rearranging (4.14), the asymptotic power of the mutual information maximization approach is then

$$\rho_{n,\infty} = \frac{\delta^2_n \nu^2_n}{\sigma^2_{n,\infty}(\sigma^2_{n,\infty} + \delta^2_n)}. \tag{4.15}$$

The long-term effects of the mutual information maximization approach can be derived using the asymptotic power allocation in (4.15). A system of equations for the long-term results can be derived by applying this power allocation to (4.9). Therefore, substituting the asymptotic power allocation in (4.15) for $\rho_{n,k+1}$, (4.9) becomes

$$\frac{\partial I_{n,k}}{\partial \rho_{n,\infty}} = \frac{\sigma^2_{n,\infty}}{2\nu^2_n} - \lambda. \tag{4.16}$$

Maximizing (4.16) requires setting it equal to zero. Solving for λ then becomes

$$\lambda = \frac{\sigma^2_{n,\infty}}{2\nu^2_n}. \tag{4.17}$$

Equation (4.17) defines the Lagrangian multiplier, λ, for the n^{th} channel, but the multiplier is common across all the channels. Deriving (4.17) for every

channel results in

$$\lambda = \frac{\sigma_{1,\infty}^2}{2\nu_1^2} = \frac{\sigma_{2,\infty}^2}{2\nu_2^2} = ... = \frac{\sigma_{n,\infty}^2}{2\nu_n^2}. \tag{4.18}$$

Generalizing (4.18) for every channel and simplifying, the long term variance at steady state for each channel can be solved numerically by

$$\frac{\sigma_{n,\infty}^2}{\nu_n^2} = \frac{\sigma_{m,\infty}^2}{\nu_m^2}, \tag{4.19}$$

where n and m refer to different channels. The power constraint requires that

$$\sum_{n=1}^{N} \frac{\nu_n^2 \delta_n^2}{\sigma_{n,\infty}^2(\sigma_{n,\infty}^2 + \delta_n^2)} = P. \tag{4.20}$$

By solving the system of equations in (4.19) and (4.20), the asymptotic parameter variance for the mutual information maximization approach can be calculated for a given measurement and dynamic variance. In contrast to Sections 4.2.2, 4.2.3, and 4.2.4, the mutual information maximization approach calculates the power allocation before each measurement occurs.

4.2.2 Entropy Asymptote Minimization

The excessive computational complexity required to maximize the mutual information in Section 4.2.1 for an optimal long-term solution restricted the approach to a short-term optimization. Because mutual information depends on the knowledge gained from the previous measurement, increasing the numerical search to multiple measurements increases the computational complexity exponentially for each additional measurement.

In contrast, the approach in this section, the asymptotic entropy minimization approach, exploits the steady state entropy that is reached when the

power allocation, measurement variance, and dynamic variance are held constant. The approach in this section calculates and minimizes these asymptotic parameter variances by, once again, setting the prior variance and posterior variance equal for each channel. This approach achieves a long term solution, whose time line before reaching a steady state is dependent on the dynamic, measurement, and initial parameter variances. Using the quadratic formula and rearranging to find the asymptotic variance, (4.14) becomes

$$\sigma_{n,\infty}^2 = -\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \frac{\delta_n^2 \nu_n^2}{\rho_{n,\infty}}}. \tag{4.21}$$

Since the steady-state variance requires a constant power allocation across time, the time subscript can be removed for power, ρ_n. Substituting into (4.5), the asymptotic entropy becomes

$$H(X_{n,\infty}) = \frac{1}{2}\log\left[2\pi e \cdot \left(-\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \frac{\delta_n^2 \nu_n^2}{\rho_n}}\right)\right]. \tag{4.22}$$

Since (4.22) is the entropy for a single channel, the total objective function must optimize across all N channels with a summation. The objective function of this summation is

$$R_{tot} = \sum_{n=1}^{N} \frac{1}{2}\log\left[2\pi e \cdot \left(-\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \frac{\delta_n^2 \nu_n^2}{\rho_n}}\right)\right]. \tag{4.23}$$

Applying the Lagrange multiplier technique with the constraint in (3.4), (4.23) becomes

$$R_{tot} = \sum_{n=1}^{N} \frac{1}{2}\log\left[2\pi e \cdot \left(-\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \frac{\delta_n^2 \nu_n^2}{\rho_n}}\right)\right] - \lambda\left(\sum_{n=1}^{N} \rho_n - P\right). \tag{4.24}$$

The derivative of (4.24), with respect to ρ_n, is

$$\frac{\partial R_{tot}}{\partial \rho_n} = \frac{\nu_n^2}{\delta_n^2\rho_n^2 + 4\nu_n^2\rho_n - \rho_n\sqrt{(\delta_n^2\rho_n)^2 + 4\delta_n^2\nu_n^2\rho_n}} - \lambda. \tag{4.25}$$

Setting (4.25) equal to zero to minimize and solving for λ yields

$$\lambda = \frac{\nu_n^2}{\delta_n^2\rho_n^2 + 4\nu_n^2\rho_n - \rho_n\sqrt{(\delta_n^2\rho_n)^2 + 4\delta_n^2\nu_n^2\rho_n}}. \tag{4.26}$$

Performing the same derivative for $\rho_1, \rho_2, ..., \rho_N$ results in

$$\begin{aligned}\lambda = \frac{\nu_1^2}{\delta_1^2\rho_1^2 + 4\nu_1^2\rho_1 - \rho_1\sqrt{(\delta_1^2\rho_1)^2 + 4\delta_1^2\nu_1^2\rho_1}} = \\ \frac{\nu_2^2}{\delta_2^2\rho_2^2 + 4\nu_2^2\rho_2 - \rho_2\sqrt{(\delta_2^2\rho_2)^2 + 4\delta_2^2\nu_2^2\rho_2}} = \\ ... = \frac{\nu_N^2}{\delta_N^2\rho_N^2 + 4\nu_N^2\rho_N - \rho_N\sqrt{(\delta_N^2\rho_N)^2 + 4\delta_N^2\nu_N^2\rho_N}}.\end{aligned} \tag{4.27}$$

Abstracting this result for the m^{th} and n^{th} channels results in a set of $N-1$ equations given by

$$\begin{aligned}\frac{\delta_n^2\rho_n^2}{\nu_n^2}\left(1 + \frac{4\nu_n^2}{\delta_n^2\rho_n} - \sqrt{1 + \frac{4\nu_n^2}{\delta_n^2\rho_n}}\right) = \\ \frac{\delta_m^2\rho_m^2}{\nu_m^2}\left(1 + \frac{4\nu_m^2}{\delta_m^2\rho_m} - \sqrt{1 + \frac{4\nu_m^2}{\delta_m^2\rho_m}}\right) \forall n \neq m,\end{aligned} \tag{4.28}$$

which, when also including (3.4), becomes a system of N equations for N unknowns. Solving this system of equations produces the optimal power allocation to minimize the long-term summation of entropy across channels.

When the measurement variance to dynamic variance is the same ratio for each channel, the power allocation for this approach simplifies to an even

distribution. This condition can be mathematically described as

$$\nu_n^2 = A\delta_n^2 \quad n = 1, 2, ..., N \tag{4.29}$$

where A is a constant. Compared to the asymptotic performance of Section 4.2.1 in (4.19) and (4.20), when (4.29) is true the mutual information maximization approach has the same performance as the asymptotic entropy minimization approach. Section 4.3 shows results when (4.29) is true. Equation (4.29) is also the only case where evenly distributing the energy is the optimal long-term solution for minimizing the entropy. Some circumstances can result in similar results between the asymptotic entropy minimization approach and evenly distributing the power (such as when resources are very scarce), even though (4.29) is the only condition when they are both mathematically equal.

4.2.3 Highest Entropy Asymptote Minimization (Minimax)

The total entropy is important to determine the uncertainty of a whole system, but some objectives require each channel uncertainty to be reduced. Each channel's uncertainty level is important for systems that require a specified amount of uncertainty to benefit from the channel. For example, a target tracking type system might require a certain amount of uncertainty to be able to continue tracking the target. Once the target's uncertainty increases beyond a system threshold, the system loses track of that target.

Threshold or peak entropy metrics are important for modeling the ability of a radar tracking system to maintain track on a specific target. When the uncertainty of a channel increases beyond the threshold, a radar system may not

have enough information to reliably update the channel via new measurements and would lose track of the moving target.

In contrast to the preceding strategies, which minimized the total entropy on either a short-term or long-term basis, the approach in this section calculates the power allocation to achieve the smallest possible entropy value for every channel. To reduce the most uncertain channels, more resources, or power, must be allocated to decrease the uncertainty as much as possible. Allocating many resources to the more uncertain channels does not leave many resources left for the less uncertain channels. Therefore, the less uncertain channels increase until all channels reach the same entropy value. At this point, the approach in this section has reduced all the entropies as much as possible. Therefore, the system of equations for this approach is derived by setting (4.22) equal for all targets. Mathematically, this system of $N-1$ equations is defined by

$$\log\left[2\pi e\cdot\left(-\frac{\delta_n^2}{2}+\sqrt{\left(\frac{\delta_n^2}{2}\right)^2+\frac{\delta_n^2\nu_n^2}{\rho_n}}\right)\right]= \log\left[2\pi e\cdot\left(-\frac{\delta_m^2}{2}+\sqrt{\left(\frac{\delta_m^2}{2}\right)^2+\frac{\delta_m^2\nu_m^2}{\rho_m}}\right)\right]\ \forall n\neq m, \quad (4.30)$$

with (3.4) being the N^{th} equation.

To calculate the asymptotic performance of this approach, we substitute the result of solving (4.30), with (3.4), into (4.22). If the asymptotic performance does not reduce each channel's uncertainty enough to meet the system requirement for channel entropy, which is determined experimentally, we reallocate the power without including the channel requiring the highest power (i.e., that channel will receive no power).

4.2.4 Variance Asymptote Minimization Below a Threshold

The approaches in sections 4.2.1, 4.2.2, and 4.2.3 are restricted between either optimizing for an average or a min-max type metric. Optimizing for the min-max metric does not consider the overall system performance, while the average metrics do not allow an upper entropy threshold to be applied. The min-max metric compromises the overall system performance by focusing only on each channel's entropy. The average type approaches ignores any channel specific requirements that could cause all channels to be inadequate. In contrast, by minimizing the average entropy while maintaining channel entropies below a threshold, both metrics can be combined with the ability to trade-off by setting the threshold value. By applying this concept, we are able to provide a compromise between the previous strategies.

The objective of this section is to reduce the asymptotic variance, given in (4.21) below a given threshold, ϵ. Mathematically, this becomes

$$\sigma^2_{n,\infty} = -\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \frac{\delta_n^2 \nu_n^2}{\rho_n}} < \epsilon. \tag{4.31}$$

Solving for the power, ρ_n, yields the power allocation required to achieve

$$\rho_n > \frac{\delta_n^2 \nu_n^2}{\left(\epsilon + \frac{\delta_n^2}{2}\right)^2 - \left(\frac{\delta_n^2}{2}\right)^2}. \tag{4.32}$$

If the power allocation calculated in (4.32) exceeds the constraint in (3.4), the channel with the highest amount of allocated power should be removed and receive zero power. Then, (4.32) is recalculated for the remaining channels, and iterate the process of removing the highest channel and recalculating until

the total power requirement is met. After the variances are reduced below the given threshold with (4.32), the residual power can be distributed with the asymptotic entropy minimization approach to reduce the overall system uncertainty.

If the given entropy threshold is high, then only a minimal amount of power (if any) will be allocated using (4.32). In this case, we allocate the resulting power allocation with the asymptotic entropy minimization approach. By setting the correct threshold, the system designer is able to trade off between minimizing the highest entropy and minimizing the overall system uncertainty.

4.3 Results - Independent Parallel Channels

To effectively compare the power allocation strategies detailed in Section 4.2, this section details multiple scenarios to highlight the different approaches. Along with the four approaches in Section 4.2, we simulate the scenarios using two baseline approaches. These baseline approaches include evenly distributing the power across channels and applying all power to the channel with the highest variance (named the greedy approach).

For each scenario, the initial parameter entropies, $H(X_{n,0})$, are determined by substituting an even power allocation into (4.22). To quantify the amount of resources available for the simulations in this section, each approach can allocate 10 units (Watts) of power per unit of time. Three of the four simulations demonstrate the performance of each approach in Section 4.2 over a certain variable. Scenario one varies the SNR values, scenario two varies the dynamics between each channel, and scenario three varies the number of channels. The fourth scenario demonstrates that the condition in (4.29) results in the mutual information maximization, asymptotic entropy minimization, and

the even distribution approaches having the same asymptotic performance.

4.3.1 Differing SNR Environment

The first scenario assigns different measurement variances for two channels while assigning them the same dynamic variances. To represent channels with slow variability, the dynamic variances were chosen to be low, with a value of 0.0001, to represent channels with low variability, while the measurement covariance matrix,

$$\Sigma_m = \begin{bmatrix} 10^4 & 0 \\ 0 & 10^8 \end{bmatrix}, \tag{4.33}$$

was selected to represent one medium and one low-SNR target. For a target tracking application, these variances are analogous to two barely maneuvering objects at different ranges. Substituting 5 units of power for each channel into (4.22) results in channel one's initial entropy being 1.02, while channel two's initial entropy is 3.32. To quantify the results, (4.19) and (4.20) are used for the mutual information maximization approach, and (4.22) for the remaining approaches, to calculate the asymptotic entropy for each channel. The resulting entropies for each channel were then averaged across the two channels for each method to calculate the values in Table 4.1. Compared to Table 4.1, the initial variances are low for one target and high for the other target. The resulting average entropies are in between both starting values.

Figure 4.1 displays the simulation results for 10,000 time steps to evaluate performance over time. Analysis of this plot shows that the results for multiple allocation strategies are close enough to overlap on this graph. Both the minimize highest entropy and greedy approaches overlap, and the even distribution approach overlaps with the asymptote optimized approach in Figure 4.1. Even

Table 4.1: Asymptotic entropy for differing SNR variances.

Approach	Asymptotic Entropy
"Even Distribution"	2.1678
"Asymptotic Optimized"	2.1678
"Mutual Information"	3.1458
"Minimize Highest"	3.1459
"Threshold"	2.2985

though the even distribution and optimized asymptotic approaches appear to have the same results, the asymptotic approach actually has marginally better performance that can be observed by zooming into the figure. The mutual information (MI) maximization performs best (has the lowest entropy) for the time period shown. However, the time line of this figure is deceptive for the long-term performance, which is apparent when comparing the performance of each approach to the asymptotic values in Table 4.1. Although the mutual information maximization approach appears better in the figure, eventually it will reach a worse steady state than the asymptotic minimized approach.

Both the greedy and minimize highest entropy approaches ignore the channel with a measurement variance of 10^4, and allocates all the energy to the channel with a measurement variance of 10^8. By only focusing on the higher measurement variance target, both approaches sacrifice the average entropy metric in order to reduce the highest entropy target as much as possible. A compromise between reducing the high measurement variance channel and minimizing the overall entropy is shown by the threshold approach. By changing the threshold value, the threshold approach varies between the asymptotic minimization approach and the minimize highest entropy approach.

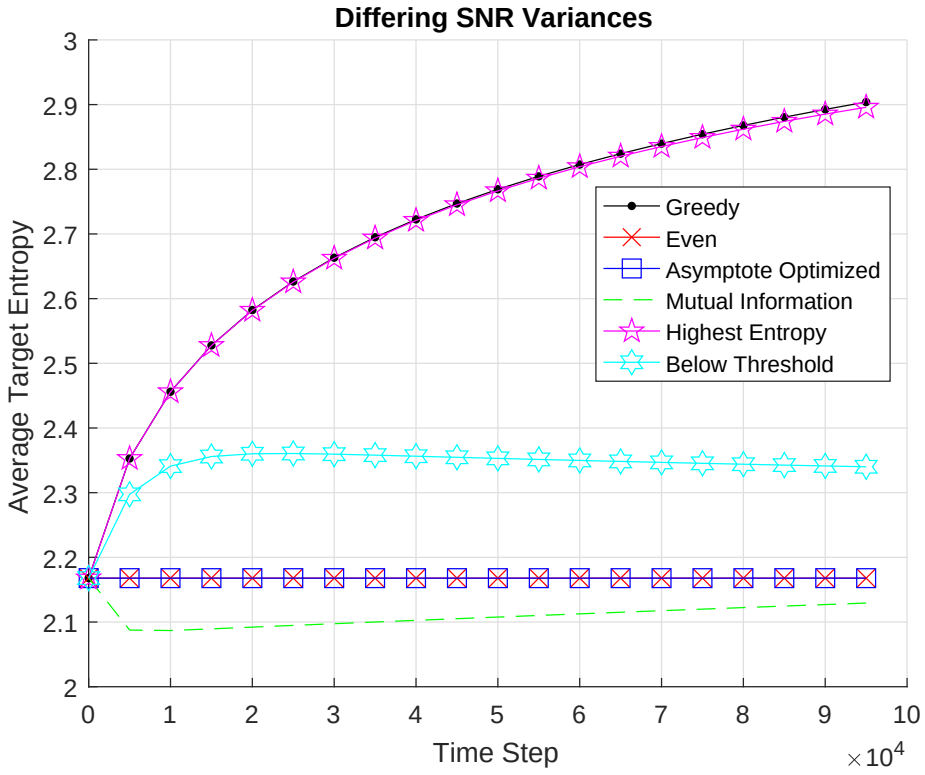

Figure 4.1: Entropy comparison for differing SNR variances.

Table 4.2: Asymptotic entropy for differing dynamic variances.

Approach	Asymptotic Entropy
"Even Distribution"	3.6539
"Asymptotic Optimized"	3.6406
"Mutual Information"	4.8290
"Minimize Highest"	4.8290
"Threshold"	3.7788

4.3.2 Differing Dynamic Environment

Complementary to the first scenario, the second scenario simulates two channels with different dynamic variances and the same measurement variance. The measurement variance for each target is set equal to 10^4, while the dynamic covariance matrix,

$$\Sigma_d = \begin{bmatrix} 10^{-2} & 0 \\ 0 & 10^4 \end{bmatrix}, \tag{4.34}$$

includes one channel with a large variability and one channel with minimal variability. For a target tracking application, the two channels would represent targets at the same distance from the sensor but with different amount of state vector variability. Substituting 5 units of power for each channel into (4.22) results in channel one's initial entropy being 2.17, while channel two's initial entropy is 6.10. Using (4.19) and (4.20) or (4.22) (depending on the approach), the asymptotic entropies are calculated and shown in Table 4.2.

Figure 4.2 displays the simulation results for 10,000 time steps to evaluate the performance over time. Analysis of this plot shows that the results of multiple allocation strategies are close enough to overlap on this graph. Once again, the asymptotic entropy minimization approach appears to overlap with the even distribution in this graph, even though Table 4.2 verifies that their

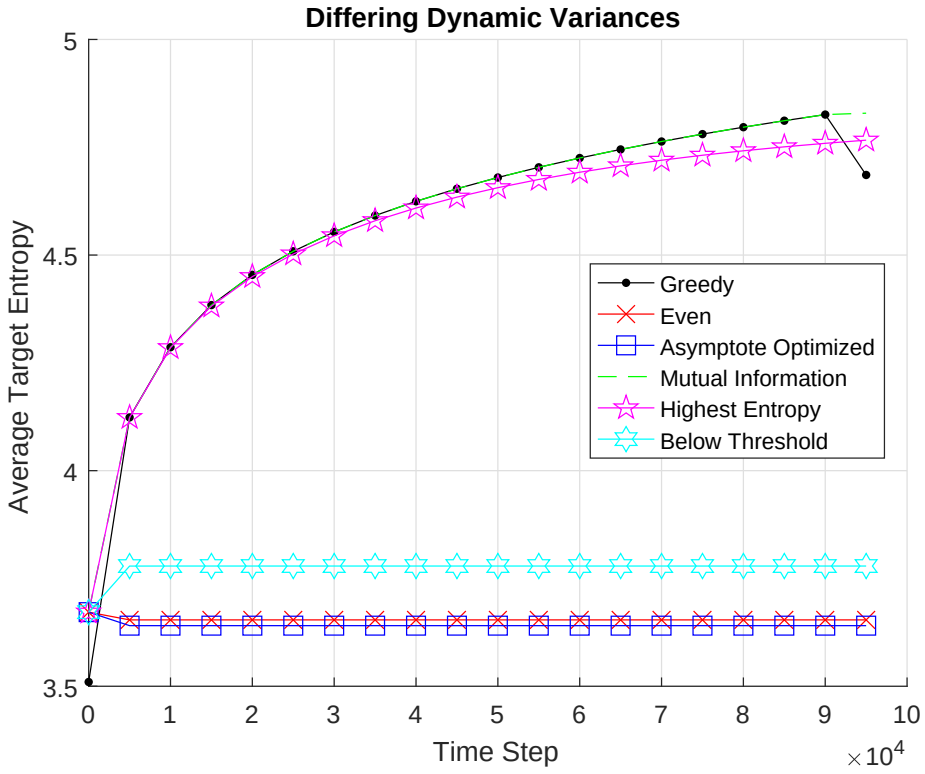

Figure 4.2: Entropy comparison for differing dynamic variances.

steady state values are actually different. The greedy, MI maximization, and minimize highest entropy approaches have similar results, until the end of the time line, due to allocating the majority of the power to the more dynamic channel.

4.3.3 Equal Variance Ratio Between Channels

By assigning the same dynamic to measurement variance ratio to each channel, we can show that this condition causes the mutual information, asymptote optimization, and even distribution approach to have the same performance. Using this criteria, the dynamic and measurement covariance matrices were

Table 4.3: Asymptotic entropy with equal variance ratio.

Approach	Asymptotic Entropy
"Even Distribution"	1.4029
"Asymptotic Optimized"	1.4029
"Mutual Information"	1.4029
"Minimize Highest"	1.4582
"Threshold"	1.4303

chosen to be

$$\Sigma_d = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \qquad \Sigma_m = \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix}. \tag{4.35}$$

Figure 4.3 and Table 4.3 demonstrate that all three approaches (mutual information maximization, asymptote entropy minimization, and even distribution) for this scenario, which abide by the requirement in (4.29), have the same allocation.

The remaining results are, predictably, in lowest-to-highest order the thresholding approach, the highest entropy minimization, and then the greedy approach.

4.3.4 Three Channels

The preceding scenarios used the overall average entropy to compare the approaches, while single channel metrics, referred to as a min-max metric, might be a more effective objective for some systems. The min-max metric aims to decrease the entropies below a specified value. By increasing the number of channels and analyzing the entropies of each target, in contrast to the average, the performance of each approach for each channel is more apparent. To this end, this simulation increases the number of channels to three with the same dynamic variance, $\delta^2 = 1$, for each channel, and a measurement covariance

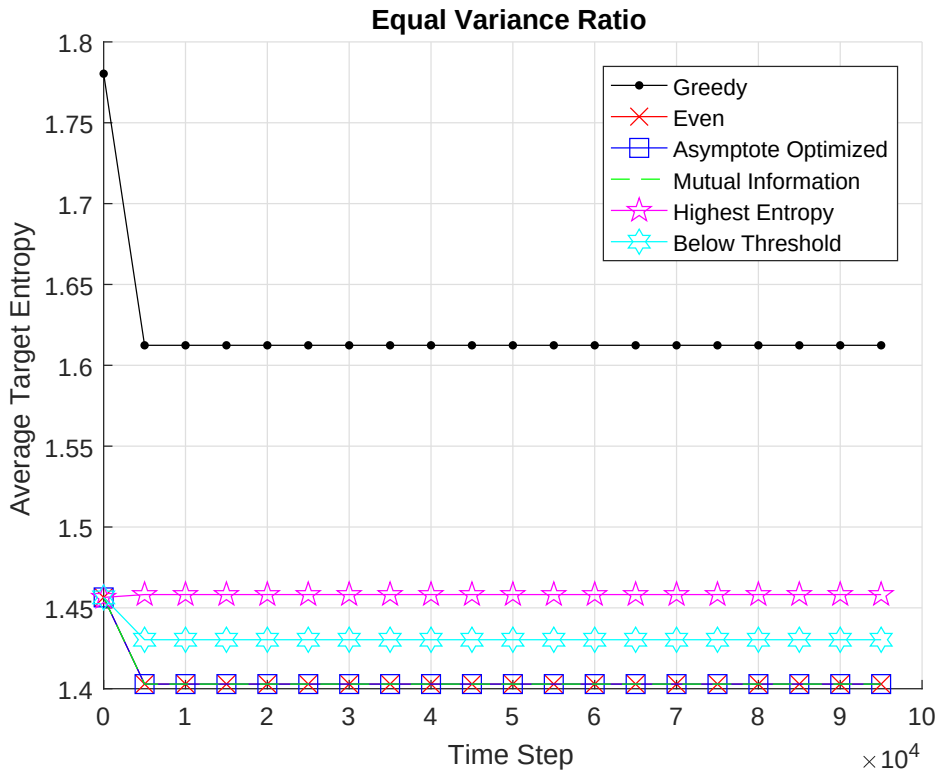

Figure 4.3: Entropy comparison for equal variance ratio between channels.

matrix of

$$\Sigma_m = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 5050 & 0 \\ 0 & 0 & 10000 \end{bmatrix}. \tag{4.36}$$

Figures 4.4 and 4.5 display each channel's entropies over time, with the highest entropy marked with a data point, for all six power allocation strategies. Ranking the approaches based on the lowest maximum asymptotic entropy, from best to worst, would be "Minimize Highest Entropy", "Minimize Below Threshold", "Greedy", "Even"/"Asymptote Optimized" (tied), and "Mutual Information Optimized".

If minimizing the channels below a specified variance threshold, in this case a threshold value of 3.4, would achieve the system objective, "Minimize Highest Entropy" and "Minimize Below Threshold" would be the only two approaches to achieve this objective for all three channels. For the tracking threshold, the "Minimize Highest Entropy" and "Minimize Below Threshold" approaches would be the only two approaches to maintain track on all three targets. The "Greedy" approach would lose all three targets, the "Mutual Information Optimized" approach would lose two targets, and the "Even"/"Asymptote Optimized" approaches would both lose one target.

The parameter model for this chapter assumes the channel dynamics are independent of each other, which is not always the case. Sometimes the parameter state variability is dependent on the surrounding parameters. For the target tracking example, a swarm of targets that move together have correlated dynamic variances. Channels with dependent state variability are investigated in next chapter.

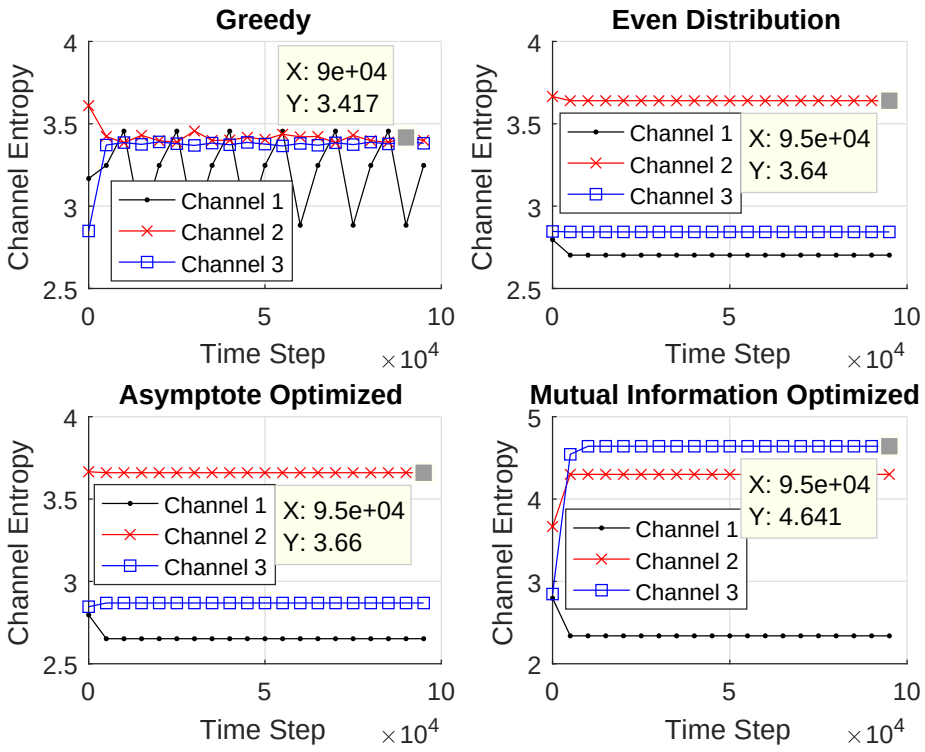

Figure 4.4: Four entropy comparisons for three channels.

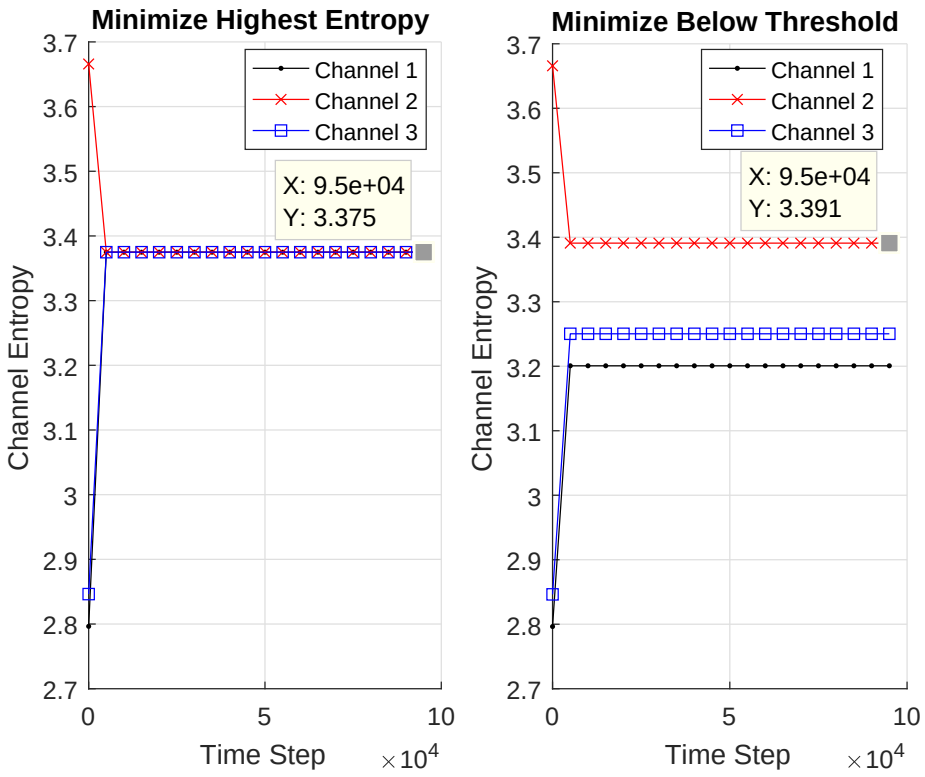

Figure 4.5: Highest entropy minimization for three channels.

Chapter 5

Correlated Targets

5.1 System Model

Although the simple model detailed in Chapter 4 is useful, it assumed an independence between the parameters' variability. While many applications estimate only parameters that adhere to the independence assumption of Chapter 4, other applications include parameters whose variability are correlated with other parameters. By removing the independence assumption, more realistic scenarios can be simulated including interactions between the parameters. In the target tracking application, correlated parameters occur when the objects are moving in a coordinated pattern. In this case, when one target changes its' state vector, another target performs a similar maneuver. The mathematical system model for this interaction between channels is given in (3.1). As will be shown, the equations derived in this chapter can be used for the independent scenarios and simplify to the equations in Chapter 4 when independent parameters are substituted into the equations. For a target tracking application, the channels from the parallel channel model in Figure 3.1 represent each target for tracking. Therefore, channels and targets are used interchangeably throughout this chapter.

As the channel input varies over time, it is modeled as a Gaussian random variable. We characterize the parameter using a Bayesian estimation paradigm with the previous time step's Gaussian distribution as the prior and updated by measuring the parameter value through the parallel channel model shown in Figure 3.1. The belief state, which is used to quantify the Bayesian approach and characterize the parameters' probability distribution [8], is

$$\mathbf{b}_k = p(\mathbf{x}_k|\mathbf{y}_k) \sim \mathcal{N}(\mu_k, \mathbf{\Sigma}_k) \tag{5.1}$$

where μ_k is the estimated parameter vector at time k and $\mathbf{\Sigma}_k$ is the belief covariance matrix at time k. Since an active sensor does not have control over the target parameter, the target parameter is modeled as varying between measurements according to a Gaussian distribution. To model this variation between measurements, the belief covariance matrix in (5.1) is increased between measurements by a constant referred to as the multi-parameter dynamic covariance, $\mathbf{\Delta}$. In contrast to $\mathbf{D}$ in Chapter 2, $\mathbf{\Delta}$ combines the dynamic variances of each parameter, δ_n^2, in a single covariance matrix. Unlike Chapter 4, $\mathbf{\Delta}$ is not necessarily diagonal, with the cross-terms accounting for correlation between targets' maneuverability.

The increase in belief covariance matrix due to the channel's variability occurs between each measurement, so if no measurement occurs, then the parameter's covariance matrix continues to grow. Although the increase occurs between each measurement, the actual dynamic covariance matrix is a constant and doesn't change, evidenced by the lack of a time index. Applying the increase in the covariance matrix due to parameter variability after the time

elapsed between measurements, (5.1) becomes

$$\hat{\mathbf{b}}_{k+1} \sim \mathcal{N}(\mu_k, \mathbf{\Sigma}_k + \mathbf{\Delta}). \tag{5.2}$$

By measuring the parameter, the increase in the belief state's uncertainty caused by the dynamic variance can be decreased, depending on the quality of the measurement. By intelligently controlling the measurement strength, a system can achieve better results for a given metric. The belief state after the system takes a measurement is

$$\mathbf{b}_{k+1} \sim \mathcal{N}\left(\left[\mathbf{P}_{k+1}\mathbf{N}^{-1} + (\mathbf{\Sigma}_k + \mathbf{\Delta})^{-1}\right]^{-1}\left[(\mathbf{\Sigma}_k + \mathbf{\Delta})\mu_k + \mathbf{N}^{-1}\mathbf{y}_{k+1}\right], \left[\boldsymbol{N}^{-1}\boldsymbol{P}_k + (\mathbf{\Sigma}_k + \mathbf{\Delta})^{-1}\right]^{-1}\right) \tag{5.3}$$

where $\mathbf{y}_{k+1}$ is the subsequent measurement [50]. Since the measurement variance, $\mathbf{N}$, and dynamic variance, $\mathbf{\Delta}$, do not change over time, denoted by the lack of a time subscript, a constant power allocation, $\mathbf{P}_{k+1}$, causes the system to converge to a steady state over time. The steady state occurs when the amount of information extracted by the measurement equals the increase in uncertainty due to the parameter's variability. The resulting steady state also occurred for the uncorrelated model in Chapter 4, which was shown in the simulations from Section 4.3.

5.2 Resource Allocation Strategies

For a large number of channels, intelligent approaches to resource allocation are important to minimize parameter uncertainty. To obtain intelligent approaches, this section derives numerical solutions for $\sqrt{\mathbf{P}_k}$ based on the metrics

given in Chapter 3. Two allocation strategies detailed in this section focus on the overall system entropy metric, while one strategy minimizes the highest parameter entropy, and the last approach maintains every parmeter's entropy below a given threshold.

5.2.1 Mutual Information Maximization

Similar to Chapter 4, the first approach maximizes the mutual information across all channels for the correlated model. After accounting for the covariance matrix increase between measurements for mutual information in (3.9), the objective function at the k^{th} time step is

$$R_{tot}^{(k)} = \frac{1}{2}\log\left[\det\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I}\right)\right]. \tag{5.4}$$

By applying the Lagrangian multiplier technique with the constraint in (3.2), (5.4) becomes

$$R_{tot}^{(k)} = \frac{1}{2}\log\left[\det\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I}\right)\right] - \lambda\left[\mathrm{tr}\left(\mathbf{P}_{k+1}\right) - P\right]. \tag{5.5}$$

Using properties from [51], the derivative of (5.4) with respect to the power allocated to channel n at time step $k+1$, $\rho_{n,k+1}$, is

$$\frac{\partial R_{tot}^{(k)}}{\partial \rho_{n,k+1}} = \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I}\right)^{-1}\mathbf{\Gamma}_{n,k}\right] - \lambda. \tag{5.6}$$

where

$$\mathbf{\Gamma}_{n,k} = \frac{\partial}{\partial \rho_{n,k+1}} \left((\mathbf{\Sigma}_k + \Delta)\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right) = \begin{pmatrix} 0 & \cdots & 0 & \frac{\sigma^2_{1,n,k}+\delta^2_{1,n}}{\nu^2_n} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{\sigma^2_{2,n,k}+\delta^2_{2,n}}{\nu^2_n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\sigma^2_{N,n,k}+\delta^2_{N,n}}{\nu^2_n} & 0 & \cdots & 0 \end{pmatrix}, \quad (5.7)$$

ν^2_n is the n^{th} diagonal value of $\mathbf{N}$, $\delta^2_{l,n}$ is the l^{th} row and n^{th} column value of the $\mathbf{\Delta}$ matrix, $\sigma^2_{l,n,k}$ is the l^{th} row and n^{th} column value of the $\mathbf{\Sigma}$ matrix at time k, and λ is the Lagrangian multiplier.

Maximizing (5.6) requires setting it equal to zero. Solving for λ then becomes

$$\lambda = \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{n,k} \right]. \quad (5.8)$$

Deriving (5.8) for every channel yields

$$\begin{aligned} \lambda = \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{1,k} \right] = \\ \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{2,k} \right] = \\ ... = \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{N,k} \right]. \end{aligned} \quad (5.9)$$

Generalizing (5.9) for every channel, maximizing mutual information is obtained by numerically solving

$$\begin{aligned} \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{n,k+1} \right] = \\ \mathrm{tr}\left[\left((\mathbf{\Sigma}_k + \mathbf{\Delta})\mathbf{N}^{-1}\mathbf{P}_{k+1} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{m,k+1} \right], \forall n \neq m \end{aligned} \quad (5.10)$$

for $\mathbf{P}_{k+1}$ with the constraint in (3.2).

The result in (5.10) converges to (4.13) by applying the independence assumption given by

$$\mathbf{\Sigma}_k = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \sigma_N \end{pmatrix} \tag{5.11}$$

and

$$\mathbf{\Delta} = \begin{pmatrix} \delta_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \delta_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \delta_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \delta_N \end{pmatrix}. \tag{5.12}$$

Similar to Chapter 4, during each measurement, if more information is gained than the uncertainty inserted by the parameter's dynamic variance, then the overall belief state entropy will ultimately decrease. By reducing the system entropy, less information exists for the system to gain through measurement. After many iterations, this process will eventually reach a point where the amount of information gained equals the uncertainty introduced through the dynamic variance at each time step. In a similar fashion, if less information is gained than the uncertainty inserted by the parameter's dynamic variance, the opposite process will happen, but with the same steady state result. This convergence process is experimentally shown in Section 5.3.1 through simulation.

Due to this convergence, allocating power to maximize mutual information eventually produces an asymptotic parameter covariance matrix, Σ_∞, where the prior covariance matrix is equal to the posterior covariance matrix. To calculate the asymptotic covariance matrix, the posterior covariance matrix after measurement in (5.3) is set equal to the prior covariance matrix (before dynamic variance is added on the previous step) to obtain

$$\boldsymbol{\Sigma}_\infty = \left[\boldsymbol{N}^{-1}\boldsymbol{P}_\infty + (\boldsymbol{\Sigma}_\infty + \boldsymbol{\Delta})^{-1}\right]^{-1}. \tag{5.13}$$

Solving for the asymptotic power matrix, $\mathbf{P}_\infty$, yields

$$\mathbf{P}_\infty = \mathbf{N}\boldsymbol{\Sigma}_\infty^{-1} - \mathbf{N}(\boldsymbol{\Sigma}_\infty + \boldsymbol{\Delta})^{-1}, \tag{5.14}$$

which depends on the measurement noise, dynamic, and parameter covariance matrices. Substituting (5.14) for the power matrix, (5.10) becomes

$$\operatorname{tr}\left[\left(\mathbf{I} + \boldsymbol{\Sigma}_\infty^{-1}\boldsymbol{\Delta}\right)^{-1}\boldsymbol{\Gamma}_{n,k}\right] = \operatorname{tr}\left[\left(\mathbf{I} + \boldsymbol{\Sigma}_\infty^{-1}\boldsymbol{\Delta}\right)^{-1}\boldsymbol{\Gamma}_{m,k}\right], \forall n \neq m. \tag{5.15}$$

To solve the system of $N-1$ equations in (5.15), the power distribution at infinity solved in (5.14) must be substituted into the power constraint to derive the N^{th} equation. The power constraint becomes

$$\operatorname{tr}\left[\mathbf{P}_\infty\right] = \operatorname{tr}\left[\mathbf{N}\boldsymbol{\Sigma}_\infty^{-1} - \mathbf{N}(\boldsymbol{\Sigma}_\infty + \boldsymbol{\Delta})^{-1}\right] = P. \tag{5.16}$$

The asymptotic variance of the mutual information maximization approach can then be calculated by solving this system of equations.

Once again, substituting (5.11) and (5.12) causes (5.15) and (5.16) to converge to (4.19) and (4.20), respectively. Equations (5.15) and (5.16) show that

even though the mutual information strategy only allocates power for the next measurement, the long-term power allocation, and subsequently the long-term performance, can still be calculated.

5.2.2 Entropy Asymptote Minimization

While the mutual information approach focuses on decreasing the uncertainty for the next time step, increasing the optimization beyond a single step results in exponentially increased mathematical complexity. On the other hand, a constant power allocation will result in long-term steady state behavior. By leveraging this consequence, a long-term solution can be derived by assuming a constant (over time) power allocation and optimizing the system's asymptotic entropy. Due to the unchanging nature of the power matrix for this approach, the power distribution at infinity is the same as the power applied at time k, and the time index can be removed. Substituting the constant power matrix, $\mathbf{P}$, and rearranging (5.13) yields

$$\boldsymbol{\Delta}^{-1}\boldsymbol{P}^{-1}\boldsymbol{N} = \boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_{\infty}\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_{\infty} + \boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_{\infty}. \tag{5.17}$$

After completing the square, (5.17) becomes

$$\boldsymbol{\Delta}^{-1}\boldsymbol{P}^{-1}\boldsymbol{N} + \frac{1}{4}\boldsymbol{I} = \left(\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_{\infty} + \frac{1}{2}\boldsymbol{I}\right)\left(\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_{\infty} + \frac{1}{2}\boldsymbol{I}\right). \tag{5.18}$$

Simplifying and solving for the asymptotic entropy results in

$$\boldsymbol{\Sigma}_{\infty} = -\frac{\boldsymbol{\Delta}}{2} \pm \left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^{2} + \boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}. \tag{5.19}$$

Substituting (5.19) into the equation for entropy in (3.6) and using the Lagrangian multiplier technique to include the power constraint in (3.2), the objective function to minimize is

$$R_{tot} = h(X_k) = \frac{1}{2}\log\left[\det\left(\pi e \cdot \left[-\frac{\mathbf{\Delta}}{2} + \left[\left(\frac{\mathbf{\Delta}}{2}\right)^2 + \mathbf{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right]\right)\right] - \lambda\left[\mathrm{tr}(\mathbf{P}) - P\right]. \quad (5.20)$$

The derivative of (5.20), derived using theorems from [51], with respect to the power of channel n is

$$\frac{\partial R_{tot}}{\partial \rho_n} = \mathrm{tr}\left[\frac{1}{2}\left(-\frac{\mathbf{\Delta}}{2} + \left[\left(\frac{\mathbf{\Delta}}{2}\right)^2 + \mathbf{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1} \left[\left(\frac{\mathbf{\Delta}}{2}\right)^2 + \mathbf{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2} \mathbf{\Upsilon}_n\right] - \lambda \quad (5.21)$$

where

$$\mathbf{\Upsilon}_n = \begin{pmatrix} 0 \cdots 0 & -\frac{\delta_{1,n}^2\nu_n^2}{\rho_n^2} & 0 \cdots 0 \\ 0 \cdots 0 & -\frac{\delta_{2,n}^2\nu_n^2}{\rho_n^2} & 0 \cdots 0 \\ \vdots \ddots \vdots & \vdots & \vdots \ddots \vdots \\ 0 \cdots 0 & -\frac{\delta_{N,n}^2\nu_n^2}{\rho_n^2} & 0 \cdots 0 \end{pmatrix}. \quad (5.22)$$

Minimizing (5.21) requires setting it equal to zero. Solving for λ then becomes

$$\lambda = \mathrm{tr}\left[\frac{1}{2}\left(-\frac{\mathbf{\Delta}}{2} + \left[\left(\frac{\mathbf{\Delta}}{2}\right)^2 + \mathbf{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1} \left[\left(\frac{\mathbf{\Delta}}{2}\right)^2 + \mathbf{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2} \mathbf{\Upsilon}_n\right]. \quad (5.23)$$

Deriving (5.23) for every channel yields

$$\lambda = \mathrm{tr}\left[\frac{1}{2}\left(-\frac{\boldsymbol{\Delta}}{2}+\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1}\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2}\boldsymbol{\Upsilon}_1\right]$$
$$= \mathrm{tr}\left[\frac{1}{2}\left(-\frac{\boldsymbol{\Delta}}{2}+\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1}\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2}\boldsymbol{\Upsilon}_2\right]$$
$$= ... = \mathrm{tr}\left[\frac{1}{2}\left(-\frac{\boldsymbol{\Delta}}{2}+\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1}\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2}\boldsymbol{\Upsilon}_N\right]. \quad (5.24)$$

Generalizing (5.24) for every channel, minimizing the asymptotic entropy is obtained by numerically solving

$$\mathrm{tr}\left[\frac{1}{2}\left(-\frac{\boldsymbol{\Delta}}{2}+\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1}\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2}\boldsymbol{\Upsilon}_n\right] =$$
$$\mathrm{tr}\left[\frac{1}{2}\left(-\frac{\boldsymbol{\Delta}}{2}+\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{1/2}\right)^{-1}\left[\left(\frac{\boldsymbol{\Delta}}{2}\right)^2+\boldsymbol{\Delta}\mathbf{P}^{-1}\mathbf{N}\right]^{-1/2}\boldsymbol{\Upsilon}_m\right],$$
$$\forall n \neq m. \quad (5.25)$$

The Lagrangian multiplier technique uses the constraint in (3.2) to integrate the objective functions for each channel as a system of equations in (5.24). By substituting another channel's derivative for λ to derive (5.25), the constraint added by the Lagrangian multiplier is removed from the set of equations. Therefore, the constraint in (3.2) must be included when solving (5.25).

When the dynamic covariance matrix has complete correlation between any

two or more channels, (5.25) is not defined due to the inverse. When this situation occurs, only the correlated channel with the lowest measurement noise variance should be included while the other completely correlated channels should be ignored. If the dynamics of multiple channels are completely correlated, then a change in one channel will be reflected with the same change in the other channel. Therefore, the sensor only needs to measure one channel, and can predict the value of both channels from one measurement. The lowest measurement noise variance channel is measured because the sensor can gain information from that channel easier than the other channels.

To verify the accuracy of (5.25), the system of equations are consistent with the uncorrelated versions in (4.28). In order to compare with the uncorrelated case in Section 4.2.2, (5.11) and (5.12) must be substituted into (5.25). After substituting, (5.25) converges to the uncorrelated version in (4.28).

5.2.3 Highest Entropy Asymptote Minimization (Minimax)

In contrast to the approaches in Sections 5.2.1 and 5.2.2, which minimize total system uncertainty, the highest entropy minimization approach reduces the channel with the highest asymptotic entropy as much as possible without increasing the other channels above it. For a tracking application, a threshold entropy could symbolize the amount of uncertainty allowable before the system loses track of a target. Therefore, the highest entropy minimization approach will calculate the lowest possible uncertainty that all the channels can simultaneously reach in their steady state. The resulting power allocation reduces the most uncertain channel, causing the other channels to increase in uncertainty, until all the channel entropies are equal. Due to this property, the

asymptotic entropy for each channel are set equal for this minimax approach.

Instead of a closed-form solution, which is not possible for this approach, the allocation is numerically calculated by equating the diagonal values of the asymptotic parameter covariance matrix. The asymptotic parameter covariance matrix is given by (5.19). Since equating (5.19) for each channel only generates $N-1$ equations, the constraint in (3.2) must be included to be able to solve the system of equations.

Although the minimax approach determines the allocation necessary to reduce the highest entropy, the resulting solution might not be sufficient to reach the system's threshold uncertainty. Since all the channels' uncertainties are equal, none of the channels would be maintained under the desired entropy. In the tracking example, the system would lose track of all the targets. Instead, the algorithm can ignore the channel requiring the most power (i.e. this channel will receive no power) and recalculate the allocation without it. This would sacrifice one target to be able to track the remaining targets. The allocation algorithm can repeat this procedure until the desired entropy is reached. In other words, the hardest channels to maintain should be removed until the residual channels can be reduced below the system's threshold.

5.2.4 Variance Asymptote Minimization Below a Threshold

The previous approaches optimized either an average-type metric or a min-max type metric, but did not provide the opportunity to compromise between the two objectives. The average-based metrics do not offer the desired customization for a specified system, while the min-max strategy doesn't perform the optimal allocation for any remaining power budget after the threshold has

been achieved. The variance asymptote minimization approach combines the two metrics to create a configurable compromise between the average type and min-max type optimization. This compromise is achieved by first calculating the power allocation necessary for every channel to reach the desired asymptotic entropy. Once all the channels have enough power to attain the threshold, the remaining power is allocated to minimize the overall system uncertainty. Similar to the minimize highest entropy asymptote approach, if not enough power is avaiable to minimize all the channels below the threshold, the channel requiring the most power must be dropped, while applying the allocation approach to the remaining targets. Using the tracking example, this approach is able to track as many targets as possible with an acceptable level of performance, while lowering the overall system uncertainty with the remaining energy.

The compromise between each approach is determined by the threshold value. A higher threshold doesn't require much power to be allocated to each channel to reach the desired threshold. Therefore, a higher threshold causes this approach to perform more like the entropy asymptote minimization approach. In contrast, a lower threshold requires more power allocated to each channel to attain the required entropy. Therefore, a smaller threshold causes this approach to allocate power closer to the highest entropy minimization approach. To calculate the allocation for this approach, the asymptotic entropy given in (5.19) is set less than the threshold matrix, $\mathbf{E}$. Since both the power allocation matrix and the threshold matrix are diagonal, setting (5.19) less than $\mathbf{E}$ reduces each diagonal of (5.19) below the system threshold ϵ. After

factoring $\mathbf{\Delta}$ from (5.19), this is mathematically shown by

$$\mathbf{E} > \mathbf{\Sigma}_{\infty} = \mathbf{\Delta}\left[\left(\mathbf{\Delta}^{-1}\mathbf{P}^{-1}\mathbf{N} + \frac{1}{4}\mathbf{I}\right)^{1/2} - \frac{1}{2}\mathbf{I}\right] \tag{5.26}$$

where

$$\mathbf{E} = \begin{pmatrix} \epsilon & 0 & \cdots & 0 \\ 0 & \epsilon & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon \end{pmatrix}. \tag{5.27}$$

Solving for $\mathbf{P}$, the amount of power for each channel to reach the required threshold is

$$\mathbf{P} > \mathbf{N}\left[\mathbf{E}\mathbf{\Delta}^{-1}\mathbf{E} + \mathbf{E}\right]^{-1}. \tag{5.28}$$

Because the threshold is a system specification related to acceptable tracking performance, it is assumed to be the same across all channels. For systems with resource variations, this threshold can be based on the resources selected for each channel. Similar to the entropy minimization approach, when two or more channels are completely correlated, a solution does not exist for (5.28). In this case, only the correlated channel with the lowest noise power should be measured. The other correlated channels should be ignored and receive no power.

The power matrix calculated by (5.28) must adhere to the power constraint in (3.2). When (3.2) is not met, sufficient resources do not exist to obtain the required performance for every channel. When this occurs, (5.28) should be recalculated without the channel allocated the most power. By replicating this procedure, a power allocation that does meet the constraint in (3.2) will be found. Once the amount of power allocated by (5.28) meets the requirement

in (3.2), then the entropy minimization approach is used for the remaining power allocation.

5.3 Results

To compare the performance of each resource allocation technique described in Section 5.2, this section highlights multiple scenarios and evaluates metrics from Chapter 3. Two baseline approaches are included for comparison purposes: 1) evenly distributing the total power across the channels at each time step, and 2) allocating all power to the channel with the highest uncertainty at each time step (also called the greedy approach).

Two metrics are selected to compare the performance of each allocation strategy. The system's average MSE, an overall system metric, is the first metric used to compare the different approaches' performance. A low average MSE indicates lower total error and better estimation of the channels as a whole. The average MSE metric is also related to the amount of total information extracted by each approach over time.

The second metric is the number of failed channels, which counts the number of channels that rise above an unacceptable amount of uncertainty, designated by a MSE threshold value. This metric relates to a system specification that defines when it is unable to reliably estimate the value of a channel whose uncertainty has become too large. For the tracking example, the possibility of losing track of a target is comparable to the failed channel metric.

While each strategy strives to minimize both metrics, the amount of available resources determines the ease with which that is possible. For this section, the simulated system is provided only 50 units (e.g., Watts) of power to allocate at each time step. When the resources are scarce, i.e. the amount of

Table 5.1: Correlated Targets Radar Parameters.

Unambiguous Range	5 km
Blind Range	100 m
Bandwidth	50 MHz
Noise Power	-201 dB

transmit power per target is low, then it is more difficult to reduce both the overall uncertainty and the number of failed channels. The performance of each approach with a limited amount of resources is numerically investigated in this section.

While the parameter variance transition between time steps is well defined by the posterior variance equation in (5.3), the initial parameter variances are not strictly defined, and must be selected by the system. In this section, the parameter variances' initial starting point is determined by calculating the asymptotic variance in (5.19) for an evenly divided amount of power. Since (5.19) might set the initial parameter variance above the failed channels threshold (causing the parameter to be initially considered failed), any initial parameter variances above the variance threshold are reduced to be the same as the threshold value.

For each simulation, the unique channel measurement variances are generated according to a uniform distribution between 1.8 and 3×10^5, and the distinctive set of parameter dynamic variances are generated according to a Rayleigh distribution with a scale parameter of 0.31. These distributions are analogous to tracking targets with a range of 1.5 to 30 km, according to the radar range equation, and accelerating at a mean of 2.3 g's [41] with the radar properties specified in Table 5.1. The noise covariance matrix, $\mathbf{N}$, and the dynamic covariance matrix, $\mathbf{\Delta}$ are modeled as constant over time.

With these specifications, this section proceeds in Section 5.3.1 with a simulation of one particular realization of measurement and dynamic variances, and comparing each approach for the average MSE metric. Although one realization of measurement and dynamic variances is useful for demonstrating how each approach reaches a steady state, it is not considerably useful for comparing the approaches for a broader range of scenarios. Section 5.3.2 addresses this issue by comparing each approach using the average MSE metric over 1,000 Monte Carlo simulations for a varying number of channels and a variance threshold of 210. To account for the second metric, the same Monte Carlo simulations as Section 5.3.2 are reported with the failed channels metric for comparison in Section 5.3.3. The simulations reported in Sections 5.3.1, 5.3.2, and 5.3.3 still included independent channels to compare the approaches from Section 5.2. In contrast, the simulation reported in Section 5.3.4 varies the amount of correlation between the targets for 100 Monte Carlo simulations with the same noise and dynamics as Sections 5.3.2 and 5.3.3 for a constant 50 channels and a constant variance threshold of 175. The average MSE metric for correlated channels is not analyzed due to the lack of variability between each approach when correlation is varied.

5.3.1 Single Realization

A single realization is important for demonstrating the typical behavior of each approach as a function of time. To this end, the average MSE performance of each approach in estimating 50 channels is displayed versus log time in Figure 5.1. For this example, the threshold approach, in contrast with the other approaches, does not reach an a steady state. It is important to note that for a variance threshold of 210, even though the majority of approaches

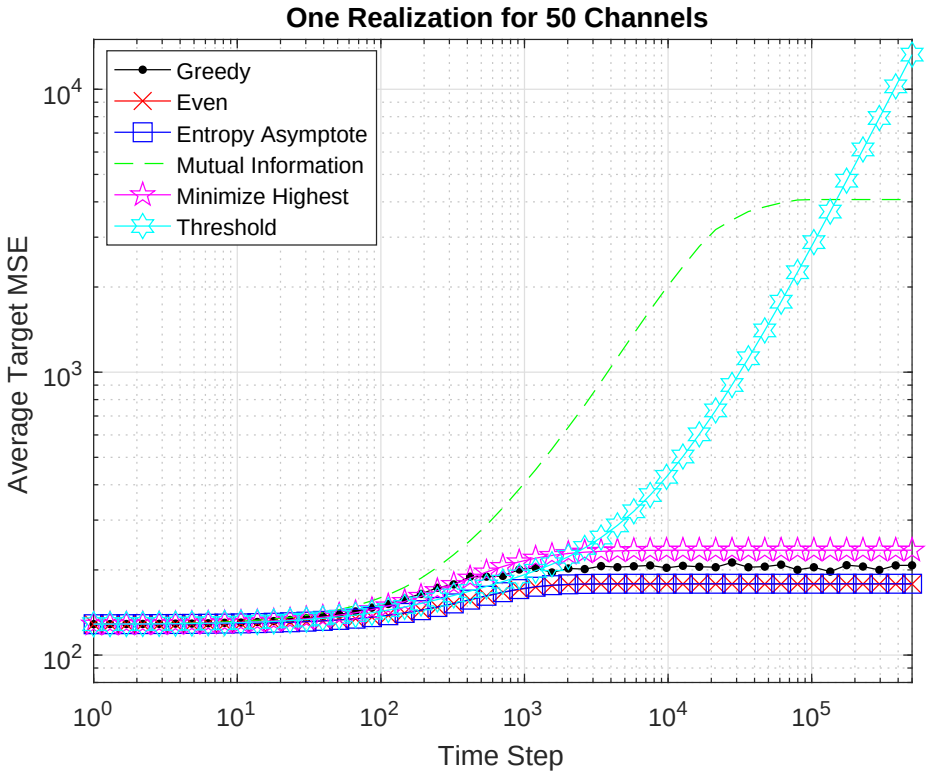

Figure 5.1: Average MSE comparison for a single realization.

have an average MSE below the threshold, the individual channels of these approaches may not necessarily be under the same threshold. To quantify the comparison, the MSE of each approach averaged across both time and targets is given in Table 5.2 with the same relative results between the approaches as Figure 5.1.

In Figure 5.1 the asymptotic entropy and the even distribution approach overlap with the lowest average MSE values. In Table 5.2, the even distribution is shown to have a slightly better result due to the rounding when calculating the power allocation for the asymptotic optimized approach. Both the minimization below a threshold and minimize highest entropy approaches have a higher average MSE in order to reduce the maximum number of individual channels under the chosen threshold. Although all the approaches

Table 5.2: MSE averaged over number of channels for single realization.

Approach	Average MSE
Greedy	204.8
Even Distribution	178.1
Asymptotic Entropy	178.7
Mutual Information	3,954.8
Minimize Highest	234.7
Threshold	6,715.6

overlap in the initial time steps of Figure 5.1, a closer look confirms that the mutual information approach initially has the lowest average entropy. This strategy is not able to maintain that benefit over the long term. As long as the variance threshold is changed accordingly, changing the number of channels will only change the amount of uncertainty for each strategy, but will not affect the order. The order in which the approaches are defined from best to worst performing is only affected a change in SNR, dynamics, or threshold (in the case of the minimize below threshold approach).

5.3.2 Monte Carlo Average Mean Squared Error

To complement the single scenario results of Section 5.3.1, Monte Carlo simulations are able to compare the approaches over a range of scenarios. Each Monte Carlo simulation captures the results for a specific scenario that are combined to generate an overall system performance for a general comparison. With the same dynamic and noise distributions as in Section 5.3.1, Figure 5.2 displays the results for a Monte Carlo simulation with a range of channels between 10 and 100. By averaging the MSE across the differing number of channels, a more quantitative result is given in Table 5.3 for comparison.

The only difference in performance order between each approach, compared

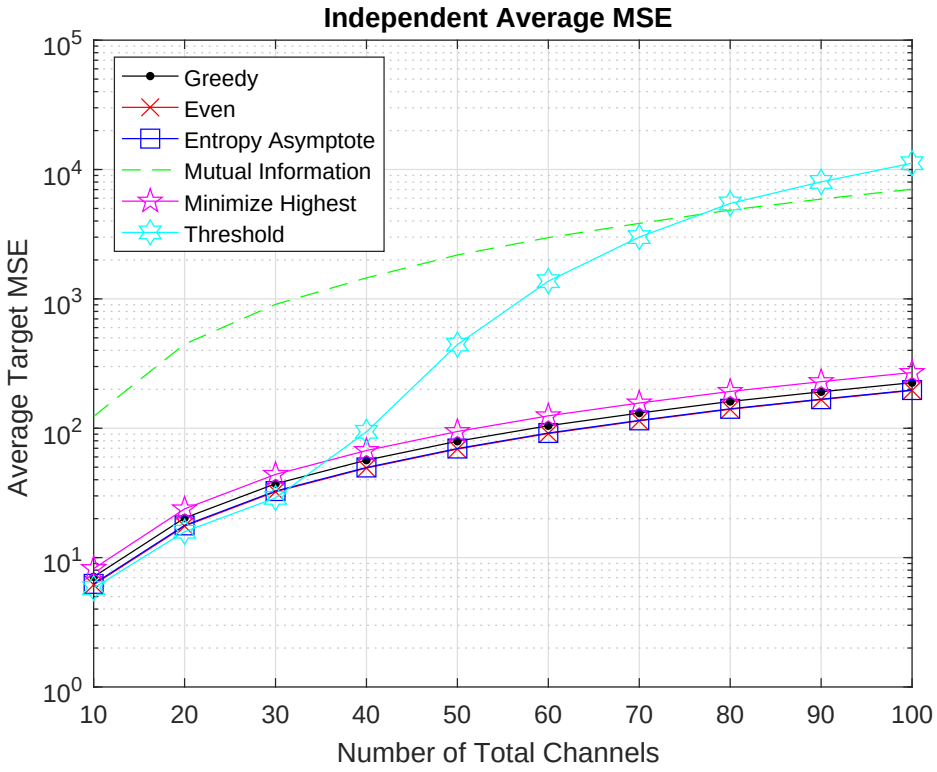

Figure 5.2: Monte Carlo simulation comparing average MSE over a varying number of targets.

Table 5.3: MSE averaged over number of simulations for Monte Carlo simulation.

Approach	Average MSE
Greedy	101.4
Even Distribution	88.2
Asymptotic Entropy	88.9
Mutual Information	2,974.4
Minimize Highest	121.0
Threshold	2,964.3

to Section 5.3.1, is the threshold strategy. This difference is caused by the constant threshold selected for every Monte Carlo trial, even while the number of channels is changing. Since the threshold stays constant while the number of channels increases, the average metric performance of the threshold strategy comparably deteriorates as the amount of power available per target decreases. As the number of channels continues to increase, the resources are not sufficient for the threshold approach to reduce the higher number of channels under the threshold. To increase the resources per target tracked, the threshold approach causes a substantial reduction in average performance by allocating zero power to the channels with the highest MSE.

5.3.3 Independent Failed Channels

The average MSE metric used for comparison in Sections 5.3.1 and 5.3.2 does not account for systems that require a limited amount of uncertainty for a channel to be useful. For a tracking application, losing track of targets due to an excess of uncertainty is more important to the continuity of a tracker versus the reduction of overall system uncertainty. Continually losing tracks wastes signal processing resources by continuously detecting, re-initializing, and then deleting the tracks. Instead, by determining which tracks are not capable of being tracked due to resource limitations, the system can remove those tracks from re-initialization. The benefit of ignoring these channels becomes clear in resource constrained scenarios.

Similar to Section 5.3.2, we simulate the failed channel metric using a Monte Carlo simulation with the same characteristics. The simulation counts and records the number of channels that increase above the given threshold. Since the metric describes the channels as failed, once a channel variance

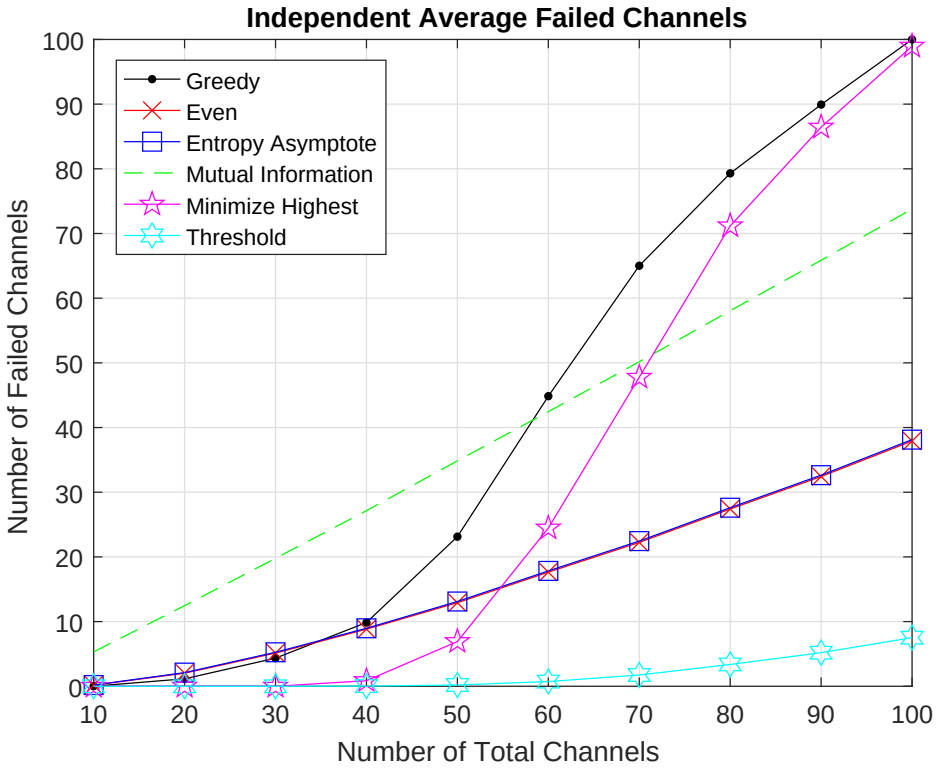

Figure 5.3: Failed channels comparison for a varying number of uncorrelated channels.

increases beyond the threshold, it is considered broken and unrecoverable. Figure 5.3 shows the number of failed channels averaged across each Monte Carlo simulation. The percentage of this average, summed and divided by the number of channels, for all the Monte Carlo simulations is given in Table 5.4.

Since the minimize below threshold approach actively aims to maintain the channels below the same threshold as the metric, it achieves the best performance for the threshold metric. The minimize highest entropy approach has good performance when the amount of resources is abundant, but starts drastically degrading to become the worst approach when resources are scarce because it can't keep all the targets under the desired threshold. Much like the minimize highest entropy approach, the greedy approach does well with an

Table 5.4: Percentage of failed channels for Monte Carlo simulation of independent channels.

Approach	Percentage of Failed Channels
Greedy	75.9%
Even Distribution	30.3%
Asymptotic Entropy	30.6%
Mutual Information	70.9%
Minimize Highest	61.2%
Threshold	3.4%

abundance of resources but drastically gets worse with less resources per channel. The even distribution and entropy asymptote minimization approaches have a similar performance with the number of failed channels metric. While these approaches are not the best for this metric, they still perform relatively well compared to the remaining strategies in the resource constrained scenarios.

5.3.4 Correlated Failed Channels

While the simulations reported in Sections 5.3.1, 5.3.2, and 5.3.4 only included channels with uncorrelated dynamics, this section varies the amount of correlation between the channels. Comparing the performance of each approach on channels with correlated dynamics is useful for channels whose values change in relation to the other channels due to external forces such as vicinity and environmental factors. For the tracking example, channels with correlated dynamics are analogous to targets that fly in formation or are affected by the same weather conditions. Instead of varying the number of targets, as in Section 5.3.3, the Monte Carlo results displayed in Figure 5.4 have a static number of channels set at 10 and the amount of correlation between channels varies. The percentage of failed channels for each approach is given in Table 5.5.

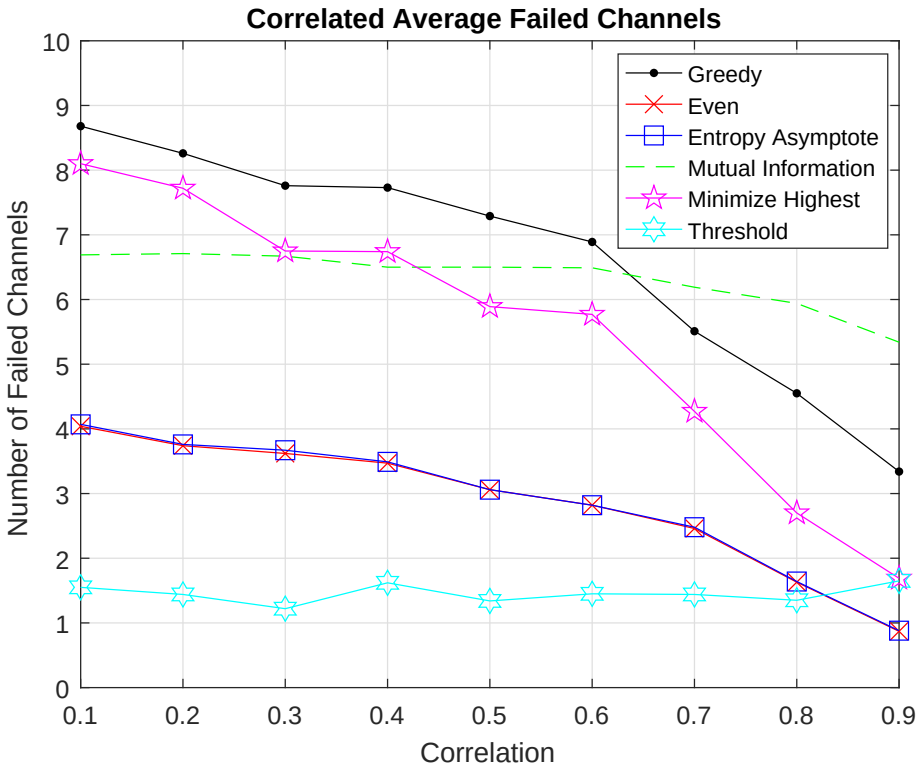

Figure 5.4: Failed channels comparison for a varying amount of correlation.

Table 5.5: Percentage of failed channels for a Monte Carlo simulation of correlated channels.

Approach	Percentage of Failed Channels
Greedy	66.7%
Even Distribution	28.6%
Asymptotic Entropy	28.7%
Mutual Information	63.4%
Minimize Highest	55.1%
Threshold	14.5%

The added correlation, compared to the independent Monte Carlo simulations in Section 5.3.3, is able to improve the performance of every approach except the thresholding approach. The approaches that improve as the correlation increases are able to prioritize the channels with a lower measurement variance to gain knowledge about all the targets. When the correlation between targets approaches one, the even distribution and minimal entropy approaches actually have an average number of failed channels less than the thresholding approach. While the thresholding approach doesn't improve with correlation, the mutual information maximization approach also is not able to capitalize much with the increased correlation compared to the other approaches, and performs the worst with high correlation.

Chapter 6

Aperture Allocation

6.1 System Model

While the models in Chapters 4 and 5 provide introductory examples regarding a system's ability to manage resources toward optimizing multiple parameter estimation, they do not realistically model the process of measuring the parameter's value, especially for a radar system. The measurement variances in Chapters 4 and 5 would realistically vary based on the signal-to-noise ratio (SNR) of the channel. To account for this modification in the measurement error, a mathematical model is defined at time instance k as

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{e}_k \quad k = 1, 2, ..., K \tag{6.1}$$

where $\mathbf{x}_k$ is the vector of parameters to be estimated, $\mathbf{e}_k$ is the measurement error vector, $\mathbf{y}_k$ is the measured data vector, and K is the total number of time steps.

In this model, we implement a measurement error variance that is a nonlinear function of SNR. To calculate the nonlinear function, we simulated the azimuth phase ratio estimation using the phase monopulse estimation technique for a pulse-Doppler radar. After simulating Monte Carlo trials for 1,000

realizations of noise, the error for each measurement was averaged for a constant SNR. The estimation error values displayed in Figure 6.1 were determined by performing the Monte Carlo simulation for a range of SNR values. The function of error versus SNR defined in Figure 6.1 determines the measurement error variance based on the measurement SNR. Since the aperture allocation affects the measurement SNR through the aperture power and gain, this chapter applies aperture allocation techniques for the nonlinear measurement model to determine the SNR and subsequently the measurement error variance for each measurement.

The nonlinear measurement model in Figure 6.1 has a low-SNR regime with a constant error for each SNR value and a high-SNR regime with an inversely linear relationship between error and SNR. The low-SNR error is constant to represent the minimal information received from a low-SNR measurement as if no measurement occurred. As shown in Figure 6.1 for the azimuth monopulse phase ratio [30], reducing the SNR below a specific SNR value does not change the measurement error variance. This behavior is typical for a non-linear estimation problem, but the SNR measurement threshold, or the knee in the curve, is dependent on the system characteristics and varies based on system sensitivity.

To derive the approaches, we separate the mathematical approximation of Figure 6.1 into a two sub-function piecewise equation, one for the low-SNR and one for high-SNR. The high-SNR regime is given in [52], while the low-SNR regime was empirically derived to match the simulated values in Figure 6.1. The piecewise error variance equation for the monopulse azimuth phase

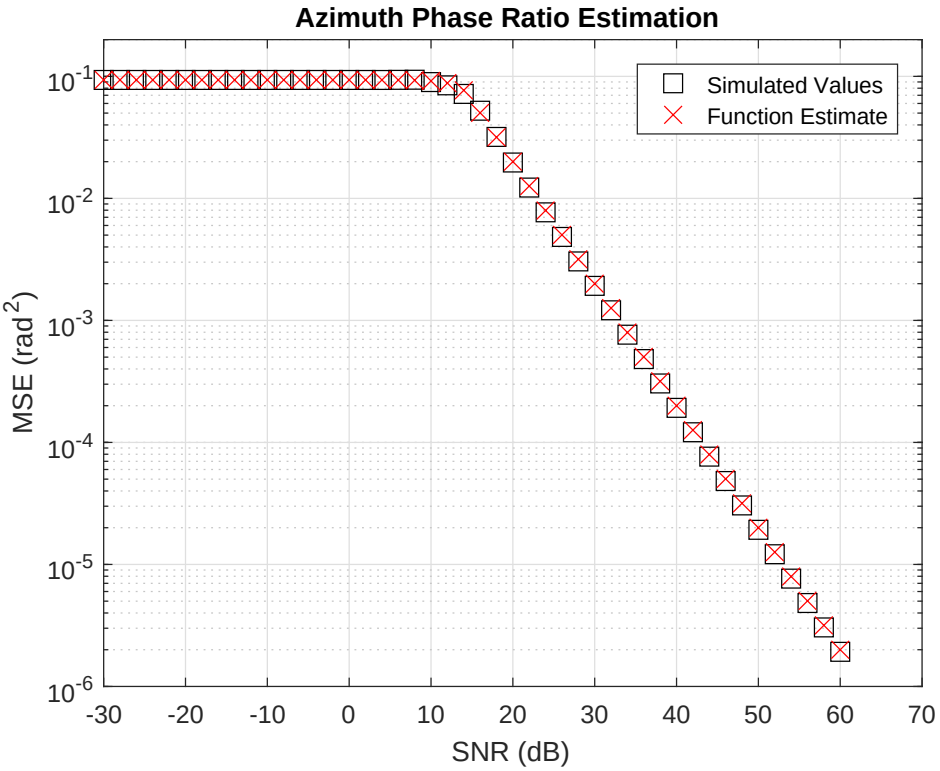

Figure 6.1: Nonlinear measurement error variance for monopulse azimuth phase ratio.

Table 6.1: Aperture Allocation Radar Parameters.

Center Frequency	5 GHz
Bandwidth	50 MHz
Pulse Repetition Frequency	5 kHz
Pulse Width	20 μs
LFM Chirp Rate	2.5×10^{12} $\frac{1}{s^2}$
Blind Range	3 km
Unambiguous Range	30 km
Number Fast Time Pulses	128
Antenna Spacing	$\frac{\lambda}{2}$ m
Number of Antenna Elements	100
Power Per Element	1 W
Noise Power	-205 dB

ratio is

$$\begin{aligned} \epsilon^2_{lowsnr_n} &= c_1\left(c_2^{[10\log_{10}(\mathrm{SNR}_n)+c_3]}\right)+c_4 : \mathrm{SNR}_n < \zeta \\ \epsilon^2_{highsnr_n} &= \frac{8}{a^2_{rx_n}\mathrm{SNR}_n} : \quad \mathrm{SNR}_n > \zeta \end{aligned} \tag{6.2}$$

where a_{rx_n} is the number of aperture elements used for monopulse processing on receive for the n^{th} target, SNR_n is the SNR achieved after pulse compression with time-bandwidth product $\tau_n\beta_n$ and integration over a CPI of M pulses for the n^{th} target, and c_1, c_2, c_3, c_4, and ζ are constants that depend on the system. ζ is a measurement threshold that separates the low-SNR regime and the high-SNR regime in Figure 6.1. The value of ζ is determined experimentally based on the sensitivity of the system. For the azimuth phase ratio monopulse estimation in Figure 6.1, c_1, c_2, c_3, c_4, and ζ are found through simulation to be -0.0031, 1.75, -11, 0.0935, and 15 dB, respectively, for this system. The radar parameters to determine these constants are given in Table 6.1.

Similar to Chapters 4 and 5, the azimuth ratio parameter is assumed to adhere to a Gaussian distribution. Due to this similarity, the azimuth ratio's

belief state is

$$b_{n,k} = p(x_{n,k}|y_{n,k}) \sim \mathcal{N}(\mu_{n,k}, \sigma^2_{n,k}). \tag{6.3}$$

Once again, the parameter's variability increases the belief state's uncertainty between each measurement, and the belief state prior to the next measurement becomes

$$\hat{b}_{n,k+1} \sim \mathcal{N}(\mu_{n,k}, \sigma^2_{n,k} + \delta^2_{n,k}). \tag{6.4}$$

By representing the parameter as a random signal with a belief state, the Bayesian estimation paradigm can be used to recursively estimate the parameter. The model in (6.1) with a Gaussian parameter and a known variance causes the belief state representation after a measurement to be

$$b_{n,k+1} \sim \mathcal{N}\left(\frac{y_{n,k+1}\left(\sigma^2_{n,k} + \delta^2_{n,k}\right) + \mu_{n,k}\epsilon^2_{n,k+1}}{\left(\sigma^2_{n,k} + \delta^2_{n,k}\right) + \epsilon^2_{n,k+1}}, \frac{\left(\sigma^2_{n,k} + \delta^2_{n,k}\right)\epsilon^2_{n,k+1}}{\left(\sigma^2_{n,k} + \delta^2_{n,k}\right) + \epsilon^2_{n,k+1}}\right) \tag{6.5}$$

where $\epsilon^2_{n,k}$ is the measurement error variance. Equation (6.5) has the same form as (4.4) with the fraction $\frac{\nu^2_n}{\rho_{n,\infty}}$ being replaced by ϵ^2_n and with ϵ^2_n represented by a piecewise equation.

For the monopulse estimation application, the processing gain and radar range equation define the SNR that appears in (6.2). Rearranging the radar range equation and including the processing gain variables, the SNR is

$$\text{SNR}_n = \tau_n \beta_n M \frac{a^2_{tx_n} P_e G_e G_{r_n} \lambda^2 \varsigma_n}{(4\pi)^3 R^4_n k T_o N_f \beta_n L_s} \tag{6.6}$$

where a_{tx_n} is the number of transmit digital array elements allocated to the n^{th} target, P_e is the per-element transmit power, G_e is gain of each element, G_{r_n} is the receive gain for the n^{th} target, λ is the operating wavelength, ς_n is the radar cross section (RCS) of the n^{th} target, k is Boltzmann's constant,

T_o is the receiver temperature, N_f is the noise figure, β_n is the bandwidth for the n^{th} channel, and L_s is the system losses. The quadratic dependence on the number of transmit antenna elements is due to the dependence of both the transmit power and transmit gain on the number of transmit elements. Equation (6.6) can be simplified into the number of transmit elements, a_{tx_n}, and the SNR for a single element, SNR_{el_n}, given by

$$\text{SNR}_n = a_{tx_n}^2 \text{SNR}_{el_n}. \tag{6.7}$$

Instead of allocating power in Chapters 4 and 5, each approach in this chapter calculates the number of transmit antenna elements, a_{tx_n}, for each channel. By calculating the number of transmit antenna elements, each approach is able to affect both the antenna gain and power transmitted toward each target. Similar to the power constraint in (3.2), the number of antenna elements in (6.6) is constrained across all targets for allocation by

$$\sum_{n=1}^{N} a_{tx_n} \leq A \tag{6.8}$$

where A is the total number of antenna elements.

6.2 Aperture Allocation Strategies

Using the model in Section 6.1, we develop a short-term objective function and three long-term objective functions to obtain four aperture allocation strategies in this section. The approaches in this section derive solutions for a_{tx_n} to represent the number of transmit antenna elements allocated to each target. The number of transmit elements affects the amount of power and gain allocated to that target, which affects the SNR through (6.7), and ultimately the

estimation performance by (6.2).

6.2.1 Mutual Information Maximization

The first proposed approach maximizes the summation of mutual information over all targets at each time step to reduce the overall system entropy metric. Due to the piecewise nature of (6.2), a system of equations cannot be easily derived to maximize the mutual information. Instead, the approach is derived by replacing $\frac{\nu_n^2}{\rho_{n,\infty}}$ in (4.6) by (6.2). After substituting, the low-SNR and high-SNR objective functions for this approach at the k^{th} time step are

$$f_k = \begin{cases} \sum_{n=1}^{N} \frac{1}{2} \log\left(1 + \frac{\sigma_{n,k}^2 + \delta_n^2}{c_1\left(c_2^{[10\log_{10}(\mathrm{SNR}_n)+c_3]}\right)+c_4}\right) & \mathrm{SNR}_n < \zeta \\ \sum_{n=1}^{N} \frac{1}{2} \log\left(1 + \frac{\left(\sigma_{n,k}^2 + \delta_n^2\right) a_{rx_n}^2 \mathrm{SNR}_n}{8}\right) & \mathrm{SNR}_n > \zeta. \end{cases} \tag{6.9}$$

Although (6.9) is not easily optimized, it can be solved, including the constraint in (6.8), with a numerical solver. A conditional statement can be implemented to select between the high-SNR and low-SNR equation. The piecewise function provides an equation to optimize, but the implementation might produce a local maximum instead of the desired global result by getting stuck on either the low- or high-SNR state. Since the parameter variance depends on SNR, which changes at each time step, the element allocations should also be calculated at each time step, which is contrasted with the approaches in Sections 6.2.2, 6.2.3, and 6.2.4.

The mutual information approach achieves good performance, but with an increased computational complexity due to frequently update the allocation.

6.2.2 MSE Asymptote Minimization

In order to alleviate the complexity burden of the mutual information maximization approach in Section 6.2.1, the asymptotic MSE minimization approach reduces the entropy by only calculating the element allocation once at the beginning of an experiment. Realistically, this approach would require recalculation every time the environment changes. By only calculating the number of elements in the beginning, the number of elements allocated to a target are constant throughout time, and the target's variance will converge to a steady state. This steady state is reached when the information gained from each measurement is equal to the uncertainty increase caused by the parameter's dynamic variance. Mathematically, this case is defined by the prior variance in (6.3) being equal to the posterior variance in (6.5), becoming

$$\sigma_{n,\infty}^2 = \frac{(\sigma_{n,\infty}^2 + \delta_n^2)\epsilon_n^2}{(\sigma_{n,\infty}^2 + \delta_n^2) + \epsilon_n^2}. \tag{6.10}$$

Solving (6.10) for the steady state variance, $\sigma_{n,\infty}^2$, yields

$$\sigma_{n,\infty}^2 = -\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \delta_n^2\epsilon_n^2} \tag{6.11}$$

Substituting (6.2) into (6.11) and summing across all channels results in the objective function for both the high- and low- SNR regimes to be

$$f_k = \begin{cases} \sum_{n=1}^{N}\left[-\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \delta_n^2\left(c_1\left(c_2^{[10\log_{10}(\mathrm{SNR}_n)+c_3]}\right)+c_4\right)}\right] & \mathrm{SNR}_n < \zeta \\ \sum_{n=1}^{N}\left[-\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \delta_n^2\frac{8}{a_{rx_n}^2\mathrm{SNR}_n}}\right] & \mathrm{SNR}_n > \zeta \end{cases} \tag{6.12}$$

Similar to the prior case, a constrained numerical solver can use the constraint in (6.8) and a conditional statement to solve (6.12). The implementation of the

piecewise function in (6.12) has the possibility of a local optimum being calculated instead of a global point due to using the conditional statement. Even though the asymptotic entropy minimization approach requires less computation, the mutual information maximization approach from Section 6.2.1 can sometimes achieve better results.

6.2.3 Parameter Variance Minimization Below a Threshold

In contrast to previous approaches, which minimized the overall system uncertainty, this approach only allocates resources to targets that can be maintained below a specified variance threshold. The targets that can't be maintained below the given threshold are ignored and are not allocated any resources. Given a specific threshold, γ, the objective function is

$$\gamma > \begin{cases} -\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \delta_n^2 \left[c_1 \left(c_2^{[10\log_{10}(\mathrm{SNR}_n)+c_3]}\right) + c_4\right]} & \mathrm{SNR}_n < \zeta \\ -\frac{\delta_n^2}{2} + \sqrt{\left(\frac{\delta_n^2}{2}\right)^2 + \delta_n^2 \frac{8}{a_{rx_n}^2 \mathrm{SNR}_n}} & \mathrm{SNR}_n > \zeta \end{cases}. \quad (6.13)$$

By substituting (6.7) for SNR_n and solving (6.13) for a_{tx_n}, the number of transmit antenna elements for each target is

$$a_{tx_n} > \begin{cases} \sqrt{\frac{10^{\frac{1}{10}\log_{c_2}\left[\frac{\gamma(\gamma+\delta_n^2)}{c_1\delta_n^2} - \frac{c_4}{c_1}\right] - \frac{c_3}{10}}}{\mathrm{SNR}_{el_n}}} & \mathrm{SNR}_n < \zeta \\ \sqrt{\frac{8\delta_n^2}{a_{rx_n}^2 \mathrm{SNR}_{el_n}\gamma(\gamma+\delta_n^2)}} & \mathrm{SNR}_n > \zeta \end{cases}. \quad (6.14)$$

Equation (6.14) calculates one allocation value for the low-SNR case and one value for the high-SNR case. Selecting the value for a_{tx_n} depends on the number of elements required to reach the high-SNR regime, which is calculated

by

$$a_{thres_n} = \sqrt{\frac{\zeta}{\mathrm{SNR}_{el_n}}}. \tag{6.15}$$

If the low-SNR equation of (6.14) is above the threshold number of elements given in (6.15), then the high-SNR equation in (6.14) should be used. If the low-SNR equation of (6.14) does not meet the threshold in (6.15), then the low-SNR equation should be used.

The results of (6.14) are sorted and allocated from lowest to highest number of elements until the remaining number of elements is not enough to satisfy the allocation minimum for any remaining targets. The remaining targets are ignored and receive no elements allocated toward them. Once there are not enough elements remaining to fulfill (6.14) for the next target, the excess elements are allocated using the asymptotic entropy minimization approach given in Section 6.2.2 for the previous targets.

6.2.4 Measurement Above SNR Threshold

Similar to the method in Section 6.2.3, this strategy increases the benefit for a select number of targets instead of allocating resources to achieve a system-wide average metric. Instead of minimizing the variance of each target, this approach maximizes the number of targets with measurements that have sufficient SNR. The number of elements required to achieve a beneficial measurement, meaning the SNR is above the noise floor (i.e. the high-SNR regime), is calculated using (6.15). Since the system might not have enough resources to satisfy the requirement of every target to reach the high-SNR regime, the results of (6.15) are sorted by target and then allocated from least to most until not enough elements remain to ensure the next target's measurement exceeds the noise floor. Any remaining resources, even though they cannot

benefit the leftover targets, are evenly distributed among the targets that will achieve beneficial measurements to increase their SNR further.

6.3 Results

This section compares the performance of each element allocation approach in Section 6.2 to two baseline approaches: 1) distributing the number of elements evenly between all the targets and 2) using the whole aperture to measure the target with the highest variance (AKA the greedy approach). The approaches are compared using both the average estimated MSE and the number of "lost targets" metrics. While a lower average estimated MSE increases the overall estimate of all the targets, the number of "lost targets" metric compares the uncertainty on a target by target basis. The "lost targets" metric specifies a MSE threshold, γ, that is analogous to a tracker losing track of a specific target.

The parameter variances are initialized by solving (6.11) for the high-SNR regime with an even power distribution, unless that variance is above γ. Any starting variances above γ are automatically reduced below γ to guarantee the initial variance does not cause the approaches to immediately lose that target.

The RCS value, range, and dynamic variance of each target are randomly selected from a probability distribution function. Both the RCS value and the range come from a uniform distribution with the range between 3km and 30km and the RCS value between 2 and 40 m^2. The dynamic variance, on the other hand, uses a Rayleigh distribution to emphasize less dynamic, or lower accelerating, targets with a scale parameter of 0.007. A scale parameter of this magnitude corresponds to a average acceleration of 2.82 g's. The simulations in this section use the same system parameters as given in Table 6.1.

Table 6.2: Average estimated MSE and percentage of lost targets comparison averaged across varying target amounts.

	Single Realization	Monte Carlo Simulations	
Approach	Average MSE	Average MSE per Target	Percentage of Lost Targets
Greedy	0.011	0.23×10^{-3}	42.3%
Even Distribution	0.008	0.23×10^{-3}	41.1%
MSE Asymptote	0.006	0.21×10^{-3}	35.8%
Mutual Information	0.008	0.20×10^{-3}	36.4%
Track Threshold	0.006	0.22×10^{-3}	31.4%
Measure Threshold	0.006	0.21×10^{-3}	32.4%

This section compares the performance of each approach for a single realization of five targets using the average MSE metric in Section 6.3.1. The single realization provides a view of the system's performance over time for a specific scenario. Although each realization will have differing results, the qualitative analysis of a single scenario provides useful insights. Then, to account for differing scenarios, 1,000 different realizations are averaged together in Sections 6.3.2 and 6.3.3 to compare the approaches using the average MSE and lost targets metric, respectively, for varying numbers of targets.

6.3.1 Single Realization

Simulating targets for extreme values of SNR and dynamic variance explores the usefulness of each approach for large variations of target characteristics. To this end, the single realization in this section included five targets with large variations: one at low-dynamics (0.01 g's) and low-SNR (range of 30 km), one at high-dynamics (9 g's) and high-SNR (range of 3 km), one at low-dynamics and high-SNR, one at high-dynamics and low-SNR, and one at medium-dynamics (4.5 g's) and medium-SNR (range of 13.5 km). The average

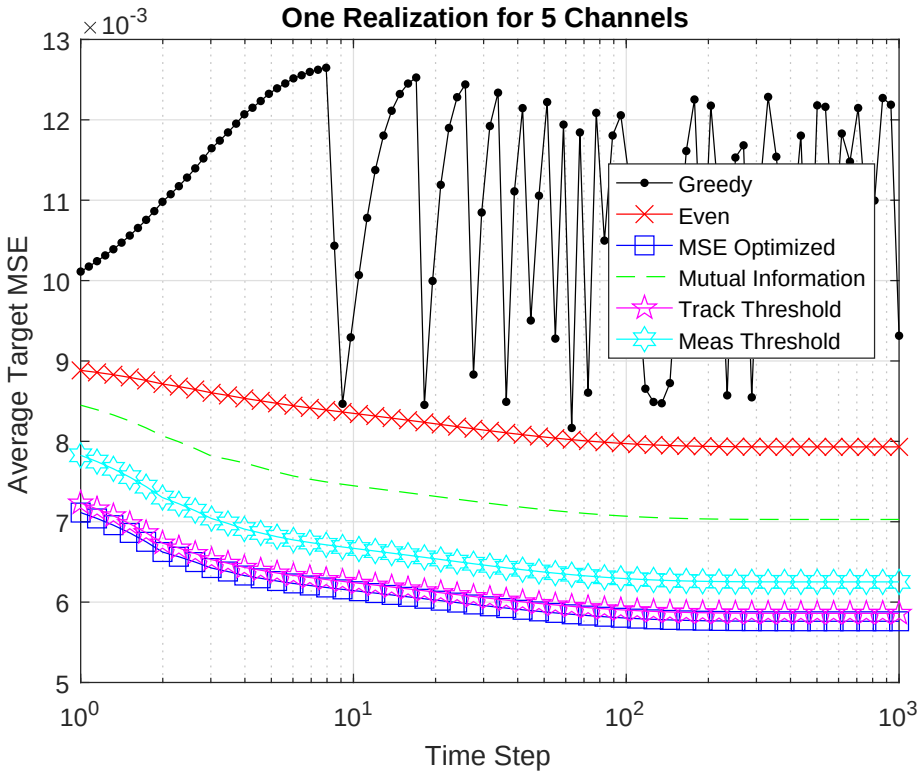

Figure 6.2: Average estimated MSE comparison for a single realization.

estimated MSE for these five targets are shown in Figure 6.2 versus log time. Figure 6.2 shows that each approach, except the greedy approach, reaches a steady state MSE value. The greedy approach does not reach a steady state because it constantly switches the allocation between two targets and never settles to a constant allocation. The values in the second column of Table 6.2 average the MSE over both time and targets to quantify each approach.

By allocating most of the resources to the medium-dynamics and medium-SNR target, the asymptotic MSE and tracking threshold approaches have the best results. In this scenario, they have extremely similar performance and almost overlap in Figure 6.2. The measurement threshold approach is close to the performance of the asymptotic MSE and tracking threshold approaches because it also allocates a little over the majority of resources to the medium-

dynamics and medium-SNR target. Although the approach still focuses on the medium level target, it spreads the resources more evenly between the two high-SNR targets and the medium target. The mutual information maximization strategy focuses the majority of the aperture on the high-SNR targets and overlaps in performance with the even distribution approach. Finally the greedy approach performs the worst because it only allocates energy to the medium and high dynamics targets.

6.3.2 Monte Carlo Average Mean Squared Error

The realization in Section 6.3.1 displays the results for a single scenario but does not provide quantitative performance over many different scenarios. Instead, Monte Carlo simulations are useful to compare the performance across many different scenarios. Figure 6.3 compares the results for each approach by averaging the MSE across time and targets, and dividing by the number of targets for 1,000 realizations. Each Monte Carlo simulation is performed for a varying number of targets between 10 and 100. The values in the third column of Table 6.2 quantify these results by averaging across the Monte Carlo simulations.

Comparing the results shows that the mutual information approach has the best average MSE but also requires the highest amount of computation. The MSE asymptote and measurement threshold approaches perform extremely similar, and are a close second best. The MSE asymptote approach is not the best, which is attributed to the possibility of reaching a local maximum instead of a global result. The tracking threshold approach rounds out the proposed approaches in fourth with the baseline approaches, greedy and even, performing the worst.

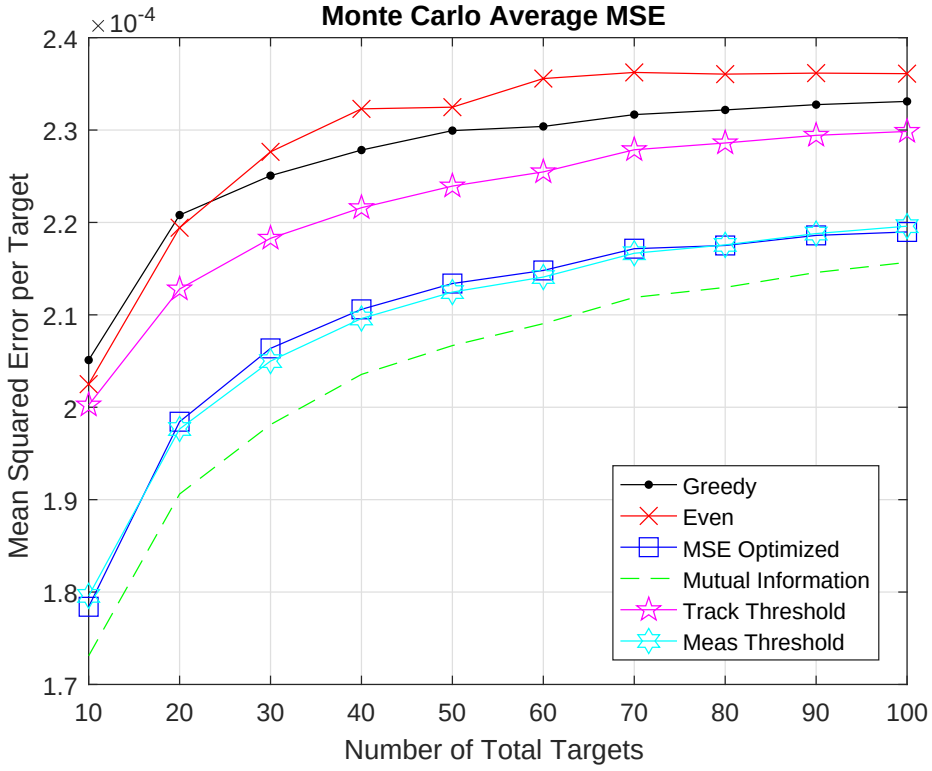

Figure 6.3: Average estimated MSE comparison for a varying number of targets.

6.3.3 Lost Targets

In contrast to previous sections, which compares the approaches using the MSE metric, this section compares the results of Monte Carlo simulations using the lost targets metric. The lost targets metric is analogous to the error becoming so high that a tracker is unable to reliably provide an estimate of the actual target location. This threshold depends on the system, but -16 dB was used for the lost target error threshold in this simulation. The Monte Carlo simulations have the same characteristics as in Section 6.3.2, only with a different metric for comparison. If the variance of a target goes above the threshold variance threshold at any point during the simulation, then the target is considered lost for the entire simulation. The average number of lost targets across 1,000 realization is shown in Figure 6.4 for a varying number of targets between 10 and 100. The fourth column of Table 6.2 also quantifies the percentage of lost targets averaged across the Monte Carlo simulations.

Unsurprisingly, the tracking threshold approach performs the best because it is actively decreasing the variance of as many targets as possible below the variance threshold. The measurement threshold approach is a close second in performance and actually performs the same as the tracking threshold approach in the highly resource constrained scenarios, when the system is tracking 100 targets with 100 elements to allocate. The MSE asymptote and mutual information maximization approaches are third and fourth, respectively, even though the mutual information approach overtakes the MSE asymptote approach in the resource constrained scenarios. Once again, the baseline approaches, the greedy and even distribution, perform the worst for the lost targets metric.

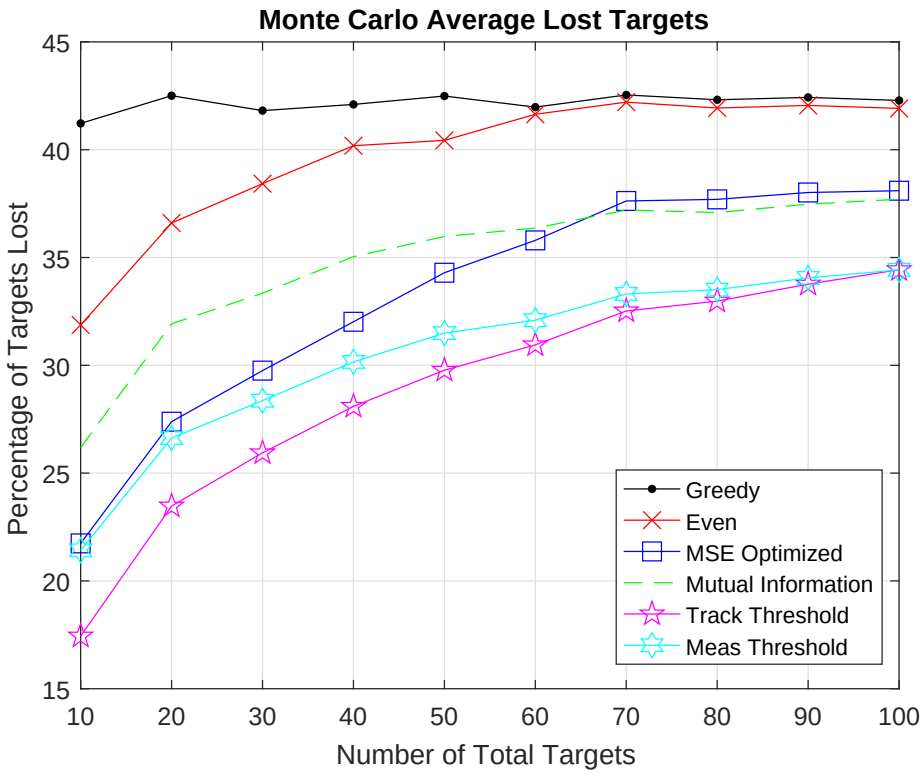

Figure 6.4: Lost targets comparison for a varying number of targets.

Chapter 7

Tracking Radar Simulation

7.1 Overview

As described in Chapter 2, trackers are able to increase the estimation accuracy of a fluctuating parameter value by combining multiple measurements. Previous chapters represented the recursive estimation using a belief state prior distribution and posterior equation. Instead of representing the estimation using the belief state posterior equation, many systems implement a Kalman filter to recursively estimate the parameter value. Along with the recursive estimation in (2.4), the Kalman filter also recursively calculates the covariance using (2.6), which is related to (5.3) by the Woodbury matrix identity [53].

Although many implementations of a Kalman filter exist, this chapter uses a constant velocity implementation, due to the fact that the velocity is considered constant in (2.12), described in Chapter 2. Although the constant velocity tracker is the simplest model, it provides the necessary constructs to evaluate the proposed allocation strategies and analyze the motivation behind each strategy for multiple scenarios. An example of the tracker's performance in two dimensions is shown in Figure 7.1. The initial estimate, represented by a dotted line, on the right side of Figure 7.1 is far off from the actual target lo-

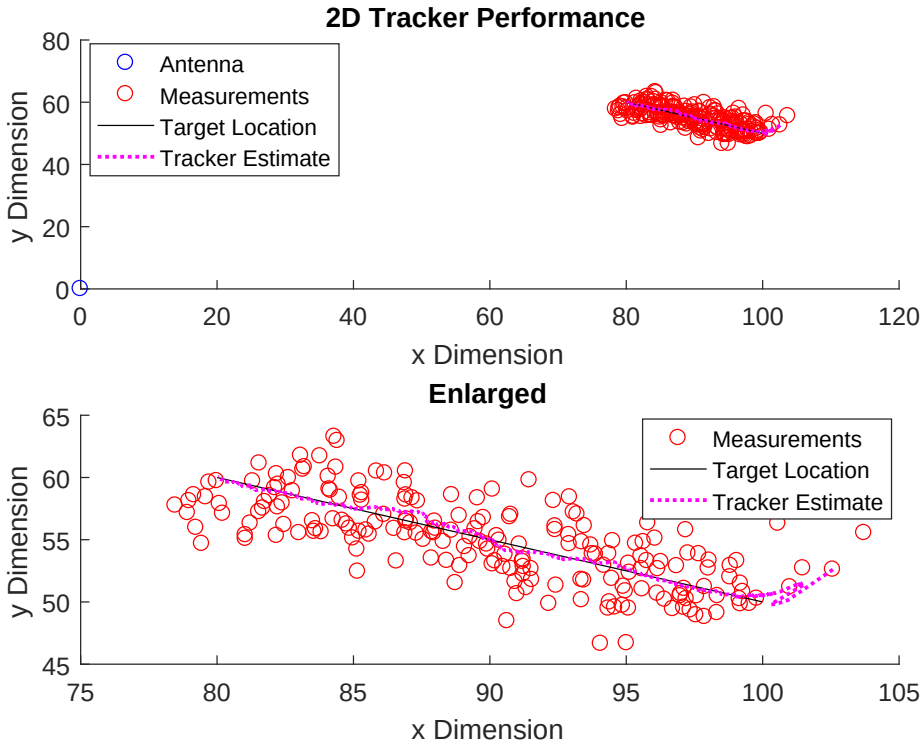

Figure 7.1: Tracker performance in two dimensions.

cation due to the initial measurement error, but the tracker's estimate quickly converges to the actual location. The error included in each measurement is transferred, but considerably reduced, to the tracker's estimate. By combining the measurements together, the tracker is able to reduce the effects of the measurement error, but not completely remove it causing the estimate to still slightly vary around the target's trajectory.

Along with the recursive Kalman filter estimator, the measurement model in Chapter 6 is extended to three dimensions to now include both range and elevation for the radar measurements. When the received measurement is located below the noise floor, the system receives no information from the measurement and must guess the parameter's state based on the possible values. Since range has a larger number of possible values and a larger resolution,

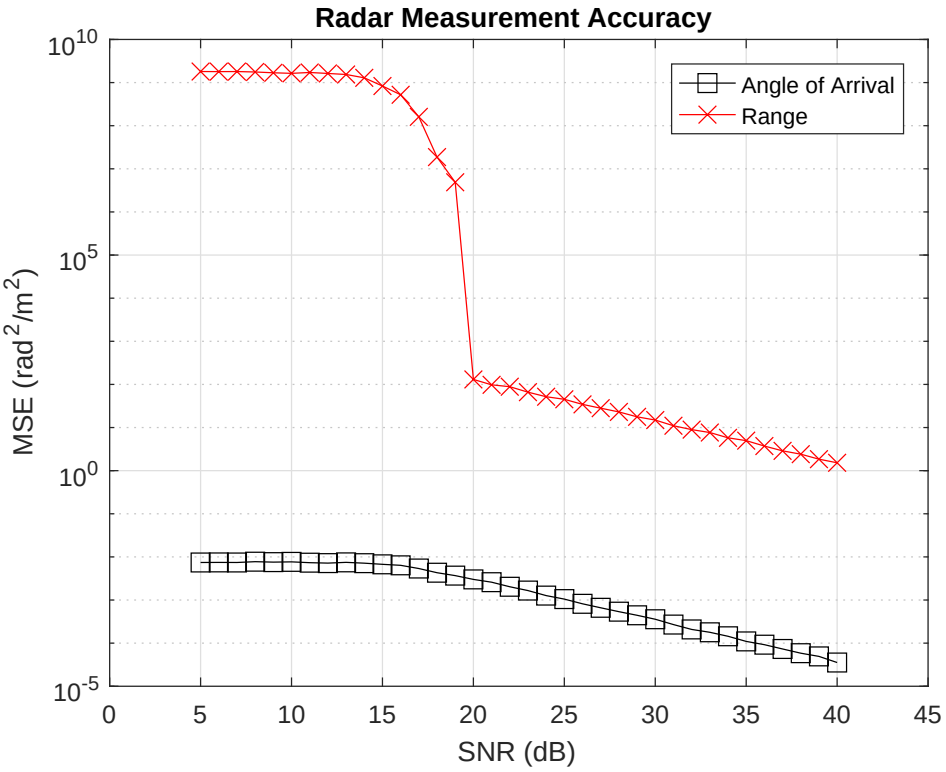

Figure 7.2: Radar range and angle of arrival measurement accuracy.

without accounting for the difference in units, than angle of arrival, the uncertainty of the range measurement is much higher than the error in the angle of arrival, as shown in Figure 7.2.

The inclusion of both range and elevation angle to the measurement model from Chapter 6 requires additional estimates to be calculated. Since the elevation angle estimate is determined through the same monopulse technique as the azimuth angle, it has the same accuracy plot in Figure 6.1 and equation in (6.2). The calculation of range is vastly different than the monopulse technique and requires a new high-SNR equation to characterize that estimate. Since the low-SNR equation in (6.2) is generic, only the constant values are necessary to change. For range, the constants c_1, c_2, c_3, c_4, and ζ were found through simulation to be -1, 1.25, 75.8, 10^9.2, and 18.5, respectively, for this

Table 7.1: Aperture Allocation Radar Parameters.

Center Frequency	50 MHz
Bandwidth	1 MHz
Pulse Repetition Frequency	1 kHz
Pulse Width	10 μs
LFM Chirp Rate	1×10^{11} $\frac{1}{s^2}$
Blind Range	1.5 km
Unambiguous Range	150 km
Number Fast Time Pulses	128
Antenna Spacing	$\frac{\lambda}{2}$ m
Number of Antenna Elements	100
Power Per Element	1 W
Noise Power	-205 dB

system. The properties of this system are shown in Table 7.1. The range high-SNR equation, from [54], for a linear frequency modulated waveform is

$$\epsilon^2_{highsnr_n} = \frac{3c^2}{8\pi^2 \mathrm{SNR}_n \beta^2} : \quad \mathrm{SNR}_n > \zeta, \tag{7.1}$$

where c is the speed of light.

By combining the realistic Kalman filter estimator with the realistic measurement model in Chapter 6, this chapter applies and analyzes the resource allocation approaches from Chapter 6 to explain their operation for multiple realistic scenarios.

7.2 Results

The results in this chapter are based on multiple scenarios designed to increase the understanding of each aperture allocation strategy from Chapter 6. The greedy and even distribution approaches are provided as baseline comparisons and will not be analyzed. Beyond demonstrating each allocation strategy using

specific scenarios, this chapter also investigates the effect of varying both the number of targets and the time between updating the resource allocation for each strategy by implementing Monte Carlo simulations for each. During the specific scenarios, the approaches recalculate the aperture allocation every 10^{th} CPI. To ease the readability of the scenario plots, the data is down-sampled to a data point every 100 CPIs.

Along with the average tracker MSE and lost targets metrics, we include a deleted tracks metric to compare the amount of resources wasted by the system tracking a target that is ultimately deleted. A tracker deletes a track when a useful measurement is not received for a specific number of time steps, or CPIs. In this case, a threshold of four CPIs is used to determine whether a track is deleted. For the tracker estimation model, the lost targets metric required a MSE of 1×10^4 before considering a target to be lost. The lost targets metric is calculated using the Euclidean distance in Cartesian coordinates for the tracker.

The radar properties for every simulation in this section are given in Table 7.1. In the Monte Carlo simulations, the targets' initial ranges are randomly selected from a uniform distribution between the two values. The targets' initial velocities are selected from a Rayleigh distribution based on a maximum of 300 m/s with a binary random variable, either a 1 or -1, to determine the direction. Targets' accelerations are determined by a Gaussian distribution based on the dynamic variance. The Monte Carlo simulation also varies the targets' RCS values based on a uniform distribution from 2 to 40 m^2.

7.2.1 Two Target Scenario #1

The first scenario is split into three related sub-scenarios, named 1A, 1B, and 1C. These scenarios are related because they all contain two targets with the same RCS of 20 m^2: one moving around the radar at a constant range and the other target starting close to and moving away from the radar. The different variants change the initial range of the first target and the dynamic variances for both targets. The dynamic variances determine the amount of typical accelerations based on a Gaussian distribution.

Scenario 1A, shown in Figure 7.3, includes two targets: one at a constant range of 50,990 m and another target starting at a range of 2,263 m, with each target having an acceleration variance of 7.4 g's. Since only two targets exist, the number of elements allocated to target #1 is analyzed, while the remaining elements are allocated to target #2. The allocation for target #1 is shown in Figure 7.4.

Initially, all of the approaches start with over 95% of the elements allocated to target #1 due to close proximity of target #2 with the radar. Since target #2 is close to the radar, it does not need many resources to achieve a good estimate. In contrast, target #1 needs more resources to increase the SNR enough to measure its location. As target #2 starts moving away from the radar, the mutual information maximization approach is the first to start allocating more resources to target #2 and reaches an even distribution around the 230^{th} CPI, when the second target is at a range of 14,950 m. In contrast, the other approaches do not reach an even distribution until CPI 900, when the second target is at 49,700 m.

As the second target moves away from the first target, the mutual information approach continues to allocate evenly until CPI 1,520 while the other

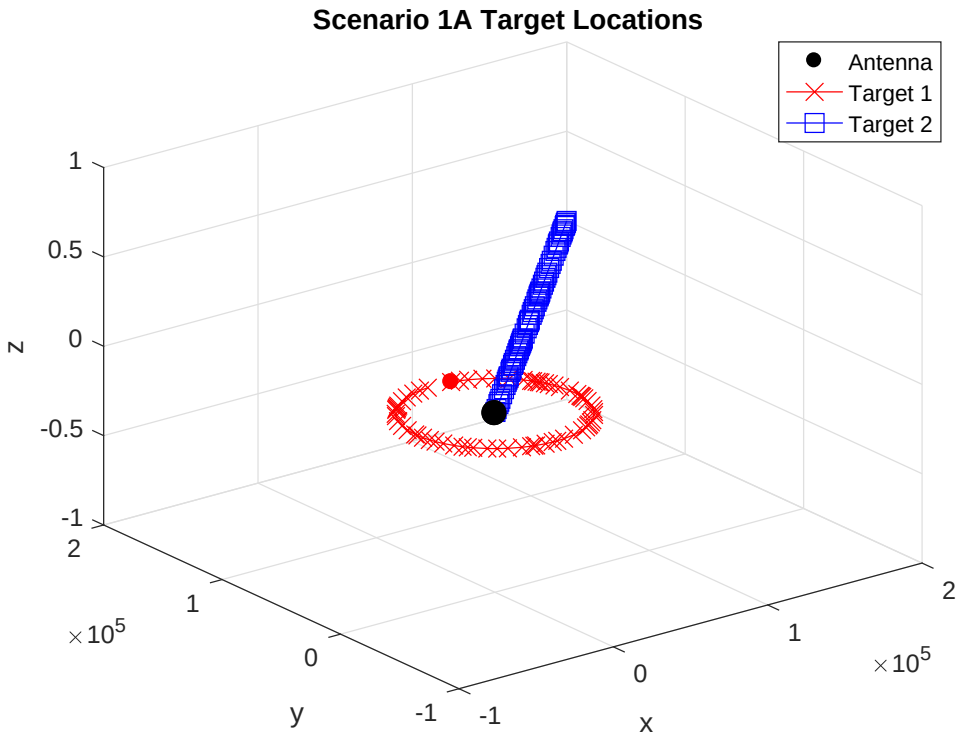

Figure 7.3: Scenario 1A trajectory.

approaches progress much quicker to 26 elements allocated to target #1. In this scenario, 26 elements satisfies (6.15) for target #1. In other words, when target #1 receives less than 26 elements, then the target's SNR is too low to generate a beneficial measurement. All of the approaches reach 26 elements around CPI 1850, except for the measurement threshold approach, which deviates and allocates all the resources to target #1.

At CPI 1850, target #2 is 87,200 meters away from the antenna and the measurement threshold approach has reached a decision point. Based on the measurement model, the minimum number of elements to achieve a useful measurement is available to either target, but not both. Therefore, the measurement threshold approach must either continue to allocate resources to target #2 and ignore target #1, or vice versa. The approach chooses target #1 and allocates all the resources toward it while the other approaches continue to split the resources between both targets. Toward the end of the scenario, around CPI 8950 and a range of 406,000 m for target #2, the mutual information maximization approach calculates that target #2's SNR is too low for any information to be gained and also allocates all the resources to target #1.

Scenario 1B is similar to 1A, with all the same parameters except that target #1 is closer to the radar at 25,500 m. The resource allocation of each strategy in scenario 1B for the first target is shown in Figure 7.6. Compared to scenario 1A, the resource allocation approaches for this scenario start out allocating fewer elements to target #1. The mutual information maximization approach also waits until the 320^{th} CPI, when the second target is at 18,854 m, to evenly distribute the resources. The other approaches reach the even point at about the same time, around CPI 430 or when the second target is

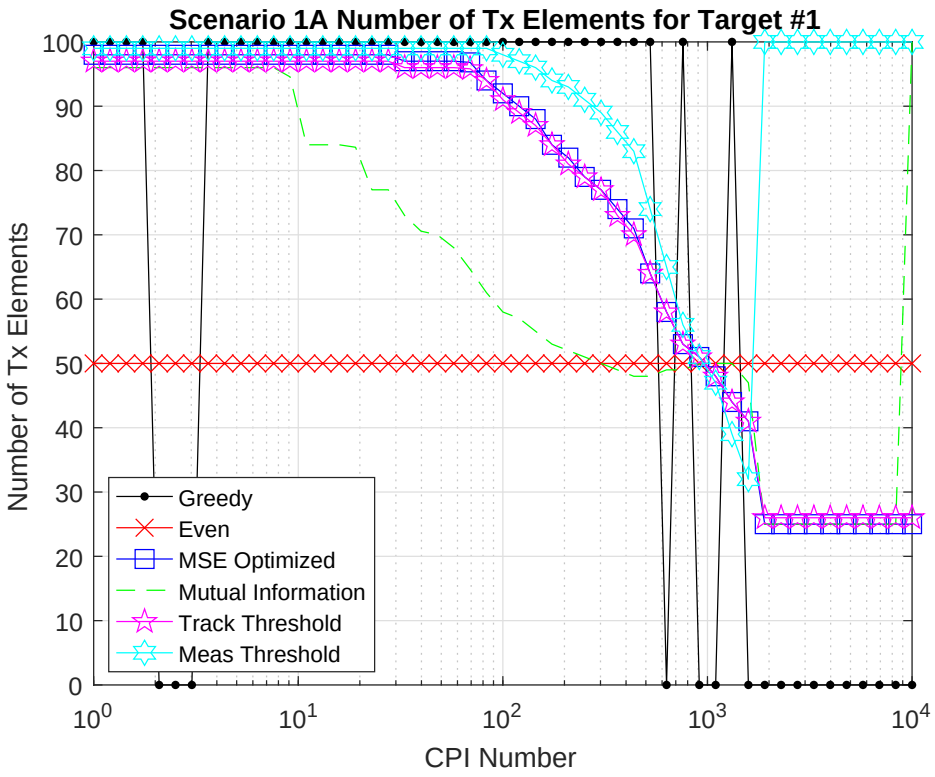

Figure 7.4: Scenario 1A resource allocation for target #1.

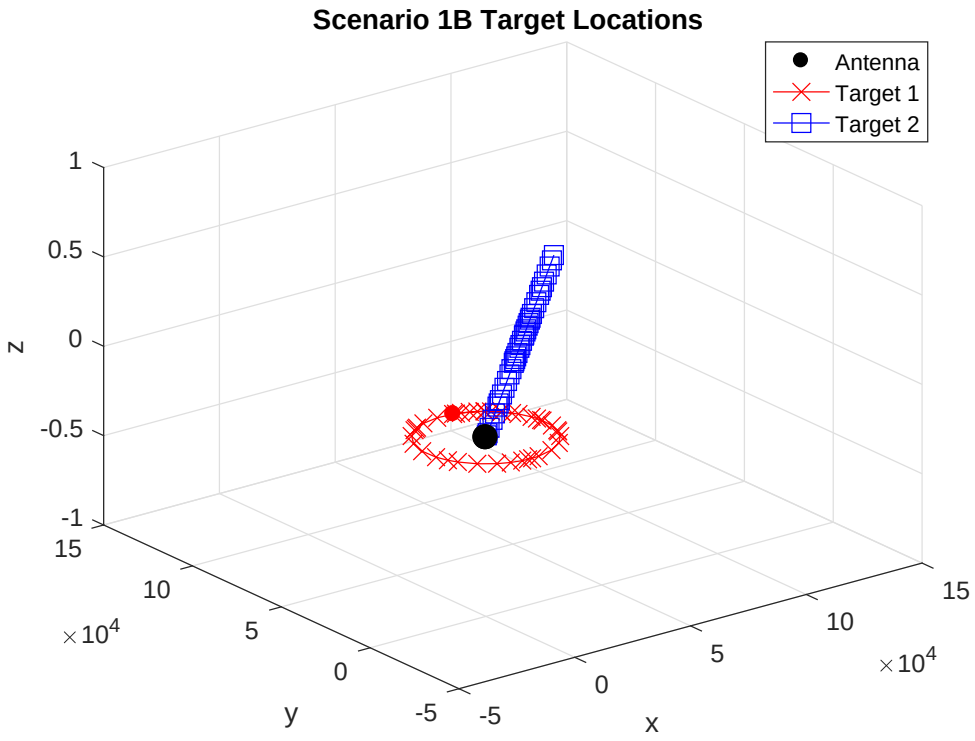

Figure 7.5: Scenario 1B trajectory.

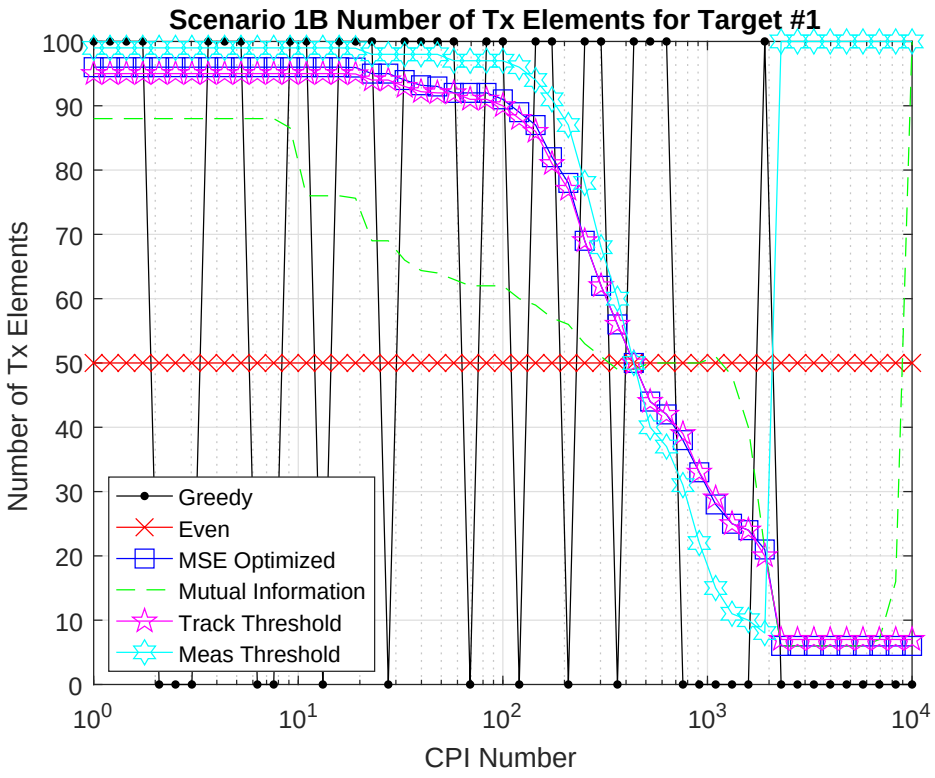

Figure 7.6: Scenario 1B resource allocation for target #1.

25,500 m away from the radar.

Since target #1 is closer than in scenario 1A, the target requires fewer elements to achieve a useful measurement. In this case, only seven elements are required to gain a beneficial measurement of target #1. The measurement threshold approach reaches this point at CPI 2050, which is later than in scenario 1A because target #2 must get far enough away to decrease its SNR enough to require greater than 93 elements for a useful measurement. Since target #1 is closer, the mutual information maximization approach also allocates all the resources to target #1 earlier than scenario 1A because more information can be gained from the first target when it is closer.

Scenario 1C has the same distances as scenario 1B but decreased the acceleration variance from 7.4 g's to 2.7 g's. Compared to scenario 1B, the

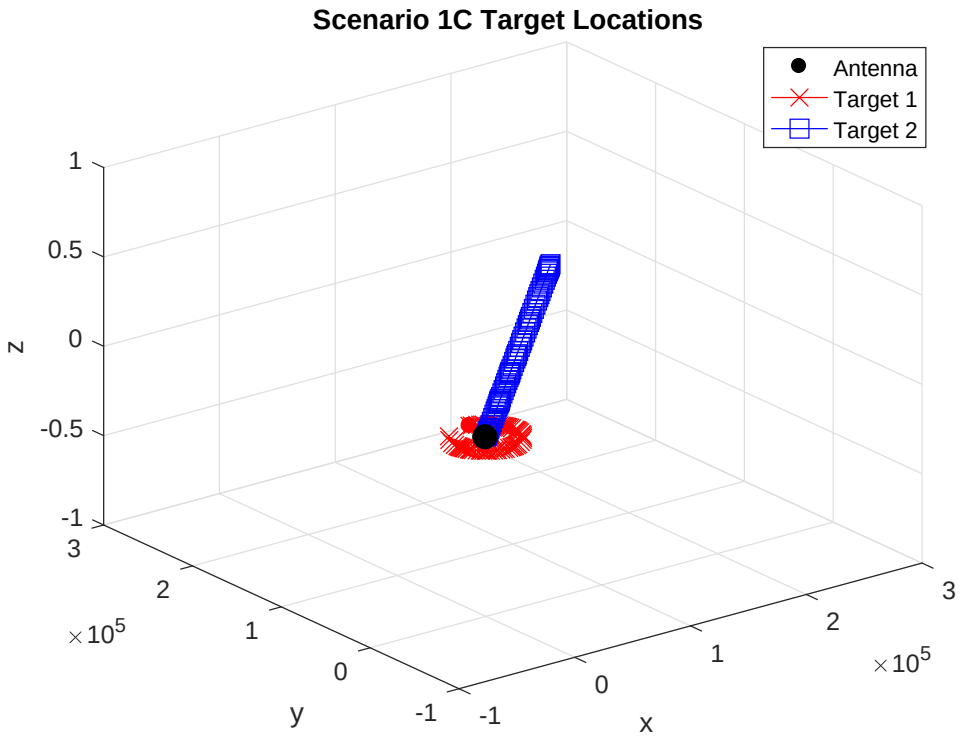

Figure 7.7: Scenario 1C trajectory.

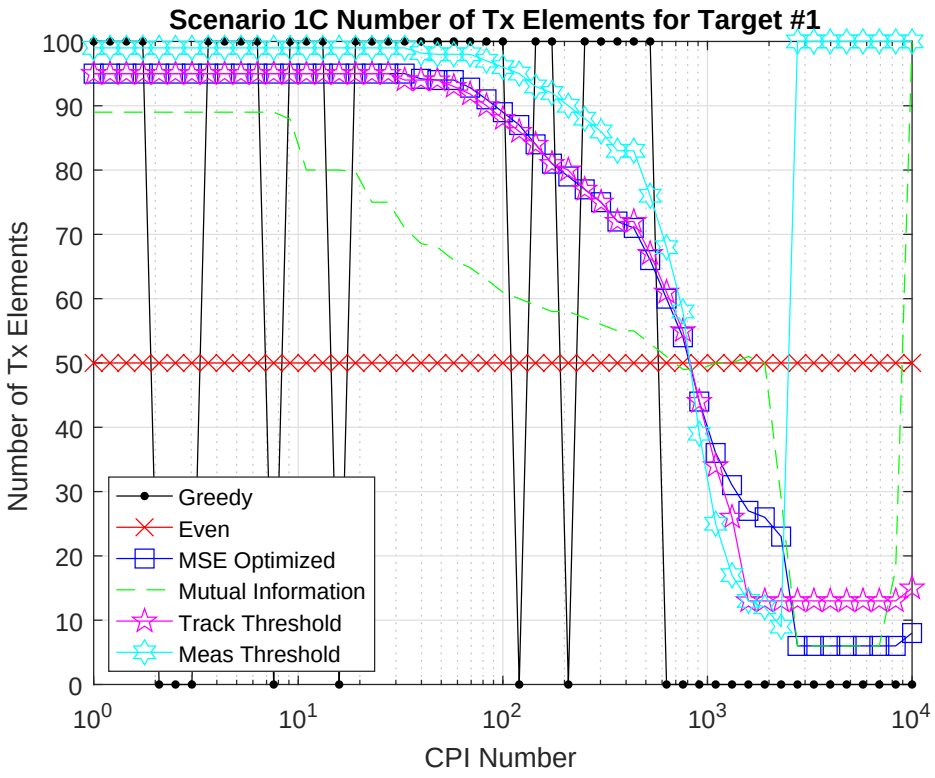

Figure 7.8: Scenario 1C resource allocation for target #1.

allocation for this scenario starts at the same amount and takes longer to reach the even distribution point at CPI 830, due to the slow acceleration of target #2. The CPI at which the measurement threshold approach reaches the decision is also later at CPI 2470. The most notable difference due to the reduction in dynamic variance is for the tracking threshold approach, which stops allocating more resources toward the second target before it reaches the minimum number for a useful measurement of target #1, which is seven elements. Also toward the end of this scenario's timeline, the tracking threshold and asymptotic MSE optimized approaches start to allocate more resources back to target #1 due to the reduction in dynamic variance.

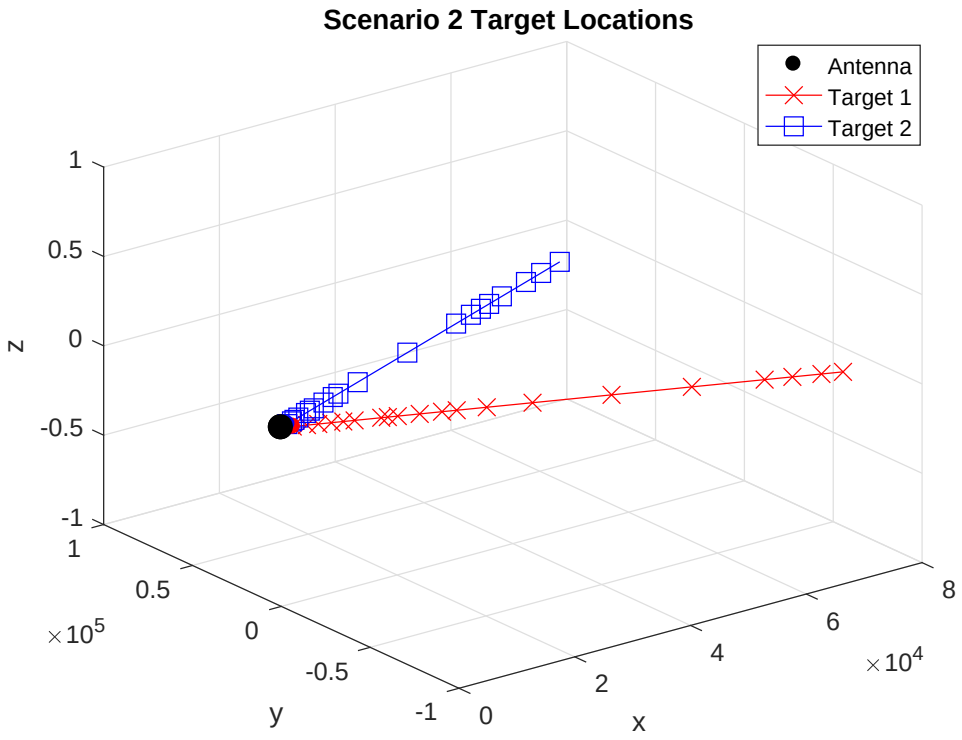

Figure 7.9: Scenario 2 trajectory.

7.2.2 Two Target Scenario #2

The second scenario replaces the constant range target from scenario #1 with another target that starts close to the radar and moves away from the radar. Therefore, this scenario, shown in Figure 7.9, simulates two targets that start close to the radar, at 2,121 m, and move away with the same acceleration variance of 7.4 g's. Once again the RCS values of both targets is 20 m^2. Although their acceleration variances are the same, the actual velocity and acceleration at a specific time are different between the targets. Figure 7.10 shows the difference in range over time due to the difference in velocities and accelerations.

The corresponding resource allocation for each method is shown in Fig-

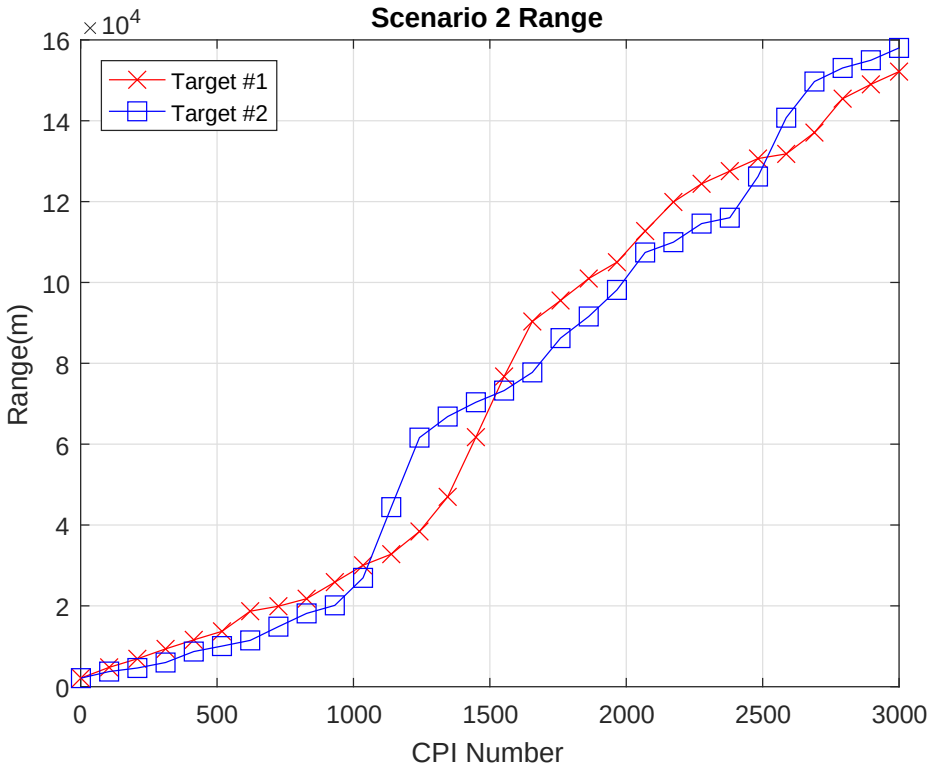

Figure 7.10: Scenario 2 range.

ure 7.11. Since the targets are initially extremely close to each other, the approaches split the resources evenly between the targets. Quickly, the first target accelerates and the approaches respond by allocating more resources to it. The approaches focus on the further target because the nearby target is close enough to gain enough information with a small amount of resources. During this time, the mutual information maximization approach only deviates by four elements from the even distribution approach.

Around CPI 1030, the targets switch relative position from the radar, with the first target closer and the second target further away. To account for this change, the approaches transition to favoring the second target with more resources. The targets cross again at CPI 1475, and the approaches start to switch back as well, until CPI 1560, when target #1 is 85,600 m and target #2 is 75,000 m away. At this point, the targets' SNRs are low enough that getting worthwhile information from both targets is difficult. Therefore, the measurement threshold approach immediately allocates all the resources to the closer target. The other approaches are slower to transition entirely to the closer target but finally reach that point around CPI 1940, when the targets are at 108,610 and 101,290 meters, respectively.

Finally around CPI 2200, at distances of 124,380 and 114,530 meters, respectively, the measurement threshold approach determines both targets to be too far away for beneficial measurements so it evenly distributes the resources between both targets. The remaining approaches continue to allocate all the resources toward the closest target, even when the targets alternate relative locations.

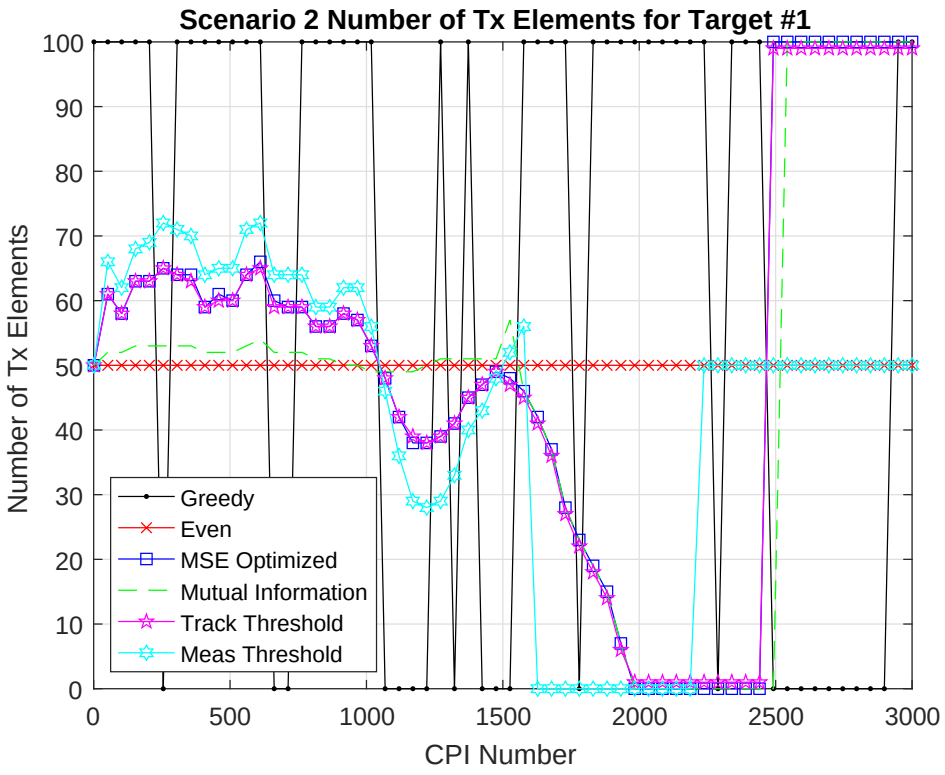

Figure 7.11: Scenario 2 resource allocation for target #1.

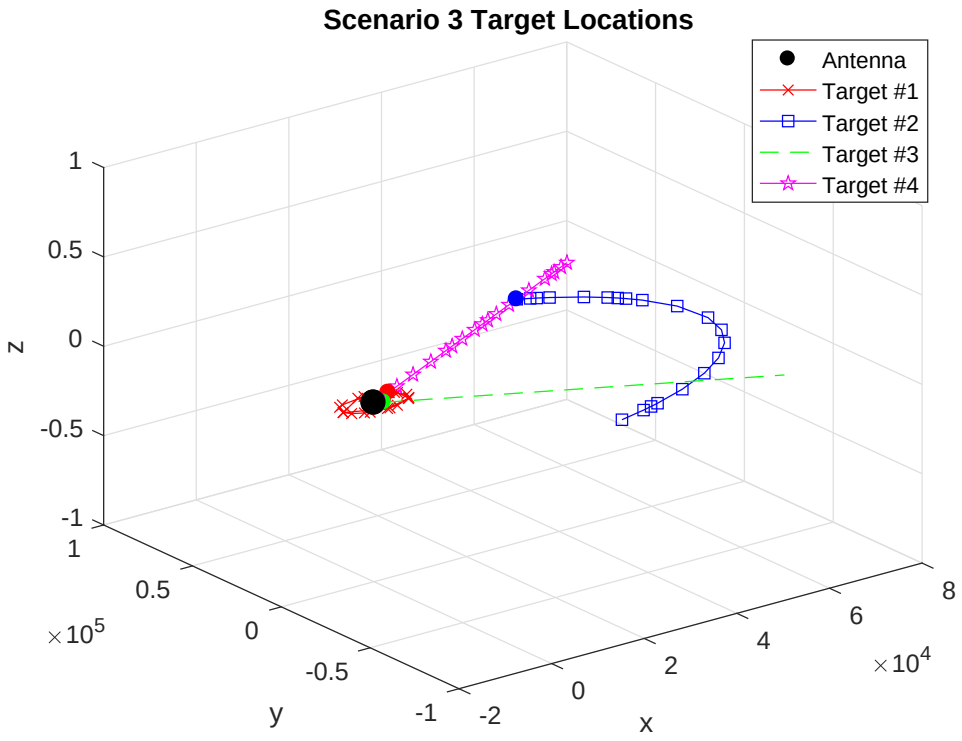

Figure 7.12: Scenario 3 trajectory.

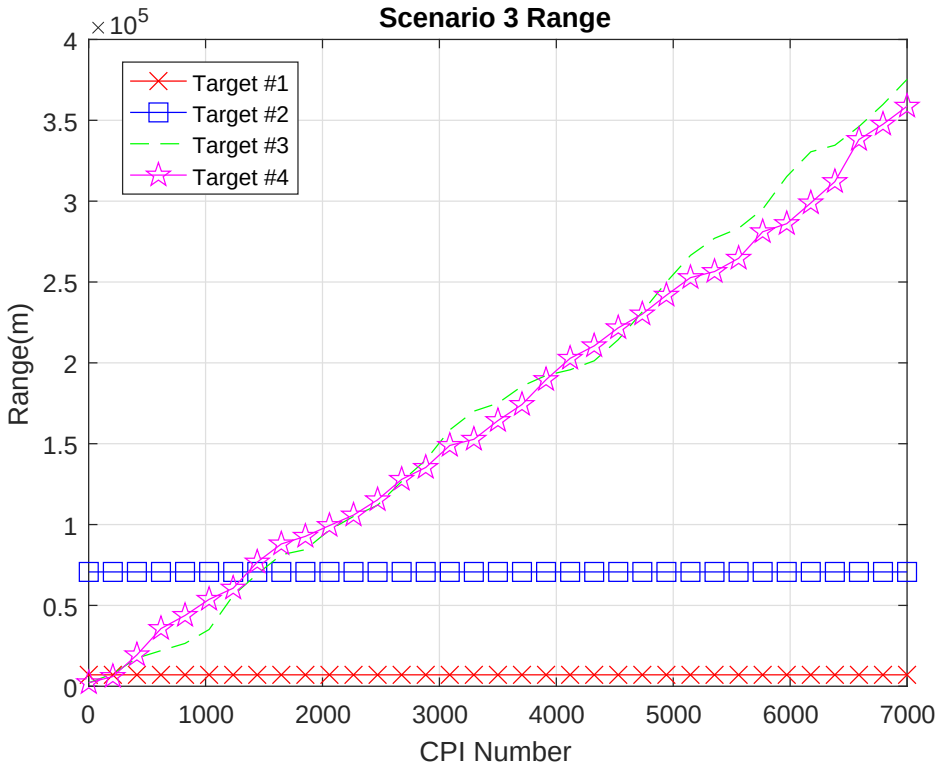

Figure 7.13: Scenario 3 range.

7.2.3 Four Target Scenario

The third and final scenario compares the allocations for more than two targets by combining the first two scenarios. In this scenario, four targets are tracked: the first two at a constant, but different, range, and the last two starting at the same range of 2,121 m, but traveling away from the radar at different speeds, but the same acceleration variance. The two constant range targets, targets #1 and two, are at 7,071 and 70,711 meters, respectively. All four targets have acceleration variances of 7.4 g's and RCS values of 20 m^2. The ranges of each target throughout the scenario are shown in Figure 7.13. Although the simulation lasted for 10,000 CPIs, some plots are reduced in scope to increase the visibility of critical points in the graph.

As seen in Figure 7.14, the measurement threshold initially favors the second target with over 95% of the resources. Due to the vicinity of targets #2, #3, and #4 to the radar, those targets require only a minimal amount of resources to achieve a useful measurement. As targets #3 and #4 move away from the radar, they require more resources to be effectively measured, so the approach starts decreasing the amount allocated to target #1 and allocating them to targets #3 and #4. Since target #4 is further away, it receives more resources to produce a useful measurement compared to target #3. This continues until CPI 1240, when targets #3 and #4 reach ranges of 63,845 and 61,625 meters. At this point, the radar does not have enough resources to measure all three targets sufficiently, and it must pick two of the three targets to measure. This method ends up splitting the resources between targets #3 and #4 because they require less resources to achieve a better SNR than target #2.

As target #4 continues moving, it finally gets further away than target #2, so the radar stops allocating resources to target #4 and only focuses on targets #2 and #3. When target #3 gets to 88,364 meters away, the radar does not have enough resources to maintain tracks for both targets #2 and #3. Therefore, the approach allocates all the resources to target #2 due to it being closer. Throughout the scenario, target #1 was constantly allocated one element because it is close enough to the radar that one element is able to achieve a useful measurement. If this target was further away, that constant allocation would be increased along with the distance until it receives resources similar to target #2.

Similar to the measurement threshold approach, the mutual information maximization approach initially allocates over 90% of the resources to target

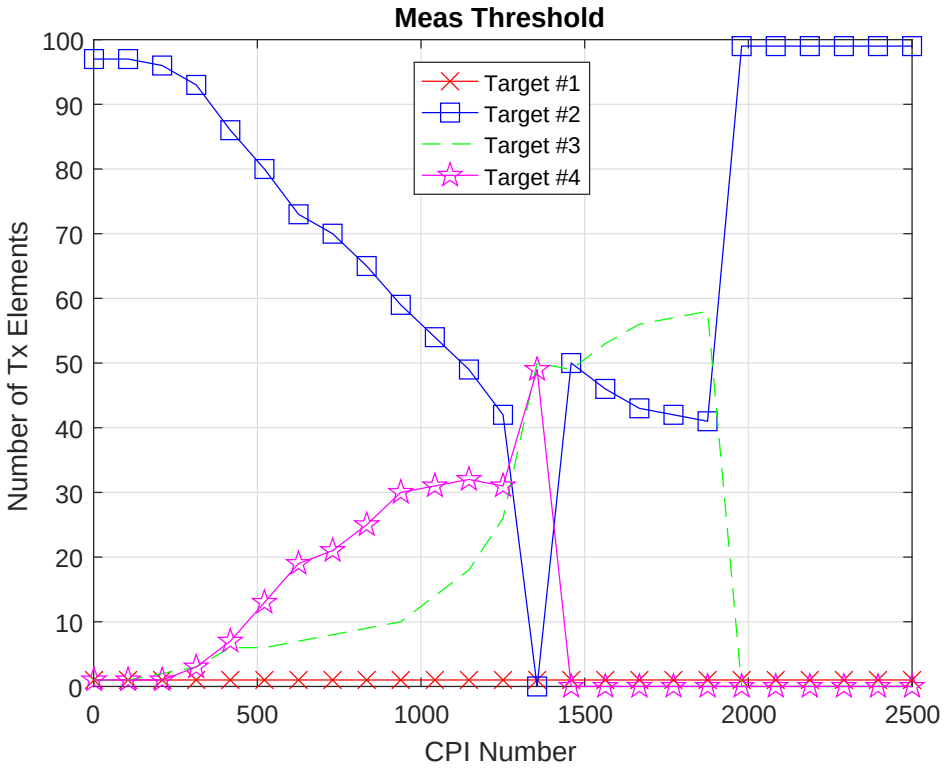

Figure 7.14: Scenario 3 resource allocation for target #1 using the measurement threshold approach.

#2. Almost immediately, though, the MI maximization approach reduces that allocation to half and allocates the remaining half to the other three targets, with target #1 getting the second most resources. As targets #3 and #4 get in the vicinity of target #1 (between CPIs 150 and 300), target #2 still gets about 50% of the resources, while the remaining resources are split evenly between the other three targets. While targets #3 and #4 have ranges between targets #1 and two (between CPIs 300 and 1200), all four targets receive some resources, although target #1 typically receives the least, with the majority of resources being determined between targets #1, #3, and #4 by their current uncertainty.

After CPI 1200, when targets #3 and #4 are in the same vicinity as target #2, the majority of resources is split between targets #3 and #4 until target #4 becomes further away than target #2 at CPI 1350. After CPI 1350, the resources are split between targets #2 and #3 while targets #1 and #4 receive a minimal amount. As targets #3 and #4 continue to get further away, the MI maximization approach continues to decrease the amount of resources allocated to those targets and increases the amount allocated to target #1 until targets #3 and #4 reach the point where no more information can be gained through measurements. The higher percentage of allocation split between targets #3 and #4 switch around CPI 2,800 and again around CPI 3,800 due to switching which one is the closer target. When targets #3 and #4 become too far away, the approach settles on allocating 57% of the resources to target #2 and 43% to target #1 for the remaining time.

Similar to the measurement threshold and MI maximization approaches, the optimized asymptotic MSE approach initially allocates over 90% of the resources toward target #2. As targets #3 and #4 move away from the radar

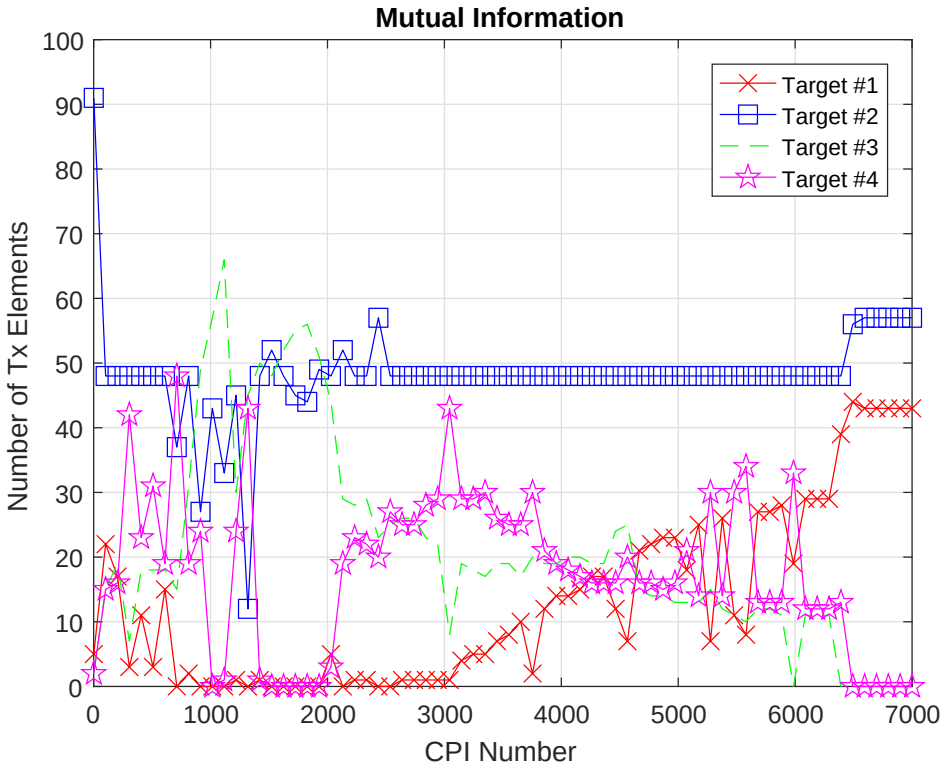

Figure 7.15: Scenario 3 resource allocation for target #1 using the mutual information maximization approach.

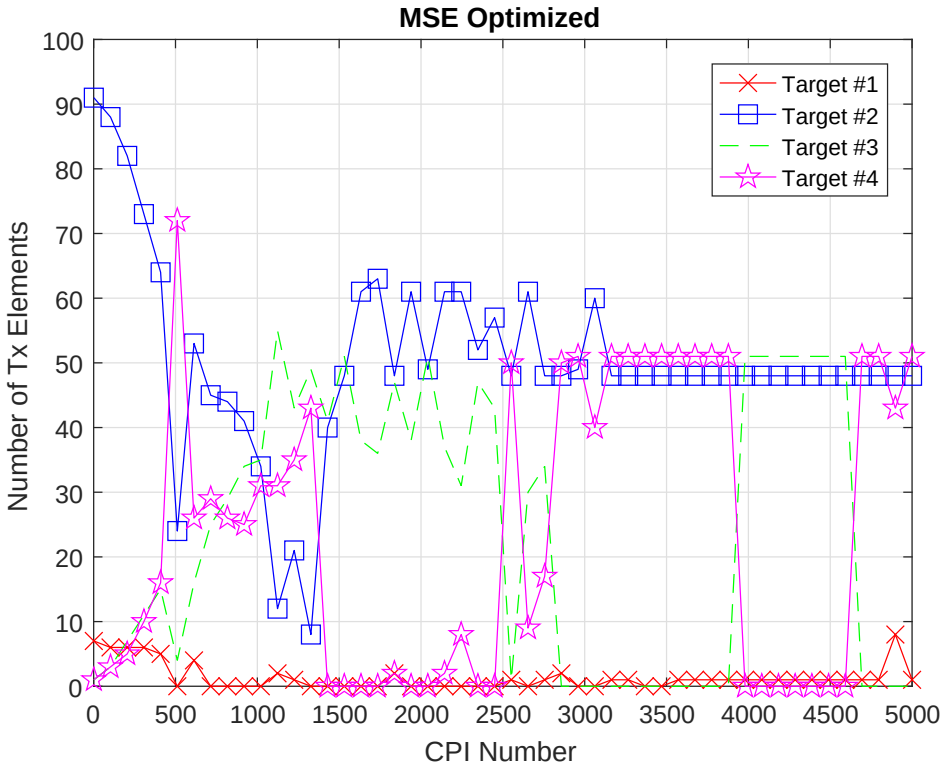

Figure 7.16: Scenario 3 resource allocation for target #1 using the MSE asymptote optimization approach.

and toward target #2, the strategy transfers more and more of target #2's resources between targets #3 and #4. Target #2 reaches zero elements around the same time as target #4 reaches the same range as target #2. At which point, the asymptotic MSE optimized approach immediately allocates half the resources to target #2, which never relinquishes less than 40% of the resources after that. The remaining resources are allocated to either the third or fourth target, with the further away target being allocated the most resources.

The tracking threshold approach is the last strategy and it follows a very similar allocation as the asymptotic MSE optimized approach. The distinct difference between the two approaches is the transition between situations inside the scenario. A different situation occurs when the targets transition

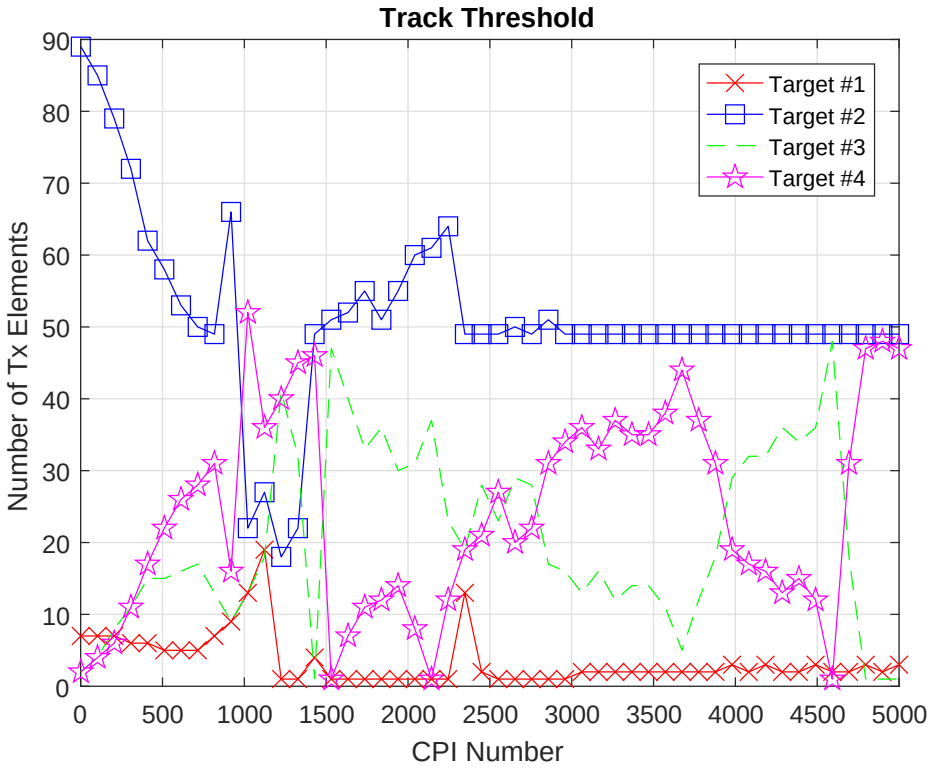

Figure 7.17: Scenario 3 resource allocation for target #1 using the tracking threshold approach.

to a different order in range (in other words, when one target passes another one in range), which causes the allocation to change. The asymptotic MSE optimized approach is more drastic in the allocation changes when a new situation occurred, while the tracking threshold approach is more gradual in its effects.

7.2.4 Varying Number of Targets

Beyond the scenarios described in Sections 7.2.1, 7.2.2, and 7.2.3, 10 Monte Carlo simulations were performed, each of which performed 1,000 realizations (with 1,000 CPIs per realization) with varying initial target locations, initial target velocities, target dynamic variances, and target RCS values. The 10

Monte Carlo simulations varied the number of targets from low (10) to high (100) in increments of 10. The amount of resources is the same between each simulation, such that the simulations with fewer targets have more resources per target while the simulations with many targets have a comparably scarce amount of resources (i.e. resource-constrained). The only aperture allocation strategy that updates the amount of resources throughout the simulation is the greedy approach. All the other approaches only perform the aperture allocation calculation in the beginning. For each Monte Carlo simulation, the average MSE, the number of lost targets, and the number of deleted tracks are used to compare each resource allocation approach.

Figure 7.18 displays the average MSE results for the varying number of targets. With an abundance of resources compared to the number of targets, the asymptotic MSE optimized, MI maximization, and measurement threshold approaches achieve the best average performance. As the amount of resources becomes more scarce, the asymptotic MSE and MI maximization perform closer to the evenly distributed approach while the measurement threshold approach performs 13% better than every other approach when resource constrained.

Figure 7.19 displays the number of deleted tracks based on the SNR levels of each target in consecutive CPIs. A track is deleted when the target's measurement is in the noise floor for five consecutive CPIs. For this metric, the measurement threshold approach performs the best for every number of target simulated, because it is actively optimized to minimize this metric. The MI maximization approach performs second best for 10 targets, but decreases considerably in performance when resources are scarce. The tracking threshold approach performs second best for the Monte Carlo simulations beyond the

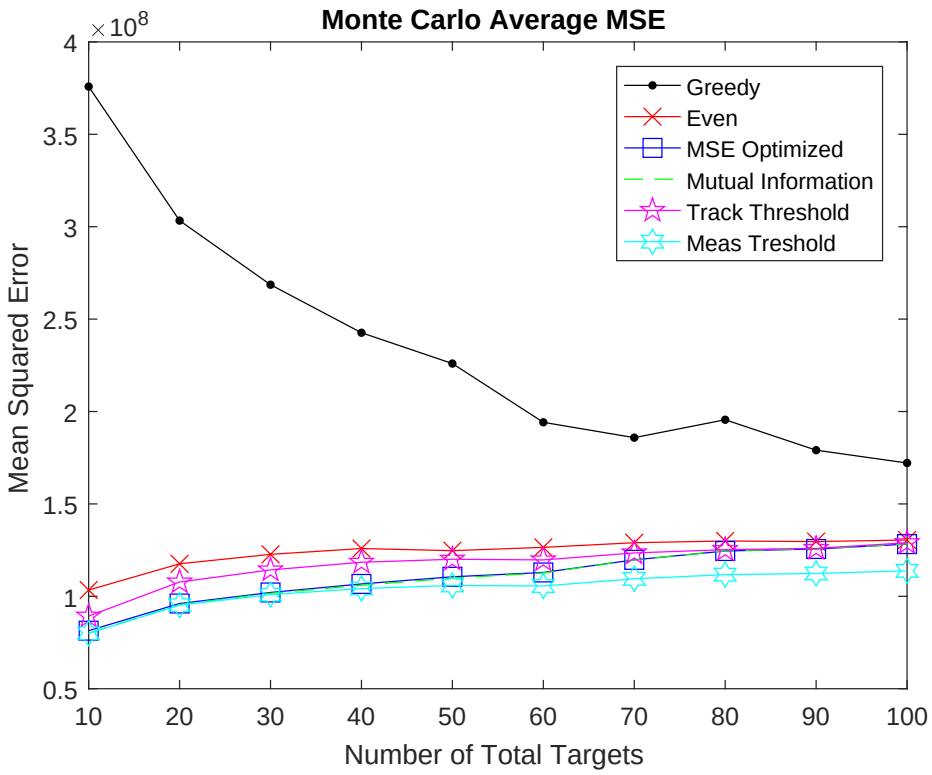

Figure 7.18: Average MSE comparison for a varying number of targets.

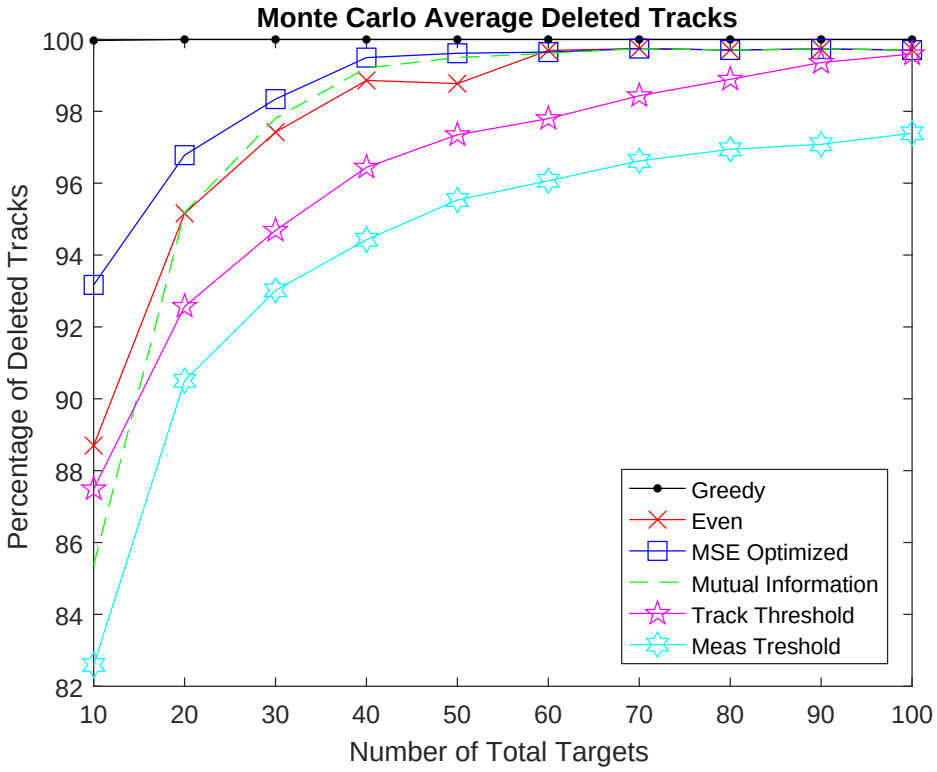

Figure 7.19: Number of deleted tracks comparison for a varying number of targets.

simulation with the least number of targets (10). The asymptotic MSE optimized approach performs worse than the MI maximization, tracking threshold, and evenly distributed approached until they all have similar performance in the resource constrained simulation.

Figure 7.20 shows the number of targets lost for each Monte Carlo simulation. Similar to the deleted tracks metric, the measurement threshold approach performs the best, and the MI maximized approach performs well when the resources are abundant. Unlike the deleted tracks metric, the asymptotic MSE optimized approach performs close to the MI maximized approach. Both the asymptotic MSE optimized and MI maximized approaches degrade in performance as the amount of resources per target decreases until they perform

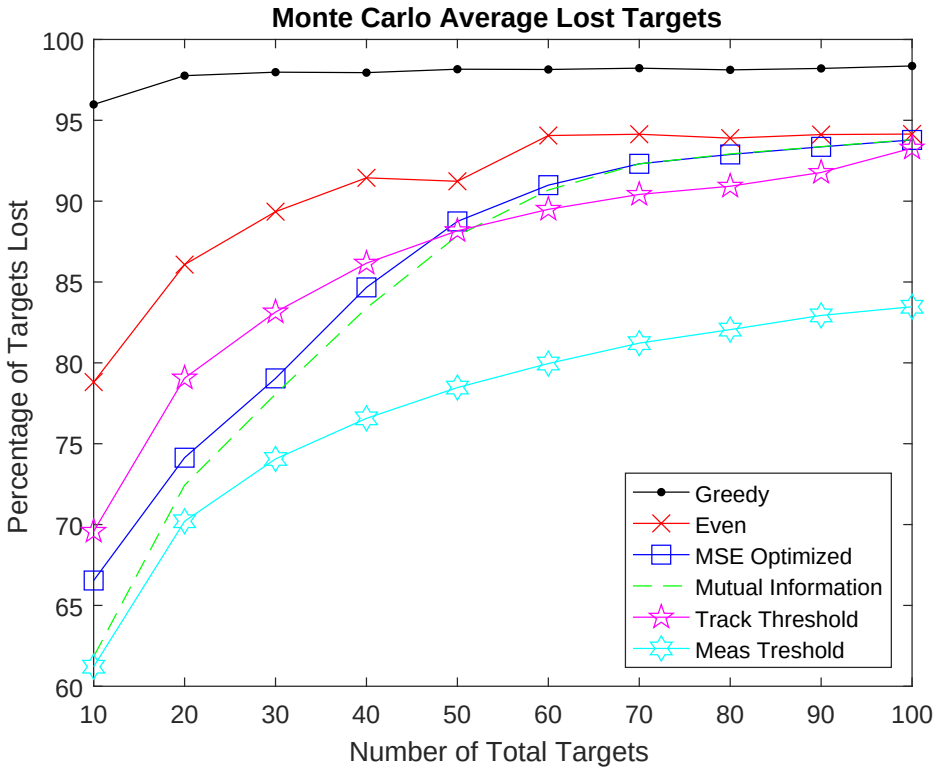

Figure 7.20: Number of lost targets comparison for a varying number of targets.

similar to the even distribution approach for resource constrained scenarios. In the resource constrained scenario, the measurement threshold approach performs about 10% better than the next closest approach.

The results in Figures 7.18, 7.19, and 7.20 are averaged over the range of targets in each figure and quantified in Table 7.2. The measurement threshold approach is obviously the best with a 6% better average MSE, 2% decrease in deleted tracks, and a 9% decrease in lost targets compared to the next closest approach. The table also shows that after the measurement threshold approach, no second best approach exists for all three metrics.

Table 7.2: Average MSE and percentage of lost targets comparison averaged across varying target amounts.

Approach	Average MSE	Percentage of Deleted Tracks	Percentage of Lost Targets
Greedy	2.34×10^8	100%	97.9%
Even Distribution	1.24×10^8	97.8%	90.7%
MSE Asymptote	1.11×10^8	98.6%	85.7%
Mutual Information	1.10×10^8	97.6%	84.7%
Track Treshold	1.17×10^8	96.3%	86.2%
Measurement Threshold	1.04×10^8	94.0%	77.0%

7.2.5 Varying Allocation Update Rates

While Section 7.2.4 varies the number of targets, this section shows the results of 10 Monte Carlo simulations, each of which performed 1,000 realizations (with 1,000 CPIs per realization), with a varied number of CPIs between each resource allocation update. By reducing the number of CPIs between resource allocations, the approaches can update the amount allocated for each target more often. Since the targets move throughout the simulation, intuitively, the approaches would perform better when they update the allocation amounts more often. While the approaches were able to reallocate the amount of resources, the number of targets were constant across simulations at 50. The same three metrics as Section 7.2.4, average MSE, the number of deleted targets, and the number of lost targets, are used for comparison.

The average MSE metric, shown in Figure 7.21, shows that the approaches do not perform drastically different when the number of updates are increased. The appearance of separation between each approach, compared to Figure 7.18, is due to the difference in the y axis scale.

The number of deleted tracks, shown in Figure 7.22, highlights a problem with the given metric. When the allocation is updated more often than

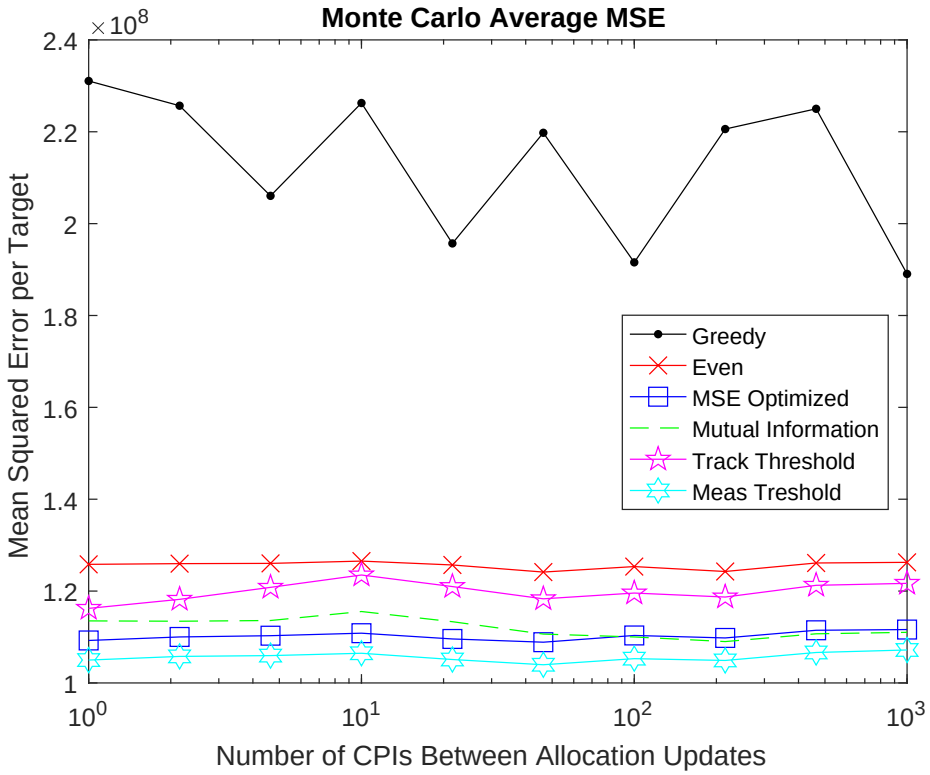

Figure 7.21: Average MSE comparison for varying update rates.

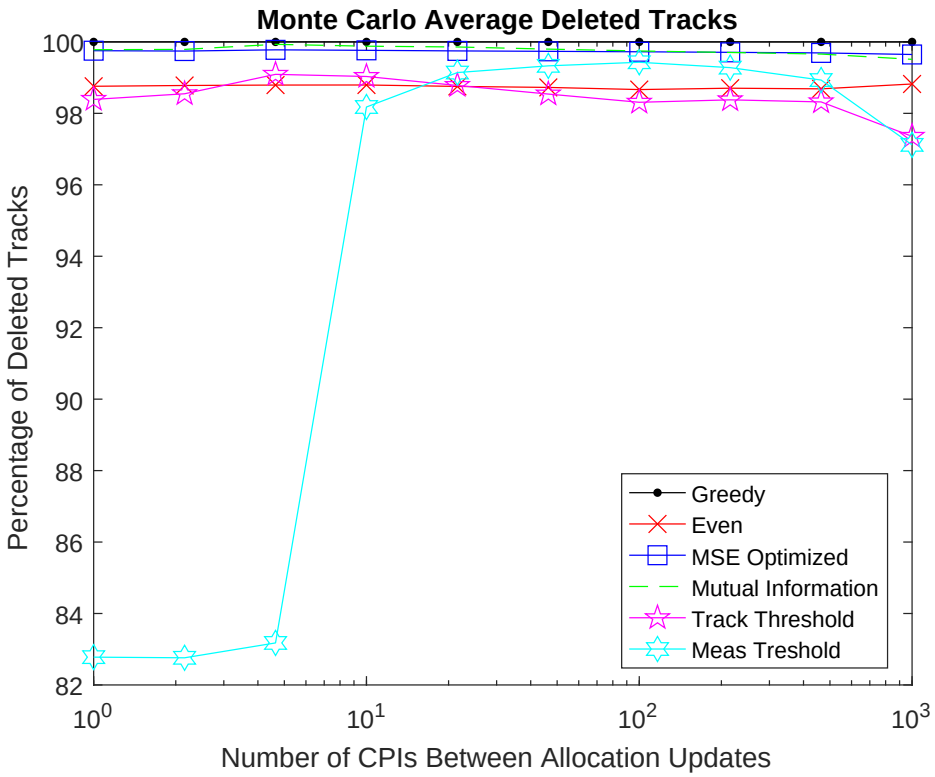

Figure 7.22: Number of deleted tracks comparison for varying update rates.

once in the beginning or less often than the number of CPIs for deletion, the measurement and tracking threshold approaches' performance degrade. Intuitively, any increase in the update frequency beyond once in the beginning should perform better, but the intermediate update frequencies do not follow this expectation. This side effect of the metric is because those approaches can change the targets receiving energy during each update. Those approaches either allocate enough resources to measure a target, or completely ignores a target causing that track to be deleted. When the amount of CPIs between each update is higher than the number of CPIs for deletion, in this case five CPIs, ignoring a target for one update cycle will cause its track to be deleted. The deleted tracks are not removed after they have been marked as deleted, so at each update, those approaches can allocate resources to targets that have already been deleted (while ignoring targets that have not been deleted yet causing them to subsequently be deleted). When the allocation update is every CPI, the frequency is high enough to switch between targets without causing their tracks to be deleted.

Each approaches' performance for the number of lost targets, shown in Figure 7.23, does not vary much when the approaches re-allocate the resources often. Although the number of lost targets metric could have the same problem as the deleted tracks metric, the use of lost targets metric allows the tracker to develop a good estimate of the parameter's state, which can continue recursively without losing track. For example, if the target is traveling in a straight line and the tracker has a strong estimate of its state, even completely ignoring the target and introducing very poor measurements into the tracker might take a while for the tracker to deviate from the predicted straight line. Therefore, switching the targets between updates might not cause as much

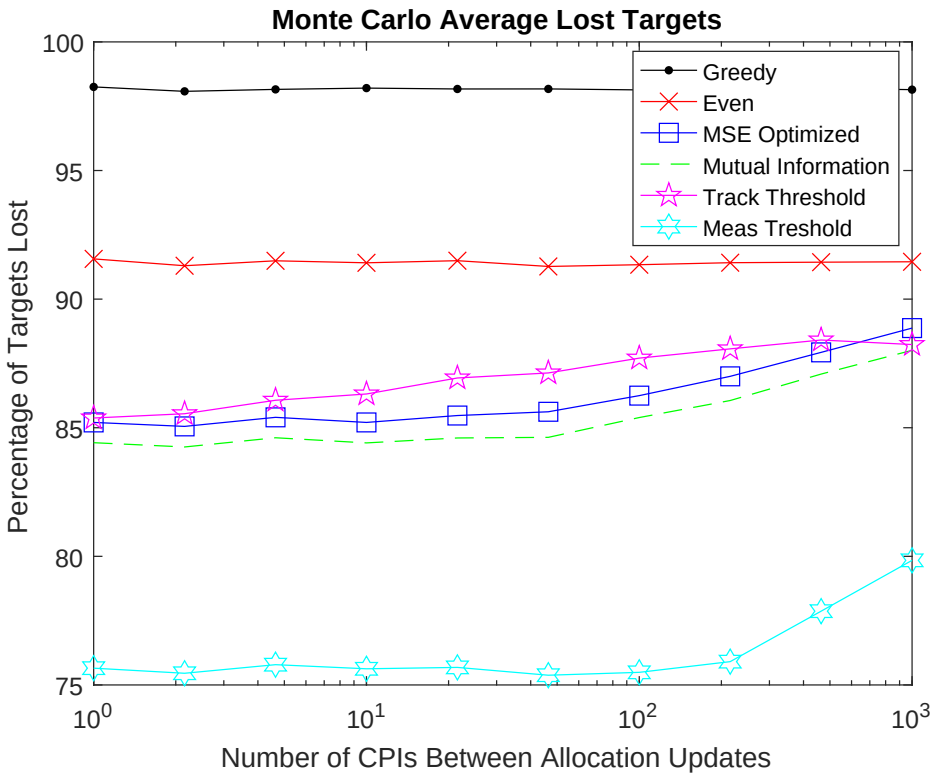

Figure 7.23: Number of lost targets comparison for varying update rates.

Table 7.3: Average MSE and percentage of lost targets comparison averaged across varying update rates.

Approach	Average MSE	Percentage of Deleted Tracks	Percentage of Lost Targets
Greedy	2.13×10^8	100%	98.2%
Even Distribution	1.26×10^8	98.8%	91.4%
MSE Asymptote	1.10×10^8	99.8%	86.2%
Mutual Information	1.12×10^8	99.8%	85.3%
Track Treshold	1.20×10^8	98.5%	87.0%
Measurement Threshold	1.06×10^8	94.0%	76.3%

degradation as compared to the deleted tracks metric.

The results in Figures 7.21, 7.22, and 7.23 are averaged over the range of targets in each figure and quantified in Table 7.3. The measurement threshold approach once again has the best performance overall with comparable results to Section 7.2.4.

Chapter 8

Conclusions and Future Work

8.1 Summary

This dissertation details multiple frameworks to model the measurement and estimation processes for a multi-target tracking radar system. These frameworks range from the simple linear measurement and Bayesian posterior equation estimation to a more realistic non-linear measurement and tracker based estimation model. This dissertation also demonstrates a straightforward model extension for tracking correlated target movements. After describing each framework, multiple aperture allocation strategies are derived by applying different optimization metrics to the given models. Finally, using information theory and tracking metrics, each aperture allocation's performance is quantified for comparison.

Results from Chapter 7 demonstrate that the most important goal for the overall system is to apply enough resources for a given target to increase its measurement above the noise floor. Allocating any resources below this amount for a given target allocates resources to a target without receiving a beneficial measurement, and ultimately wastes the resources without receiving any information about the target. Once the measurements are increased

beyond the noise floor, the allocation strategies in Chapters 4 or 5 should be applied for the targets receiving energy to optimize for the selected metric.

8.1.1 Future Work

Although many allocation approaches were derived for models with varying degrees of realism, much more work is necessary before these strategies are applied to real systems. Chapter 7 demonstrated that the best approach is increasing the target measurements above the noise floor, but it did not compare the best approach for the remaining aperture. To determine the best approach for the remaining aperture, approaches from the linear measurement model in Chapter 4 or 5 should be compared using the tracker based estimation model. Combining the linear measurement and tracker based estimation models will demonstrate the best approach to use with the remaining aperture.

Also, the results demonstrated for this dissertation assumed a spatial independence between targets, meaning that each target is separated enough in angle that their corresponding beams do not overlap and cause interaction between measurements. This assumption simplified the mathematical derivations for this dissertation but would not be true for an actual system. In this case, one measurement can receive information about two targets, although better resolution may be necessary to differentiate any targets located spatially close together.

The approaches described here focused on optimizing the performance based on the Gaussian distribution posterior variance equation, but did not consider the resiliency of a tracker to continue tracking non-maneuvering targets, i.e. targets that are constantly moving in a straight line. Typically, targets with a large dynamic variance can change its current trajectory quickly,

but a large dynamic variance does not mean that the target actually will deviate from a straight line. By incorporating the probability of a target actually maneuvering into the allocation technique, the aperture allocation strategies could implement a priority structure to time interleave measurements accordingly to achieve better performance.

So far, the estimation model has received exact SNR values to calculate the necessary allocation, but an actual system would have to estimate this value based on the received signal strength. This SNR estimation is difficult to achieve in real systems due to variability in components, target reflectivity, and environmental effects over time, but this estimation is critical to allocate the optimal amount of resources.

So far, we have focused on Gaussian distributed parameters throughout this dissertation but not all parameters, such as RCS variability, operate according to this assumption. Extending these approaches to non-Gaussian parameters and using a particle filter tracker to estimate the results may be beneficial for realistic scenarios.

Finally, this dissertation has focused on aperture allocation, but all-digital arrays have many other degrees of freedom that can be leveraged for optimal tracking performance. By including those other degrees of freedom, such as waveform design, fast time, and slow time scheduling, along with the spatial aperture allocation, the radar can efficiently optimize the resources for the necessary tracking performance.

References

[1] S. Pillai, H. Oh, D. Youla, and J. Guerci, "Optimum transmit-receiver design in the presence of signal-dependent interference and noise", *IEEE Trans. on Information Theory*, vol. 46, no. 2, pp. 577–584, 2000.

[2] S. D. Blunt and E. Mokole, "Overview of radar waveform diversity", *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 11, pp. 2–42, 2016.

[3] B. W. Jung, R. Adve, J. Chun, and M. C. Wicks, "Detection performance using frequency diversity with distributed sensors", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 1800–1813, 2011.

[4] N. A. Goodman and D. Bruyere, "Optimum and decentralized detection for multistatic airborne radar", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 806–813, 2007.

[5] C. Fulton, M. Yeary, D. Thompson, J. Lake, and A. Mitchell, "Digital phased arrays: Challenges and opportunities", *Proceedings of the IEEE*, vol. 104, no. 3, pp. 487–503, 2016.

[6] S. H. Talisa, K. W. O'Haver, T. M. Comberiate, M. D. Sharp, and O. F. Somerlock, "Benefits of digital phased array radars", *Proceedings of the IEEE*, vol. 104, no. 3, pp. 530–543, 2016.

[7] A. Charlish, K. Woodbridge, and H. Griffiths, "Information theoretic measures for mfr tracking control", in *Proc. 2010 IEEE Radar Conference*, IEEE, 2010, pp. 987–992.

[8] A. Charlish and F. Hoffmann, "Anticipation in cognitive radar using stochastic control", in *Radar Conference (RadarCon), 2015 IEEE*, IEEE, 2015, pp. 1692–1697.

[9] A. Charlish, K. Woodbridge, and H. Griffiths, "Phased array radar resource management using continuous double auction", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 2212–2224, 2015.

[10] F. Katsilieris, A. Charlish, and Y. Boers, "Towards an online, adaptive algorithm for radar surveillance control", in *Proc. 2012 Workshop on Sensor Data Fusion: Trends, Solutions, Applications*, IEEE, 2012, pp. 66–71.

[11] J. Yan, H. Liu, W. Pu, H. Liu, Z. Liu, and Z. Bao, "Joint threshold adjustment and power allocation for cognitive target tracking in asynchronous radar network", *IEEE Transactions on Signal Processing*, vol. 65, no. 12, pp. 3094–3106, 2017.

[12] J. Yan, B. Jiu, H. Liu, B. Chen, and Z. Bao, "Prior knowledge-based simultaneous multibeam power allocation algorithm for cognitive multiple targets tracking in clutter", *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 512–527, 2015.

[13] W. Huleihel, J. Tabrikian, and T. Shavit, "Optimal adaptive waveform design for cognitive mimo radar", *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 5075–5089, 2013.

[14] H. Godrich, A. P. Petropulu, and H. V. Poor, "Power allocation strategies for target localization in distributed multiple-radar architectures", *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3226–3240, 2011.

[15] N. Garcia, A. M. Haimovich, M. Coulon, and M. Lops, "Resource allocation in mimo radar with multiple targets for non-coherent localization", *IEEE Transactions on Signal Processing*, vol. 62, no. 10, pp. 2656–2666, 2014.

[16] X. Song, N. Zheng, and T. Bai, "Resource allocation schemes for multiple targets tracking in distributed mimo radar systems", *International Journal of Antennas and Propagation*, vol. 2017, 2017.

[17] T. Aittomäki, H. Godrich, H. V. Poor, and V. Koivunen, "Resource allocation for target detection in distributed mimo radars", in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, IEEE, 2011, pp. 873–877.

[18] H. Chen, S. Ta, and B. Sun, "Cooperative game approach to power allocation for target tracking in distributed mimo radar sensor networks", *IEEE Sensors Journal*, vol. 15, no. 10, pp. 5423–5432, 2015.

[19] P. Chavali and A. Nehorai, "Scheduling and power allocation in a cognitive radar network for multiple-target tracking", *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 715–729, 2012.

[20] K. J. Hintz and E. S. McVey, "Multi-process constrained estimation", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 1, pp. 237–244, 1991.

[21] D. R. Fuhrmann, "One-step optimal measurement selection for linear gaussian estimation problems", in *Waveform Diversity and Design Conference, 2007. International*, IEEE, 2007, pp. 224–227.

[22] M. R. Bell, "Information theory and radar waveform design", *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1578–1597, 1993.

[23] J. Chen, F. Wang, and J. Zhou, "Information content based optimal radar waveform design: Lpi's purpose", *Entropy*, vol. 19, no. 5, p. 210, 2017.

[24] M. Bica, K.-W. Huang, V. Koivunen, and U. Mitra, "Mutual information based radar waveform design for joint radar and cellular communication systems", in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, IEEE, 2016, pp. 3671–3675.

[25] Y. Yang and R. S. Blum, "Mimo radar waveform design based on mutual information and minimum mean-square error estimation", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, 2007.

[26] R. A. Romero, J. Bae, and N. A. Goodman, "Theory and application of snr and mutual information matched illumination waveforms", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 912–927, 2011.

[27] R. Romero and N. Goodman, "Waveform design in signal-dependent interference and application to target recognition with multiple transmissions", *IET radar, sonar & navigation*, vol. 3, no. 4, pp. 328–340, 2009.

[28] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[29] R. E. Kalman, "A new approach to linear filtering and prediction problems", *Journal of basic engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[30] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[31] J. L. Lagrange, *Mécanique analytique*. Mallet-Bachelier, 1853, vol. 1.

[32] S. M. Kay, *Fundamentals of statistical signal processing estimation theory*. Prentice Hall PTR, 1993.

[33] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell, *Bayesian multiple target tracking*. Artech House, 2013.

[34] K. Ramachandra, *Kalman filtering techniques for radar tracking*. CRC Press, 2000.

[35] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems", in *Signal processing, sensor fusion, and target recognition VI*, International Society for Optics and Photonics, vol. 3068, 1997, pp. 182–194.

[36] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory", *Journal of basic engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[37] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics", *Journal of Geophysical Research: Oceans*, vol. 99, no. C5, pp. 10 143–10 162, 1994.

[38] F. Daum, "Nonlinear filters: Beyond the kalman filter", *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 57–69, 2005.

[39] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: Particle filters for tracking applications*. Artech house, 2003.

[40] R. A. Davis, "Gaussian process", in *Encyclopedia of Environmetrics*. American Cancer Society, 2006, ISBN: 9780470057339. [Online]. Avail-

able: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470057339.vag002.

[41] A. J. Haug, *bayesian estimation and tracking: a practical guide*. John Wiley & Sons, 2012.

[42] S. S. Blackman, "Multiple-target tracking with radar applications", *Dedham, MA, Artech House, Inc., 1986, 463 p.*, 1986.

[43] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking: Dynamic models", in *Signal and Data Processing of Small Targets 2000*, International Society for Optics and Photonics, vol. 4048, 2000, pp. 212–236.

[44] A. Hoover. (2018). Ece 8540 analysis of tracking systems, [Online]. Available: http://http://cecas.clemson.edu/~ahoover/ece854/ (visited on 03/18/2019).

[45] D. Lucking and N. Goodman, "Resource allocation for multi-variate dynamic gaussian estimation", in *Radar Conference (RadarConf), 2018 IEEE*, IEEE, 2018, pp. 0582–0587.

[46] D. Wackerly, W. Mendenhall, and R. L. Scheaffer, *Mathematical statistics with applications*. Cengage Learning, 2014.

[47] J. She, F. Wang, and J. Zhou, "A novel sensor selection and power allocation algorithm for multiple-target tracking in an lpi radar network", *Sensors*, vol. 16, no. 12, p. 2193, 2016.

[48] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures", *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[49] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in gaussian channels", *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[50] K. P. Murphy, "Conjugate bayesian analysis of the gaussian distribution", *def*, vol. 1, no. $2\sigma 2$, p. 16, 2007.

[51] K. B. Petersen, M. S. Pedersen, *et al.*, "The matrix cookbook", *Technical University of Denmark*, vol. 7, p. 15, 2008.

[52] L Brennan, "Angular accuracy of a phased array radar", *IRE Transactions on antennas and propagation*, vol. 9, no. 3, pp. 268–275, 1961.

[53] W. W. Hager, "Updating the inverse of a matrix", *SIAM review*, vol. 31, no. 2, pp. 221–239, 1989.

[54] M. A. Richards, J. Scheer, and W. A. Holm, *Principles of modern radar: basic principles*. Scitech Publishing, 2010.