The Pennsylvania State University
The Graduate School

OPERATOR-THEORETIC AND DATA-DRIVEN APPROACHES TO RADAR MODELING AND SIGNAL PROCESSING

A Dissertation in
Electrical Engineering
by
Caden J. Pici

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2023

The dissertation of Caden J. Pici was reviewed and approved by the following:

Ram M. Narayanan
Professor of Electrical Engineering
Dissertation Advisor
Chair of Committee

Timothy J. Kane
Professor of Electrical Engineering

Constantino Lagoa
Professor of Electrical Engineering

Asok Ray
Distinguished Professor of Mechanical Engineering and Mathematics

Sastry Kompella
Special Member

Madhavan Swaminathan
Professor of Electrical Engineering
Head of the Electrical Engineering Department

ABSTRACT

Optimizing radar signal processing algorithms is typically addressed under various assumptions about the nature of the targets, clutter, and noise. This can include target motion and signature, clutter distribution, and Gaussian nature of the noise. In this work, a generalized linear operator based model of the radar problem is presented which combines aspects of a couple different works in literature, and overcomes the limitations present in the linear time-invariant model often used. From the general equation for monostatic radar, connections are made to the work presented in the remainder of the thesis, in which the following chapters study different aspects of that equation under different assumptions and conditions. This includes a filtering method for adapting noise radar waveforms to the eigenfunction solution of maximizing the signal-to-interference-plus-noise ratio in the presence of colored noise. The optimal solution to this problem is limiting in that it requires knowledge of the target impulse response. Two target classification receivers are presented which rely on having a library of target response data across different aspect angles. One solution is based on performing a sparse regression onto the space of possible frequency responses. Though providing the best accuracy, this is a slower algorithm. This is addressed by providing a target-tailored matched filter band which is shown to have promise in classifying targets in the simulation study. Finally, an existing stochastic differential equation model of sea clutter is reviewed, and a data-driven framework relying on Koopman operator theory and dynamic mode decomposition is used to discover a model of sea clutter dynamics. This model generates a higher dimensional linear operator that is able to reasonably well model the temporal dynamics of the sea clutter stochastic differential equations. From this model, an anomaly detection approach is studied, which is able to detect changes

in the sea state, and shows potential for future work on small target detection in the presence of sea clutter.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Ram Narayanan, for providing me the opportunity to go to graduate school to continue my education, get the research experience needed to continue my career. Dr. Narayanan provided me the guidance and invaluable connections that allowed me to get to where I am today. I would also like to extend my sincere gratitude to Dr. Sastry Kompella for being a great supervisor during our overlapping time at NRL. His mentorship, enthusiasm for research, and constant support made the last couple of years of finishing grad school while working at NRL an enjoyable experience. Dr. Kompella's support made the transition to working at NRL in the middle of grad school to finish my PhD possible. I would also like to extend a big thank you to Dr. Asok Ray. Dr. Ray exposed me to a new world of research with the numerous classes he taught, and with his encouragement for me to complete a Master's degree in Mathematics. Thank you to all of my committe members for their time. Additionally, I would like to thank all of the friends and colleagues that I've met along the way.

Most importantly, I would like to thank my parents. Without their love, support, and continued belief in me, I don't know that I would be where I am today.

This work was partially funded by the Office of Naval Research Grant N00014-17-1-2386, and by the U.S. Naval Research Laboratory (NRL) Karle Fellowship program. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the funding sponsor.

DEDICATION

To my family.

CHAPTER 1

INTRODUCTION

1.1 Motivation

This dissertation largely looks at linear operator theoretic and data-driven approaches to different radar signal processing problems. The operator theoretic viewpoint begins by defining the radar problem in a way that models the received signal as the action of linear operators on the transmitted signal. This has prior been proposed in literature to be a more general framework than the common linear time-invariant model, as it can include the effects of Doppler. Of specific interest in this model is to look at a particularly difficult kind of time varying clutter, which is known as sea clutter. This is a term that describes the electromagnetic backscatter from the ocean surface. This has been shown to be a nonlinear dynamical system, and this dissertation proposes to use techniques from Koopman operator theory to approach this problem. The Koopman operator is an infinite dimensional linear operator that can be used to model the dynamics of finite dimensional nonlinear systems. A review of this theory, as well as data-driven approaches to its finite dimensional approximation, are given in a later chapter. In addition to the data-driven techniques associated with the aforementioned area of study, this dissertation also looks at data-driven approaches to overcome some limitations in radar target modeling often found in literature. It is common to assume some knowledge of a target's impulse or frequency response, and proceed

with different kinds of analysis from there. This is believed to be too limiting of an assumption in practice, as different target aspect angles can have drastically different responses. This dissertation proposes two classification receiver designs that can begin to address this limitation when the radar has a library of target response data to inform the receiver design.

1.2 Contributions

- Combined some existing literature into a unified model of radar processing by taking advantage of certain classes of linear operators that generalizes linear-time invariant models and can account for Doppler.
- Developed a target classification receiver approach using sparse regression onto a library of possible target response data.
- Developed a target classification receiver based on a target tailored matched filter bank that takes advantage of a library of possible target response data.
- Demonstrated a filtering approach to spectrally shape a noise radar waveform to an optimal SINR maximizing solution.
- Demonstrated a novel approach to modeling sea clutter by building a data-driven model of sea clutter stochastic differential equation dynamics using Koopman operator theory and dynamic mode decomposition.
- Demonstrated the potential of the aforementioned data-driven model for anomaly detection of sea clutter stochastic differential equation data.

1.3 Dissertation Outline

In chapter 2, a general model of the radar problem in the presence of Doppler is reviewed, and combines a couple results to give a more complete framework. The general approach is to define the component operators to be from a more general class than the somewhat standard linear-time invariant model. Analysis using these operators leads to more general versions of common results found in literature. From the general equation of a target in the presence of clutter and noise, certain

unique interpretations are made that connect this work to the other chapters in the dissertation.

In chapter 3, two target classification receivers are presented that take advantage of having libraries of target response data of various target aspect angles, to determine which target is present in a return. The first method involves performing a sparse regression onto the library of possible target responses, and is shown to have very good classification results. The downside of this approach is the computation time needed to get a result. In the second approach, a faster matched filter bank approach is presented that classifies target returns from a bank of target tailored matched filters. The library of response data allows for the derivation of a series of matched filters that captures the correlation structure of the data. Results are demonstrated using an electromagnetically simulated library of four target responses, each containing 180 degrees of aspect angles.

In chapter 4, a filtering approach is shown that spectrally shapes a noise radar waveform to fit the eigenfunction solution of maximizing the signal-to-interference-plus-noise ratio in the presence of colored noise. This takes advantage of an approach from existing literature that derives the eigenfunction solution, and is a special case of the model presented in chapter 2.

In chapter 5, Koopman operator theory and dynamic mode decomposition is reviewed, and the potential for developing data-driven models of dynamical systems is demonstrated.

In chapter 6, a data driven approach is presented to discover a model for the dynamics of sea clutter stochastic differential equations. This approach relies on Koopman operator theory and dynamic mode decomposition which were review in the previous chapter. Directly from measurement data, a model for the temporal dynamics is discovered that maps the data to a high dimensional linear operator. The quality of the model to reconstruct the data is shown. This shows a novel approach to approaching the sea clutter problem, with potential application outlined in the next chapter.

In chapter 7, the previously outlined approach is used for the purpose of anomaly detection. The learned model is applied to a data set not used to train the model, that contains some sort of anomaly inserted into a section of the time series. The possibility for using the DMD model to detect anomalies in sea clutter data is shown to have promise for the application of the technique for small target detection

in the presence of sea clutter. This work sets the stage for future work in further analyzing and optimizing this new target detection approach.

CHAPTER 2

LINEAR OPERATOR MODEL OF RADAR

2.1 Summary

The purpose of this chapter is to present a generalized equation that describes the received signal at the radar as

$$s(t) = Tw(t) + Cw(t) + n(t) \tag{2.1}$$

where T is a linear operator representing the effect of a target on a waveform, C a linear operator describing the effect of clutter, and $n(t)$ is the additive noise component. At the end of this chapter, this version of the radar problem is connected on a chapter by chapter basis to work in the remainder of the dissertation.

It is common in literature to model a radar scene as a linear operator that acts on the transmitted waveform to generate the received signal. Often, this takes the well understood form of a linear time-invariant (LTI) system. This can be a good approximation in the static scenario, without motion in the radar or scene, but is inadequate in Doppler scenarios. This can be understood mathematically as LTI systems having sinusoids as the eigenfunctions, which does not offer a means of modeling Doppler shifts. This was highlighted in [1], where the authors addressed this by examining a larger class of linear operators, known as Hilbert-Schmidt

operators, for narrowband radar scene modeling. In this chapter, the background on the radar problem is presented based primarily around the operator theoretic model in [1], and is extended to the wideband case using some ideas in [2] and tools from [3]. In addition, some additional interpretations and extensions are presented that led to, and motivate, later chapters.

In this chapter, the purpose is to present the radar background in the setting presented in [1], with some further development and interpretation, which will be shown to be a generalized interpretation of the problem. The general idea of this chapter is to depart from the common conventions of modeling problems in radar as simple convolutions in the setting of LTI operators, and present the problem in a more general linear operator framework in which LTI systems are a specific case. To this end, this background chapter will contain more detail than common treatments of radar basics. A hope is that, besides just being interesting in the mathematical sense, the proposed framework can potentially lead to new insights on the problem. Studying the underlying properties and constraints of the class of admissible operators also leads to generalized versions of popular results in literature. Additionally, this approach highlights the mathematical similarities between signal analysis and quantum mechanics (though not the interpretation). Besides the general potential for this structure, thinking of problems in a functional analytic and linear operator sense is what this author has learned to do, and has lead to much of the later research results in this dissertation. Some basic background information on Hilbert spaces and Hilbert-Schmidt operators can be found in Appendix A.

2.2 Time and Frequency Operators

Much of the work in this chapter relies on the idea of displacement operators. In particular, the primary motivation for this framework came from [1], and in that work, a general form of the displacement operator is given for the narrowband assumption. In this section, tools are outlined that will allow for an explicit representation of the displacement operators discussed in the next section, both in the narrowband and wideband regimes.

In general, much of time-frequency analysis can be cast as operators acting on signals. Basic building blocks to this include the time and frequency shift of

signals, as in the displacement operator. An understanding of an operator theoretic viewpoint of time-frequency analysis can be found in [3]. In this section, an explicit representation of the time and frequency shift operators in the notation of [3] will be defined. In general, these shift operators are but special cases of the exponential representation of the Taylor series expansion.

Definition 2.2.1. Define the time operator $\mathcal{T}$ as

$$\mathcal{T} = -\frac{1}{j}\frac{d}{d\omega}$$

Definition 2.2.2. Define the frequency operator $\mathcal{W}$ as

$$\mathcal{W} = \frac{1}{j}\frac{d}{dt}$$

A time shift in a signal can be written based on these operators, where a time shift can be written as

$$e^{j\tau\mathcal{W}} f(t) = f(t+\tau)$$

and on frequency functions as

$$e^{-j\theta\mathcal{T}} S(\omega) = S(\omega+\theta)$$

This can be shown for the time shift using the Taylor series expansion as

$$\begin{aligned} e^{i\tau\mathcal{W}} f(t) &= \sum_{n=0}^{\infty} \frac{(j\tau)^n \mathcal{W}^n}{n!} f(t) \\ &= \sum_{n=0}^{\infty} \frac{(\tau)^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}t^n} f(t) \\ &= f(t+\tau) \end{aligned}$$

The frequency shift can be proven similarly. In addition to being able to represent shifts, it is useful in calculating expected values and higher moments of signals [3]. When dealing with multiple operators acting on a function, one way of interpreting this action is through what is known as the commutator, which is defined for general operators $\mathcal{A}$ and $\mathcal{B}$ as

$$[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$$

with associated anti-commutator defined as

$$[\mathcal{A}, \mathcal{B}]_+ = \mathcal{A}\mathcal{B} + \mathcal{B}\mathcal{A}$$

Definition 2.2.3. The scale operator can be defined using the anti-commutator of the time and frequency operators as

$$\mathcal{C} = \frac{1}{2}[\mathcal{T}, \mathcal{W}]_+ \tag{2.2}$$

This has the property that it transforms a signal according to

$$e^{i\sigma\mathcal{C}}s(t) = e^{\sigma/2}s(e^{\sigma}t) \tag{2.3}$$

This has the effect of rescaling the signal by a compression factor, and is relevant to the wideband version of some signals analysis [3]. Note that by using $\ln\sigma$ this reduces to $e^{i\ln\sigma\mathcal{C}}s(t) = \sqrt{\sigma}s(\sigma t)$

2.3 Displacement Operator

The action of a target on a waveform from the point of view of the receiver is to time delay, and possibly frequency shift the transmitted waveform. Often, a radar scene is modeled as an LTI system, and has value in modeling static scenes. However, this is not the case for a scene involving Doppler, since sinusoids are eigenfunctions of LTI systems. This is the motivation behind the Hilbert-Schmidt operator formalism presented in [1], as this is a larger class of linear operators on $L^2(\mathbb{R})$ that can accommodate Doppler shifts. Fundamental to the construction is the idea of a displacement operator to describe the effects of a scene on the waveform, though only the narrowband case is considered in [1]. The goal is to define a wideband and narrowband displacement operator $D(\alpha, \tau)$ used to describe the action of a point scatterer on a waveform. This will then be put into the context of a general extended target.

2.3.1 Wideband Case

In the wideband representation of the return signal, the waveform is time delayed and scaled to accommodate doppler. Consider the wideband displacement operator

$D_{WB}(\alpha, \tau) : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ that shifts the signal in time by τ and scales by α as

$$D_{WB}(\alpha, \tau) f(t) = \sqrt{\alpha} f(\alpha(t - \tau)) \tag{2.4}$$

where for speed of light c and radial velocity v the scaling parameter is

$$\alpha = \frac{c - v}{c + v} \tag{2.5}$$

Using the previous tools, this can also be written as

$$\begin{aligned} D_{WB}(\alpha, \tau) f(t) &= e^{i \ln \alpha \mathcal{C}} f(t - \tau) \\ &= e^{i \ln \alpha \mathcal{C}} e^{-i\tau \mathcal{W}} f(t) \end{aligned} \tag{2.6}$$

That is, the returned signal can be modeled as the action of the scale change and time shift operators. In this setting, the transmitted waveform can be thought of as the means to measure the observables associated with the target or scene.

2.3.2 Narrowband Case

In the narrowband assumption, the Doppler shift is modeled as a frequency shift of the transmitted waveform. Consider the displacement operator $D(\alpha, \tau) : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by

$$D(\alpha, \tau) f(t) = e^{i\alpha t} e^{-i\tau \mathcal{W}} f(t) \tag{2.7}$$

i.e., the operator that shifts the signal $f(t)$ by τ in time and by α in frequency. This is the most common approach to modeling the radar problem, as most problems fall under the narrowband assumption [4]. This is, in fact, the only situation considered in [1]. Going forward the narrowband version of the displacement operator will be considered unless stated otherwise. This is because the narrowband assumption is typically considered valid for most radar applications. Where appropriate, the wideband version will be denoted with the D_{WB} subscript.

With this displacement operator, another version of the Hilbert-Schmidt operator can be defined. Of particular interest in this chapter, is the following version of a Hilbert-Schmidt class operator. This version is used in [1] to formalize the radar problem, and will be summarized in this chapter.

Definition 2.3.1. A Hilbert-Schmidt class operator $S : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is one that can be written as a superposition of displacement operators

$$S = \int_{\mathbb{R}} \int_{\mathbb{R}} s(\alpha, \tau) D(\alpha, \tau) d\alpha d\tau \tag{2.8}$$

where the coefficients $s(\alpha, \tau)$ form a square summable function of the variables α and τ; i.e,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |s(\alpha, \tau)|^2 d\alpha d\tau < \infty \tag{2.9}$$

Note that for the case of $\alpha = 0$, and using $h(\tau) = s(0, \tau)$, that this reduces to

$$Sf(t) = \int_{\mathbb{R}} h(\tau) f(t - \tau) d\tau, \tag{2.10}$$

and is just the standard definition of convolution. This shows that LTI operators are a subclass of Hilbert-Schmidt operators.

2.4 Scene and Target Modeling

The H-S formalism allows for an effective way to capture the effects of radar backscatter. This section summarizes some of these ideas captured in [1]. For a monostatic radar scenario operating under a narrowband and far-field assumption, consider a target consisting of a collection of point scatterers. The Hilbert-Schmidt displacement operator description of this scenario is

$$S = \sum_{(\alpha, \tau)} s_t(\alpha, \tau) D(\alpha, \tau) \tag{2.11}$$

where $s_t(\alpha, \tau)$ is the complex weight representing the strength of the point scatterer at (α, τ), and the sum is taken over all constituent point scatterer's delays τ and Doppler shifts α. Now, consider a situation in which the constituent scatterers are clustered around a target center at delay and Doppler coordinate (α_0, τ_0). The collection of additional scatterers can be described to be in a neighborhood of this target center $N(\alpha_0, \tau_0)$. In this case, it is possible to rewrite (2.11) as

$$S = \left[\sum_{(\alpha, \tau) \in N(\alpha_0, \tau_0)} s_t(\alpha, \tau) D(\alpha, \tau) \right] D(\alpha_0, \tau_0) \tag{2.12}$$

The displacement operator $D(\alpha_0, \tau_0)$ shifts the response of the cluster to the range and radial velocity of the target center.

A clutter operator can be similarly defined and expressed as

$$C = \sum_{(\alpha,\tau)} s_c(\alpha, \tau) D(\alpha, \tau) \tag{2.13}$$

The weight functions, $s_c : \mathbb{R}^2 \to \mathbb{C}$, can determine the nature of the clutter. There can be advantages to a continuum model, such as in a stochastic case, where (2.13) can be rewritten as

$$C = \int_{\mathbb{R}} \int_{\mathbb{R}} s_c(\alpha, \tau) D(\alpha, \tau) d\alpha d\tau \tag{2.14}$$

where $s_c \in L^2(\mathbb{R}^2)$. There will be a further discussion about the nature of this clutter term in certain scenarios later in the chapter.

2.5 Received Signal Models

This section considers modeling the received signal in a monostatic radar with baseband waveform $w \in L^2(\mathbb{R})$. The return from a single point target and then an extended target in the noise only case is first considered, and then the general case of clutter and noise. This again follows some of the results in [1], with the addition of proofs that further elucidate some of the resulting expressions.

2.5.1 Detection in AWGN

Consider the waveform converted to a carrier frequency Ω with $D(\Omega, 0)w$. The scene consists of one constituent illuminating point scatterer at distance r, velocity v, delay $\tau_0 = 2r/c$, and frequency shift $\alpha_0 \approx 2v\Omega/c$. The signal is corrupted by additive white Gaussian noise with variance σ^2. After mixing to baseband, this results in the following received signal

$$\begin{aligned} x(t) &= D(-\Omega, 0) s(\alpha_0, \tau_0) D(\alpha_0, \tau_0) D(\Omega, 0) w(t) \\ &= e^{-i\Omega\tau_0} e^{i\alpha_0 t} s(\alpha_0, \tau_0) w(t - \tau_0) \\ &= s(\alpha_0, \tau_0) e^{-i\Omega\tau_0} D(\alpha_0, \tau_0) w(t) \end{aligned} \tag{2.15}$$

For the case of an extended target in zero-mean additive white Gaussian noise $n(t)$, the return signal is complicated by the presence of multiple scatterers. The received signal in this case becomes

$$\begin{aligned} x(t) &= \sum_{(\alpha,\tau)} e^{-i\Omega\tau} s_t(\alpha,\tau) D(\alpha,\tau) w(t) + n(t) \\ &= Tw(t) + n(t) \end{aligned} \tag{2.16}$$

where $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the H-S class target response operator defined by

$$T = \sum_{(\alpha,\tau)} e^{-i\Omega\tau} s_t(\alpha,\tau) D(\alpha,\tau)$$

2.5.2 Detection in Clutter and Noise

In addition to the model in the previous section, the transmitted signal is additionally affected by a H-S class clutter operator C. The received signal is then

$$s(t) = Tw(t) + Cw(t) + n(t) \tag{2.17}$$

or when processed with a H-S operator H at decision instant t_0 this expression can be written

$$r(t_0) = (HTw + HCw + Hn)(t_0) \tag{2.18}$$

In many ways (2.17) is the fundamental equation of radar signal processing. A lot of work in literature involves looking at this equation, often in the LTI case, under different kinds of constraints, considerations, and optimization goals. This could include waveform optimization, clutter modeling, target modeling and classification, etc. The rest of the work in this dissertation is no different in this regard. Later in this chapter, the connections between this equation and other chapters will be further explained and highlighted. The signal-to-interference-plus noise (SINR) ratio corresponding to (2.18) is given by

$$\mathrm{SINR}(t_0) = \frac{|HTW(t_0)|^2}{\mathbb{E}[|HCw(t_0) + Hn(t_0)|^2]} \tag{2.19}$$

A typical objective of radar is to maximize the SINR at time t_0 by suitable choice of H. Another somewhat common objective function considered in literature

is the mutual information between the target and waveform. As the receiver, H, is a H-S operator on $L^2(\mathbb{R})$, it can be represented as an integral operator with kernel $\Phi : \mathbb{R} \to \mathbb{C}$. That is, the receiver acting on some function f, where $f \in L^2(\mathbb{R})$, can be written as

$$Hf(t) = \int_{\mathbb{R}} \Phi(t,\tau) f(\tau) d\tau)$$

where the noise term in the denominator can now be written as

$$\begin{aligned} \mathbb{E}[|Hn(t)|^2] &= \int\int \Phi(t,\tau)\overline{\Phi(t,u)}\mathbb{E}[n(u)n(\tau)]dud\tau \\ &= \sigma^2 \int |\Phi(t,\tau)|^2 d\tau \end{aligned} \tag{2.20}$$

To simplify notation, define $h_{t_0} = \Phi(t_0, \cdot)$, where $h \in L^2(\mathbb{R})$. This simplifies the above expression to $\mathbb{E}[Hn(t)]^2 = \sigma^2 \|h_{t_0}\|^2$. Similarly, the clutter term can be written as

$$\begin{aligned} \mathbb{E}[|HCw(t_0)|^2] &= \int\int h_{t_0}\overline{h_{t_0}}\mathbb{E}[]Cw(\tau)\overline{Cw(u)}]d\tau du \\ &= \langle h_{t_0}, G_C h_{t_0} \rangle \end{aligned} \tag{2.21}$$

The term G_C is a non-negative definite Hermitian operator defined by

$$G_C f(t) = \int_{\mathbb{R}} \mathbb{E}[\overline{Cw(t)}Cw(t)]f(\tau)d\tau$$

The clutter operator here is modeled as random, and which results in the expectation in this expression. In addition, the clutter is assumed to be independent to the noise. However, for deterministic clutter, the clutter operator can be written as

$$G_C = CP_wC^\dagger \tag{2.22}$$

where $P_w f = \langle f, w \rangle w$ for $f \in L^2(\mathbb{R})$. In words, this is the rank-one projection operator from $L^2(\mathbb{R})$ onto the one-dimensional subspace spanned by w. As the proof was not given in [1], it can be shown in the following.

$$\begin{aligned}\|HCw(t_0)\|^2 &= \int\int h_{t_0}(\tau)\overline{h_{t_0}(u)}Cw(\tau)\overline{Cw(u)}d\tau du \\ &= \int h_{t_0}(\tau)C\langle\overline{h_{t_0}}, Cw\rangle w(\tau)d\tau \\ &= \int h_{t_0}(\tau)C\langle C^\dagger\overline{h_{t_0}}, w\rangle w(\tau)d\tau \\ &= \langle h_{t_0}, CP_wC^\dagger h_{t_0}\rangle\end{aligned}$$

Similarly, the signal term in the numerator of the SINR expression may be written in inner product form as

$$|HTw(t_0)|^2 = \langle h_{t_0}, P_{Tw}h_{t_0}\rangle \tag{2.23}$$

where P_{Tw} is the rank-one projection operator from $L^2(\mathbb{R})$ onto the one-dimensional subspace spanned by Tw, and is non-negative definite. Now, the SINR expression (2.19) can be rewritten as

$$\text{SINR}(t_0) = \frac{\langle h_{t_0}, P_{Tw}h_{t_0}\rangle}{\langle h_{t_0}, (\sigma^2\mathbb{I} + G_C)h_{t_0}\rangle} \tag{2.24}$$

where since G_C is non-negative definite, the operator $(\sigma^2\mathbb{I} + G_C)$ is positive definite as long as $\sigma^2 > 0$. Thus, maximizing the SINR is a generalized Rayleigh quotient, or a generalized eigenvalue problem. The h_{t_0} that maximizes this expression is

$$h_{t_0} = h_{\max} = (\sigma^2\mathbb{I} + G_C)^{-1/2}Tw \tag{2.25}$$

which agrees with the usual notion of a matched filter. Now, denoting $G = (\sigma^2\mathbb{I}+G_C)$ and noting that this is positive definite and self-adjoint, the maximum value of this expression can be written for $f = Tw$ as

$$\begin{aligned}\text{SINR}_{max} &= \frac{\langle G^{-1/2}f, P_{Tw}G^{-1/2}f\rangle}{\langle G^{-1/2}f, G^{1/2}f\rangle} \\ &= \frac{\langle f, G^{-1/2}P_{Tw}G^{-1/2}f\rangle}{\langle f, f\rangle}\end{aligned} \tag{2.26}$$

Again, this is a Rayleigh quotient with a maximum value corresponding to the maximum eigenvalue of the non-negative definite operator $G^{-1/2}P_{Tw}G^{-1/2}$. However, this is a rank-one operator so the one positive eigenvalue is equal to the trace of the operator. Therefore,

$$\begin{aligned} \text{SINR}_{max} &= tr(G^{-1/2} P_{Tw} G^{-1/2}) \\ &= \langle G^{-1/2} f, G^{-1/2} f \rangle \\ &= \langle Tw, (\sigma^2 \mathbb{I} + G_C)^{-1} Tw \rangle \end{aligned} \tag{2.27}$$

The SINR maximizing waveform can then be written as

$$w_{max} = \arg\max_{w \in W} \langle w, T^\dagger (\sigma^2 \mathbb{I} + G_C)^{-1} Tw \rangle \tag{2.28}$$

where W is the class of allowable waveforms. Certain energy and constant modulus constraints can limit this class in practice. More will be discussed about this optimization problem later in this chapter. The next sections will include a further look and interpretations of the framework presented in [1] that are relevant to later work in this dissertation.

2.6 Ambiguity Function

The radar ambiguity function is a two dimensional function of the delay and Doppler, or τ and α. It can be thought of as a two dimensional matched filter, representing the return from a moving target. Additionally, it can be used as a means to evaluate radar waveforms. In this section, operator theoretic considerations to defining the ambiguity function will be presented. The explicit forms of the displacement operators given previously will be relevant, and takes advantage of some of the results in [2], and tools in [3]. However, while [2] only considers point targets, it is extended here to extended targets. This section takes advantage of the mathematical formalism connecting time-frequency analysis and quantum mechanics. In particular, these tools allow the ambiguity function to be written as an expected value of an operator over a density, similar to defining the expected value of an observable in quantum mechanics. Though interesting as a mathematical trick in its own right, this interpretation could lead to new results in exploring other interactions of the environment with a waveform. The following first goes over the standard definition for a point target, where the returned signal is a delayed and shifted version of the original, and then gets into a definition for an extended target.

2.6.1 Point Target

2.6.1.1 Wideband

The wideband ambiguity function can be written as

$$\begin{aligned}
\chi_{WB}(\omega,\tau) &= \sqrt{\alpha}\int_{-\infty}^{\infty} s(t)s^*(\alpha(t+\tau))dt \\
&= \sqrt{\alpha}\int_{-\infty}^{\infty} s^*(t)D_{WB}(\alpha,\tau)s(t)dt \\
&= \sqrt{\alpha}\int_{-\infty}^{\infty} D_{WB}(\alpha,\tau)|s(t)|^2dt \\
&= \sqrt{\alpha}\langle s(t), D_{WB}(\alpha,\tau)s(t)\rangle \\
&= \sqrt{\alpha}\langle D_{WB}(\alpha,\tau)\rangle \\
&= \sqrt{\alpha}\langle e^{i\ln\alpha\mathcal{C}}e^{-i\tau\mathcal{W}}\rangle
\end{aligned} \tag{2.29}$$

where the wideband notation is used to denote the alternate form of the displacement operator, and the final line follows from the physics notation as defining the expected value of an observable A over a density as $\langle A\rangle$. This gives the ambiguity function as the expected value of the operators on a density $|s(t)|^2$.

2.6.1.2 Narrowband

The point target ambiguity function can then be written as

$$\begin{aligned}
\chi(\omega,\tau) &= \int_{-\infty}^{\infty} s(t)s^*(t+\tau)e^{-i\omega t}dt \\
&= \int_{-\infty}^{\infty} s^*(t)D(\alpha,\tau)s(t)dt \\
&= \int_{-\infty}^{\infty} D_N(\alpha,\tau)|s(t)|^2dt \\
&= \langle s(t), D_N(\alpha,\tau)s(t)\rangle \\
&= \langle D_N(\alpha,\tau)\rangle \\
&= \langle e^{i\alpha t}e^{-i\tau\mathcal{W}}\rangle
\end{aligned} \tag{2.30}$$

A standard property of the ambiguity function is that for the squared magnitude, all values are less than or equal to the maximum value at the delay and Doppler coordinates $(0,0)$. This maximum value corresponds to the waveform energy squared, which will be denoted E_s^2.

2.6.2 Extended Target

In an extended target model, the ambiguity function will be a function of both the transmit waveform and the target response. In this framework, that will take the form

$$\int_{-\infty}^{\infty} s^*(t)Ts(t)dt = \langle Ts(t), s(t)\rangle \tag{2.31}$$

The maximum value of this extended target version of the ambiguity function is interesting, as it is slightly more complicated than the standard point target scenario. Looking at the squared magnitude, and invoking the Cauchy-Schwarz inequality, the maximum can be found as

$$|\langle Ts, s\rangle|^2 \leq \langle s, s\rangle\langle Ts, Ts\rangle = \|s\|^2\|Ts\|^2 \tag{2.32}$$

Assuming constant waveform energy, the right side will be maximized for the waveform being the eigenfunction corresponding to the maximum eigenvalue, $\lambda_{\max}$ of the non-negative definite Hermitian operator $T^\dagger T$. This property can then be written as

$$|\langle Ts, s\rangle|^2 \leq \lambda_{\max}\|s_{\max}\|^2 \tag{2.33}$$

While a traditional ambiguity function has a peak of $\|s\|^2$, the eigenwaveform used on an extended target is amplified by $\lambda_{\max}$.

2.6.3 Additional Interactions

The operator viewpoint of the radar ambiguity function can give a framework to allow for additional signal interactions beyond just time-delay and Doppler [2]. Signal interactions between the transmitted signal and the target and environment can include additional effects such as micro-Doppler signatures, or electromagnetic scattering parameters such as the reflection coefficient or dielectric properties. In the framework of the linear operator model, this can be interpreted as the action of additional operators on the ambiguity kernel. One such approach to this is presented in [2], where these additional interactions are considered through the action of a generator in a commutator equation. Series expansion gives a familiar

exponential Taylor series expression for the operator. See the aforementioned paper for additional details on this argument. In general, if $\hat{H}$ is an operator generating an additional signal interaction, then this operator can be combined in the kernel as

$$\chi(\omega, \tau) = \langle \hat{H}_1 ... \hat{H}_n e^{i \ln \alpha \mathcal{C}} e^{-i\tau \mathcal{W}} \rangle \tag{2.34}$$

It is hypothesized that this approach can lead to additional processing methods and unify some existing physics inspired signal analysis. Further work will consider the explicit structure of additional operators that capture additional interactions and properties in the scattered signal.

2.7 Chapter 3 Target Operator Considerations

This section highlights connections between this linear operator model of radar, and the work presented in Chapter 3 of this thesis. Specifically, practical considerations for the higher dimensional structure of the target operator before being projected onto the transmitted waveform. This framework will later be used for the purpose of target classification.

Often in literature, the target operator is modeled with some LTI response function, and algorithms may be built upon this assumption. For example, many works consider optimal waveform shaping under this assumption. The analysis thus far in a generalized H-S setting is similar. The scene operators will be projected onto the one-dimensional subspace spanned by the waveform. Hidden in both this, and the typical model is that there is also a projection onto the radar's, or the observers, location. This is because a target operator will vary as a function of the radar's location, as the scattering centers will change with orientation. In radar, this is called the aspect angle of a target. Implications of this could include that the target operator could be considered time varying, or that the larger dimensional space that the target operator lives in can be exploited.

Consider, again, the following radar scene model

$$S = \int_{\mathbb{R}} \int_{\mathbb{R}} s(\alpha, \tau) D(\alpha, \tau) d\alpha d\tau$$

A way to write this that accounts explicitly for the information being left behind

after acting on a waveform is given below

$$S = SP_w + R \tag{2.35}$$

where $Rw = 0$, and P_w is the one-dimensional orthogonal projection defined previously. Contained explicitly in this representation is the target operator information in R that effectively gets left behind. This can be interpreted in a way, as the waveform selecting a one dimensional subspace of the higher dimensional scene or target operator. This is reminiscent of ideas in compressed sensing. Typical detection algorithms, such as the matched filter, make this rank one projection assumption implicitly. An example of how this interpretation of the target operator might be exploited is shown in chapter 3. There, a study of using a higher dimensional version of this target operator in receiver design for the static target scenario is given. Two receivers are presented that can take advantage of the information in this higher dimensional target space.

2.8 Chapter 4 Optimizing Waveform in Presence of Additive Colored Noise

This section studies the problem of maximizing the SINR of the returned signal in the presence of additive colored noise, and is the basis of Chapter 4 noise radar waveform design work. These noise characteristics for example could emerge from additional interference or jamming effects. The goal is to show a generalized version of a similar LTI matrix derivation in [5,6]. First, a characterization of the mean and covariance operator is given in the Hilbert space setting. The covariance operator is then used in defining a noise whitening filter.

2.8.1 Covariance Operator

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X : \Omega \to \mathcal{H}$ be a random variable in $\mathcal{H}$ with probability measure P. Furthermore, assume that $\mathbb{E}\|X\|^2 = \int_{\mathcal{H}} \|x\|^2 dP(x) < \infty$. This condition is sufficient for the existence of both the mean and covariance operator. Define the functional $f : \mathcal{H} \to \mathbb{R}$ by $f(y) = \int_{\mathcal{H}} \langle x, y \rangle dP(x)$. This is a linear functional that can be shown to be bounded on the Hilbert space.

Then, by the Riesz representation theorem there exists unique $\mu \in \mathcal{H}$ such that $f(y) = \langle \mu, y \rangle, \forall y \in \mathcal{H}$. This can equivalently be written as

$$\mathbb{E}\langle X, y \rangle = \langle \mu, y \rangle \tag{2.36}$$

where μ is the expectation of the random variable, now denoted as $\mathbb{E}(X)$. Now, consider a bilinear functional $\varphi : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ given by

$$\varphi(y, z) = \int_{\mathcal{H}} \langle x - \mu, y \rangle \langle x - \mu, z \rangle dP(x)$$

where $\mu = \mathbb{E}(X)$. This functional can be shown to be bounded, so therefore by the Riesz representation theorem there exists a unique bounded linear operator, $\mathcal{C} : \mathcal{H} \to \mathcal{H}$, such that

$$\begin{aligned} \langle \mathcal{C}y, z \rangle = \varphi(y, z) &= \int_{\mathcal{H}} \langle x - \mu, y \rangle \langle x - \mu, z \rangle dP(x) \\ &= \mathbb{E}(\langle X - \mu, y \rangle \langle X - \mu, z \rangle), \quad \forall y, z \in \mathcal{H} \end{aligned} \tag{2.37}$$

Now, $\mathcal{C}$ is called the covariance operator of X. This will additionally be denoted by $\text{cov}(X)$. This can be shown to be a positive self-adjoint operator on $\mathcal{H}$, and additionally to have a finite trace. It is therefore also a Hilbert-Schmidt operator.

2.8.2 Jointly Optimizing Transmitter and Receiver

Consider the received signal comprising a target in the presence of additive colored noise

$$s(t) = Tw + n \tag{2.38}$$

The optimum receiver can be shown to be a noise whitening filter, followed by a white noise matched filter. The whitening filter is a function of the covariance of the noise. If $R = \text{cov}(n)$ denotes the noise covariance, then the corresponding whitening filter is given by

$$\hat{H} = R^{-1/2} \tag{2.39}$$

The noise whitening property can be verified as follows:

$$\begin{aligned}
\mathrm{cov}(\hat{H}n) &= \langle \mathcal{C}\hat{H}n, \hat{H}n\rangle \\
&= \int_{\mathcal{H}} \langle x-\mu, \hat{H}n\rangle\langle x-\mu, \hat{H}n\rangle dP(x) \\
&= \int_{\mathcal{H}} \langle x-\mu, (\langle \mathcal{C}n, n\rangle)^{-1/2}n\rangle\langle x-\mu, (\langle \mathcal{C}n, n\rangle)^{-1/2}n\rangle dP(x) \\
&= (\langle \mathcal{C}n, n\rangle)^{-1}\int_{\mathcal{H}} \langle x-\mu, n\rangle\langle x-\mu, n\rangle dP(x) \\
&= (\langle \mathcal{C}n, n\rangle)^{-1}(\langle \mathcal{C}n, n\rangle) \\
&= \mathbb{I}
\end{aligned}$$

The SNR can now be expressed as

$$\mathrm{SINR} = \frac{|H\hat{H}Tw|^2}{\mathbb{E}[|H\hat{H}n|^2]} \tag{2.40}$$

Now, from previous analysis define $f = \hat{H}Tw$ and $G = \mathbb{I}$, resulting in the following expression for the maximum of the SNR

$$\begin{aligned}
\mathrm{SINR}_{\max} &= \langle \hat{H}Tw, \hat{H}Tw\rangle \\
&= \langle w, T^{\dagger}\hat{H}^{\dagger}\hat{H}Tw\rangle
\end{aligned} \tag{2.41}$$

The optimal waveform can now be expressed as finding

$$w_{\max} = \arg\max_{w} \langle w, T^{\dagger}\hat{H}^{\dagger}\hat{H}Tw\rangle \tag{2.42}$$

This now gives a generalized version of the similar result found in [5, 6] for the finite-dimensional matrix case. This framework, for the LTI case, will be the focus of a noise radar waveform approach in chapter 4.

2.9 Chapter 7 Clutter Operator Considerations

In this dissertation, Chapters 6 and 7 present a framework for learning a model of electromagnetic scattering from a random dynamical media. Specifically, the scattered signal from sea clutter. The work is applied to a physically motivated model of sea clutter resulting in a model governed by a set of stochastic differential equations. Here, the clutter operator can be shown to reduce to the scattering

model that leads to these SDEs under certain assumptions.

Consider again the clutter operator representation given by

$$C = \sum_{(\alpha,\tau)} s_c(\alpha,\tau) D(\alpha,\tau)$$

As mentioned previously, the weight functions, $s_c : \mathbb{R}^2 \to \mathbb{C}$, can determine the nature of the clutter. They can be deterministic, Gaussian, a Markov random field, or have some other statistical characterization [1]. In all of these cases, the clutter is assumed to be a time independent operator. A comparably more difficult scene for radar is that of sea clutter, or the electromagnetic backscatter from the ocean surface. Unlike many land scenarios, sea clutter is a dynamical system, and can if fact be shown to be a nonlinear dynamical system [7]. This kind of clutter changes on appreciable enough time scales that it can cause difficulties for small target detection on the ocean surface. For the Rayleight scattering case, explicitly writing the clutter operator in terms of the scatterer population size N, and adding a time dependence to the weight functions, gives

$$C_t^{(N)} = \sum_{j=1}^{N} s_t^j(\alpha,\tau) D^j(\alpha,\tau) \tag{2.43}$$

The weights can be decomposed as consisting of a component scatterer's weight and relative phase as

$$s_t^j = a_j \exp(i\phi_t^{(j)}) \tag{2.44}$$

Putting this back into the clutter operator equation gives

$$C_t^{(N)} = \sum_{j=1}^{N} a_j \exp(i\phi_t^{(j)}) D^j(\alpha,\tau) \tag{2.45}$$

These component phases, $\phi_t^{(j)}$, can be considered a stochastic process. If they are considered to evolve as a collection of displaced Wiener processes, then under certain conditions, the resulting system can be shown to evolve according to a stochastic differential equation. This is known as Field's model of electromagnetic scattering from random media [8]. This is a very interesting interpretation of clutter in its own right and is overviewed and studied in chapters 6 and 7 of this dissertation.

CHAPTER 3

TARGET CLASSIFICATION RECEIVERS

3.1 Introduction

The radar target classification problem seeks to not only detect the presence of a target in a scene, but to exploit the nature of the reflection to determine the source of the returns. In particular, the problem could be cast as determining the target from some prior list of known classes. Designing a radar system for the task of classification can include optimization of the transmit waveform, receiver, post-processing algorithms, or some combination thereof. Inherent in the objective of classification is the problem of class representation. This factors into which features can be exploited in the classification procedure. Possibilities could include anything from an understanding of a target's radar cross section (RCS), transfer function, Doppler or micro-Doppler signature, etc. The nature of the signals or features to be used depends on the location of the classification operation in the signal processing chain. There are advantages and disadvantages to classifying at different locations, such as the computational cost of the algorithm, the time scale at which decisions are desired to be made, etc.

An often considered version of the radar classification problem is in the domain of synthetic aperture radar (SAR), in which a series of returns is first processed into

images, and the classification is performed on the images. This is a longer, more computationally expensive approach if real time classification is desired. Another approach to the problem is to attempt to classify the target returns directly. This is often seen in the context of waveform optimization through some metric such as signal-to-interference-plus-noise ratio (SINR) or mutual information (MI) to some ensemble of target impulse response functions. Results towards this front seem to usually start with the assumption of some prior knowledge of the target responses.

In practice, the difficulty in assigning returns to different target classes not only increases with the number of possible targets, but also with the return variations from different target aspect angles [9]. In the case of simple targets, such as a sphere or a corner reflector, analytical expressions describing characteristics are known, like the RCS as a function of viewing angle. Of more practical interest involves more complex targets such as vehicles, ships, or planes. These can generally be considered to fall under the category of extended targets, which are targets of a large enough physical size that they no longer behave like point targets. The returns from point targets are often considered to behave as a time delayed, scaled, and possibly frequency shifted version of the transmitted waveform. As a result of the target's size, extended targets exhibit interference and resonance effects in the scattered field, resulting in a more complex reflected signal [10]. This is especially true when considering wideband waveforms, which will henceforth be the use case in this chapter, wherein the range resolution cells are small enough that most practical targets would be distributed across some number of range gates. The aspect angle uncertainty in such scenarios adds to the complexity, as different parts of a target could contribute to different range gates depending on the angle. This, along with interference effects, can lead to rapid changes in scattering characteristics with aspect angle [9].

Target representation in the classification problem is therefore an important problem to consider, as one target would consist of what could be many possible subclasses. An often considered approach to characterizing these effects in signal processing is to consider the target as having an impulse or frequency response which is applied to the waveform in the received signal model. It is thus important to develop or simulate libraries of these profiles for use in the classification algorithm. Electromagnetic and CAD software packages provide a means for generating the frequency dependent scattering information of objects. This could include RCS or

impulse response profiles. These could also be used to generate varying look or aspect angles of the objects, and hence make dictionaries of target class information. Work that utilized this approach can be found in [11] and [12, Ch. 14], wherein techniques for waveform optimization for target classification were covered. The waveform optimization approaches, as outlined in [13], consider either a known transfer function, or treat it as a stochastic process. For simulation, a publicly found CAD model of an F-16 jet was used in finite difference time domain (FDTD) based software to generate frequency response profiles over some variation of target aspect angles. In [11], these simulated target profiles were then used for waveform design in automatic target recognition. Angle variations were dealt with by averaging templates over some sector of angles, and assuming knowledge about the bearing information of a target to reduce the number of possible angular sectors. Simulated results reduce the number of hypotheses to just four in each trial. Bayesian updates over repeatedly adapted transmissions was performed to achieve desired classification accuracy, as in [14]. There, matched illumination waveforms are combined with a sequential hypothesis testing procedure to identify a return from among a set of targets with known impulse response profiles. The approach relies on regularly adapting waveforms in response to previous measurements.

Much of the work on classification by optimizing radar waveforms for extended targets that are represented as an impulse or frequency response stems from [10]. The theory and procedure for maximizing the signal-to-noise ratio (SNR) and MI for extended targets was introduced in [10]. In [15], the problem of target aspect uncertainty is studied in the context of finding the optimal transmission pulse shape. For target identification, the approach is to maximize the square of the Mahalanobis distance. In the case of aspect angle uncertainty, averages are taken over the aspect dependence. In the recent work [16], the focus is on a nonadaptive procedure for optimal classification-based waveform design using multiclass Fisher analysis. The class description assumptions are that each class contains angle dependent subclasses of impulse/frequency responses that are modeled as complex Gaussian random vectors. Objective functions were derived under this assumption for waveform design, and for different class and subclass numbers. In [17], the problem of target classification and adaptive waveform design in cognitive radar networks was considered. Target classes are represented by sets of impulse responses from varying aspect angles. Aspect uncertainty is dealt with by assuming that the

system consists of a network of radars monitoring a moving target, and the multiple observations from the multiple radars can be used to estimate the unknown aspect angle. Another domain of target recognition involves micro-Doppler signature classification [18]. Micromotion in certain kinds of targets introduces sideband behavior in the spectral signature of the radar returns. Many different approaches have been proposed for modeling, characterizing, and classifying such features [19,20]. A good overview on many different aspects of target recognition and classification can be found in [9].

A commonality among many approaches, possibly stemming from [10], is that targets are represented by known transfer functions, and the waveform effectively captures that function through convolution, and carries it back to the receiver. Detection and classification decisions are then made based on that return. The benefit of capturing that transfer function with some kind of optimized waveform is to have control over what information is carried back or accentuated, to improve performance according to some metric. More precisely, the energy spectrum of the waveform can be constructed such that it interacts with the target's equivalent spectral response to maximize some objective function. The approach seems to become limited with growing libraries of response profiles and varying aspect angles. A common approach to deal with this seems to be in the spirit of [14], which involves pulse-to-pulse sequential hypothesis testing and intermediate waveform adaptions. In addition, it seems common to assume that target motion would allow the system to first estimate the orientation, and also to average over sections of angles to reduce the possible classes. In this chapter, instead the focus is on optimizing the receiver operation under the assumption of having libraries of response profiles. The objective is to have a classification approach that is invariant to the unknown target aspect angle, and does not rely on sequential or adaptive methods for operation.

Our contributions include two approaches to the classification problem using libraries of response profiles. The first is based on sparse regression techniques to classify both target class and aspect angle. From the signal model, it is argued that this is the best approach to solving the problem directly. That is, determining which element from the large dictionary of target and aspect angle profiles generated the return. However, this approach could have its limitations. This includes the computational cost with solving the associated optimization problem. For this reason, an approach is developed that attempts to solve the problem in a way that

is invariant to the specific target aspect angle. These two approaches could be seen as either complementary or independent, but both attempt to classify based on the same underlying signal model and structure of the target class dictionaries. This sparse regression approach will be used as a baseline for the second method, which utilizes a matched filtering approach for target classification over the same dictionaries of response profiles containing varying target aspect angles. The proposed matched filtering approach is to find an aspect angle invariant class representation and filtering scheme for classification, and therefore does not rely on or make assumptions about target orientation to reduce class possibilities.

This chapter is organized as follows. First the signal model is outlined in Section 3.2. Using the signal model, the target classification approaches are outlined and/or derived in Sections 3.3 and 3.4 for the sparse regression and the matched filtering approaches, respectively. The approaches are then verified in simulation using RCS data over four simulated target CAD models in Section 3.5. Section VI provides conclusions.

3.2 Signal Model

Consider a pulsed radar system operating in a scenario where an extended target has already been detected, and needs to be classified into a predefined list of targets. The discrete representations of all signals will be considered. Let $\mathbf{x} \in \mathbb{C}^m$ be a vector of discrete samples of an arbitrary transmit waveform over some bandwidth. The targets are assumed extended targets, and can be represented as linear time-invariant systems for non-moving targets and radar platform. Then, the reflected signal from a target is the convolution between a complex target impulse response, $\mathbf{g}_k \in \mathbb{C}^m$, and the transmitted waveform. The subscript, $k \in \{1, \ldots, M\}$, in the impulse response denotes that the target response comes from one of M possible classes. In addition, the signal is corrupted by clutter $\mathbf{c} \in \mathbb{C}^m$, and additive white Gaussian noise $\mathbf{n} \in \mathbb{C}^m$, as shown in (3.1).

$$\mathbf{y} = \mathbf{x} * (\mathbf{g}_k + \mathbf{c}) + \mathbf{n} \tag{3.1}$$

For an extended target, the returned signal across some bandwidth, incidence angle, and aspect angle can be characterized by an associated impulse response. As

mentioned, profiles can be simulated in electromagnetic software for some discretized variation of these parameters, and a dictionary of impulse response profiles can be generated for each target class. For the k-th target, these dictionaries will be represented by an $m \times n$ matrix $\mathbf{\Phi}_k$, where m is the number of discrete impulse or frequency response samples, and n is the number of simulated azimuth or incidence angles.

It is considered in the radar scenario that detection has already been performed, and the target is assumed to be from a set of known target classes. A given target return will be assumed to come from a discrete angle, corresponding to a response represented by a column of the dictionary. That is, for an $n \times 1$ vector $\mathbf{s}$ with a one at the location of the correct target angle, and zeroes everywhere else, the discrete target impulse response for target k can be represented as in (3.2). The aspect angle vector, $\mathbf{s}$, is unknown during operation. This problem setup could allow for the use of sparse regression techniques to find the unknown $\mathbf{s}$. This is considered in the next section.

$$\mathbf{g}_k = \mathbf{\Phi}_k \mathbf{s} \tag{3.2}$$

The received signal can then be modeled as in (3.3). This equation has the two unknowns of the class and aspect angle represented by the target class dictionary matrix $\mathbf{\Phi}_k$ and aspect angle vector $\mathbf{s}$. Consider target dictionaries each containing 180 different aspect angles at one incidence angle. Then, trying to distinguish between two different targets at unknown aspect results in a classification problem over 360 possible subclasses. Since the target echo will come from only one target and aspect angle, a sparsity promoting algorithm would be a direct way to find the specific profile that generated the return. As will be seen in the next section, this would involve solving an optimization problem that may not be suitable for real time operation. As the number of aspect angles in a dictionary and number of possible classes increases, the number of possible subclasses for an algorithm to sort through could become prohibitively large. For this reason, finding a single classification approach that is invariant to the specific aspect angle that generated the target echo is a motivating problem of this work.

$$\mathbf{y} = \mathbf{x} * \mathbf{\Phi}_k \mathbf{s} + \mathbf{x} * \mathbf{c} + \mathbf{n} \tag{3.3}$$

As mentioned, the target dictionary could alternatively be represented by frequency response or RCS profiles, which are both related to the impulse response. Well known is the frequency response, which is just the Fourier transform of the impulse response. In this case, the dictionary will be denoted as $\hat{\mathbf{\Phi}}_k$. This could in practice be scaled and limited to the frequencies contained in the waveform bandwidth. However, the relation to the RCS is not as direct.

Radar cross section is a measure of the ratio of the scattered power density, P_{scat}, at a distance R from the target to the incident power density, P_{inc}, at the target. The general definition of RCS is shown in (3.4) [21]. It would therefore make sense for it to be related to the energy spectral density (ESD) of the target response.

$$\sigma = \lim_{R\to\infty} 4\pi R^2 \frac{P_{\text{scat}}}{P_{\text{inc}}} \tag{3.4}$$

Recalling the signal processing relations for the ESD, first denote the ESD of waveform $\mathbf{x}(t)$ as $\mathbf{\Psi}_x(f)$. Then the filtering of this signal with a filter having frequency response $H(f)$, results in the output $\mathbf{y}(t)$ having ESD $\mathbf{\Psi}_y(f) = |H(f)|^2\mathbf{\Psi}_x(f)$. Therefore, the RCS will be treated as the squared magnitude of the frequency response of the target.

3.3 Sparse Identification of Target Class and Aspect Angle

Motivated by the structure of the signal model, first considered is the problem of target class and aspect angle classification. Such a problem is a desired first step as it would output the most information, namely the class and angle. This is done without making any simplifying assumptions about the nature of the problem, such as having already estimated a target's heading, and hence that the target is moving in the first place. If a dictionary consists of M targets with n discrete angle responses, this will result in a classification problem over Mn classes, which will increase the complexity of the problem, and make it comparatively difficult for real time application. This will therefore be considered as the baseline approach, in which a filtering-based scheme will be compared later.

Built upon the previous signal formulation, the approach is to model the

problem as a sparse regression or optimization over the unknown aspect angle vector. This involves either an ℓ_1 minimization or regularization which promotes sparse solutions to the problem, as can be commonly seen in compressed sensing literature. Therefore, the comparative speed of the parallelized aspect invariant filtering algorithm in the previous section is sacrificed to obtain more information in form of the exact target aspect angle.

Regression seeks to take the predictions and outcomes of a system respectively contained in a matrix $\mathbf{A}$ and a vector $\mathbf{b}$, and find the vector $\mathbf{x}$ most related with the outcomes in $\mathbf{b}$ according to some metric. This is written mathematically as $\mathbf{Ax} = \mathbf{b}$, and the regression problem can be cast in various ways that typically come down to minimizing some ℓ_p norm. If $p = 2$, then this leads to the common least-squares problem, and if $p = 1$, then this is a sparsity promoting norm often used as a convex relaxation of the ℓ_0 pseudo-norm.

In least-squares regression, the vector $\mathbf{x}$ ends up with energy distributed throughout its elements, or in other words, it ends up with many non-zero coefficients. Under the sparsity assumption, it is desired to find the solution with mostly zeroes except for a small number of coefficients. In this chapter, the least absolute shrinkage and selection operator (LASSO) [22] regression technique is used, which adds an l_1 penalty to the least-squares problem to better promote more sparse solutions. The signal model presented in this chapter naturally leads to sparse solution techniques, as is outlined next.

The frequency domain noise only representation of (3.3) will be used for the regression problem to allow for a simpler direct comparison with the returned signal. This is done by considering all relevant signals across discretized frequencies in the bandwidth of interest. The structure of (3.2) does not change, as the columns will now contain the frequency response at each discrete angle. Since the multiplications will now be element-wise, the resulting Hadamard product will be done be representing elements of the transmit waveform in a diagonal matrix, $D_{\hat{\mathbf{x}}} = \text{diag}(\mathbf{x})$. The resulting returned signal is shown in (3.5a). Combining the waveform and frequency response dictionary into $\hat{\tilde{\mathbf{\Phi}}}_k$ gives the simplified expression in (3.5b).

$$\hat{\mathbf{y}} = D_{\hat{\mathbf{x}}}\hat{\mathbf{\Phi}}_k\mathbf{s} + D_{\hat{\mathbf{x}}}\hat{\mathbf{c}} + \hat{\mathbf{n}} \tag{3.5a}$$

$$= \hat{\tilde{\mathbf{\Phi}}}_k \mathbf{s} + D_{\hat{\mathbf{x}}} \hat{\mathbf{c}} + \hat{\mathbf{n}} \tag{3.5b}$$

A dictionary of all target class frequency response profiles is shown in (3.6), where $\mathbf{\Phi}$ is now an $m \times Mn$ matrix. A target return will still be assumed to come from one discrete angle from one target class.

$$\mathbf{\Phi} = [\hat{\tilde{\mathbf{\Phi}}}_1 \hat{\tilde{\mathbf{\Phi}}}_2 \dots \hat{\tilde{\mathbf{\Phi}}}_M] \tag{3.6}$$

The returned signal $\hat{\mathbf{y}}$ is then a noisy return of some column of this matrix. The regression problem will seek to find the sparsest $Mn \times 1$ aspect angle vector, $\mathbf{s}'$, that matches returned signal to the dictionary of all classes. There are different ways to cast this as an optimization problem, with perhaps the most common being to minimize the ℓ_1 norm of the unknown vector $\mathbf{s}'$ subject to the equation constraint. Different formulations of the problem could result in different tools and algorithms to find the solution. Due to this, a proper complexity analysis will not be included in this chapter. Instead, as a general statement, the iterative algorithms needed to solve the optimization will be more computationally expensive than a standard filtering operation from digital signal processing. The form of the optimization problem chosen in this chapter is the ℓ_1 penalized least-squares for the case of zero clutter in (3.7). The chosen approach is the sparsity promoting LASSO regression technique [22], which varies the penalty parameter λ through a range of values, and can typically be cross-validated to minimize the mean squared error (MSE), but could be chosen to more strongly weight the penalty term.

$$\mathbf{s}' = \arg\min_{\mathbf{s}'} \|\mathbf{\Phi}\mathbf{s}' - \hat{\mathbf{y}}\|_2 + \lambda \|\mathbf{s}'\|_1 \tag{3.7}$$

This approach will differ from the matched filtering approach in the next section in that it involves an optimization and classification over a much larger number of possibilities. Additionally, this is computationally slower than the parallelizable filtering approach, specifically if it can be made to be invariant to aspect angle. This would particularly be true in the case of a low SNR, where the solution would be more difficult to fit. Additionally, this requires an optimization among perfectly known target response profiles. This may not be a reasonable assumption in practice. Combined, this may make the sparse regression based classification not likely a good candidate for real time classification. The approach in the next

section addresses this problem by relaxing the need to classify aspect angle, and finding an approach relying on parallelizable standard digital filtering.

3.4 Aspect Angle Invariant Matched Filter for Classification

Inherent in the sparse regression problem is the large number of classes that result from classifying both target class and angle. In relaxing the goal of determining the aspect angle, the possible number of classes drops to M, which is a significant decrease. In practice, this may be a reasonable assumption, as target class would likely be the more important parameter. This is the approach taken here. The next target classification approach is to develop a filtering scheme, such that the output for the correct target is maximized from among all possible classes.

In finding a filter which is matched to the target's response, there are practical issues to be considered due to the unknown target orientation and variations target response that result. The desired filtering scheme would account for these variations through an orientation invariant filter, or set of filters, matched to the correct target. Consider the binary hypothesis testing problem of deciding between two possible targets.

$$\begin{aligned} H_1: &\quad \mathbf{x} * \mathbf{g}_1 + \mathbf{x} * \mathbf{c} + \mathbf{n} \\ H_2: &\quad \mathbf{x} * \mathbf{g}_2 + \mathbf{x} * \mathbf{c} + \mathbf{n} \end{aligned} \tag{3.8}$$

where $\mathbf{g}_1$ and $\mathbf{g}_2$ are the different target impulse responses. It is assumed that detection has been performed such that a target is found present. The objective is to find matched filter tailored to a target's response, followed by a classification decision to the largest output. This is diagrammatically shown in Fig. 3.1, where the yet to be determined filtering for target k is represented as $\mathbf{h}_k$. Detecting one of M known signals using a bank of matched filters, specifically in the case of additive white Gaussian noise, is a problem that shows up in digital communications. The case of designing the matched filtering blocks under dictionaries of possible target responses is shown here.

It is desired to derive a data-driven matched filtering scheme tailored to the dictionaries of target responses. The case of known, or simulated, target impulse

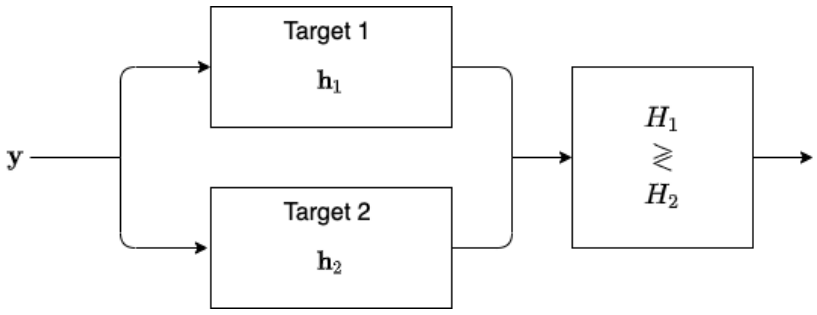


Figure 3.1: Block diagram of target signal model for binary hypothesis testing.

responses will be considered. As mentioned in the previous section, the impulse response can be related to the target RCS, another simulated output of electromagnetic modeling software. The result will thus be derived for known impulse response, and extend it as discussed to dictionaries of RCS data.

Usual derivations of the matched filter involve finding some filter, $\mathbf{h}$, that maximizes the output signal-to-noise ratio (SNR). This results in a filter based on the transmitted waveform that will find a peak correlation with the received signal at the correct delay. For a point target, where the returned signal is just a scaled, time-delayed, and possibly frequency shifted version of the transmitted waveform, this undoubtedly results in one optimal matched filter. As can be seen in [10] and [13], in the case of an extended target, there is also a dependence on the target response.

Instead of finding a filter that maximizes SNR for detection performance, it is desired to maximize the output SNR for the correct class relative to all other classes. The signal and noise components of the signal can be broken down into

$$\begin{aligned} \mathbf{z}_s &= \mathbf{h}_k * \mathbf{x} * \mathbf{g}_k \\ \mathbf{z}_n &= \mathbf{h}_k * (\mathbf{x} * \mathbf{c} + \mathbf{n}) \end{aligned} \tag{3.9}$$

For simplification, denote the causal convolution matrix of the waveform as $\mathbf{X}$, given by

$$\mathbf{X} = \begin{bmatrix} x(1) & 0 & & \cdots & 0 \\ x(2) & x(1) & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & x(m) & x(m-1) \\ 0 & \cdots & & 0 & x(m) \end{bmatrix}, \tag{3.10}$$

Then, the modified dictionary is written as $\tilde{\mathbf{\Phi}}_k = \mathbf{X}\mathbf{\Phi}_k$. At the point of complete overlap between the filter and received signal, for the k'th target class, the SNR can be represented as in (3.11). The numerator is the energy in the signal component of the received vector. The waveform spectrum could also be divided out on receive, which would then impact the denominator of the expressions. The numerator is given by the squared magnitude of the inner product of the filter and signal. This is similarly done in the denominator for the filter and stochastic noise vector.

$$\begin{aligned} \mathrm{SNR}^k &= \frac{|\mathbf{z}_s|^2}{\mathbb{E}\left[|\mathbf{z}_n|^2\right]} \\ &= \frac{|\mathbf{h}_k^* \tilde{\mathbf{\Phi}}_k \mathbf{s}|^2}{\mathbb{E}\left[|\mathbf{h}_k^* \mathbf{X}\mathbf{c} + \mathbf{h}_k^* \mathbf{n}|^2\right]} \end{aligned} \tag{3.11}$$

For uncorrelated clutter and noise components, the denominator simplifies to

$$\begin{aligned} \mathbb{E}\left[|\mathbf{h}_k^* \mathbf{X}\mathbf{c} + \mathbf{h}_k^* \mathbf{n}|^2\right] &= \mathbf{h}_k^* \mathbf{R}_c \mathbf{h}_k + \mathbf{h}_k^* \mathbf{R}_n \mathbf{h}_k \\ &= \mathbf{h}_k^* \mathbf{R}_{c+n} \mathbf{h}_k \end{aligned}$$

where $\mathbf{R}_n = \mathbb{E}[\mathbf{n}\mathbf{n}^*]$ and $\mathbf{R}_c = \mathbb{E}[\mathbf{X}\mathbf{c}\mathbf{c}^*\mathbf{X}^*]$ are the covariance matrices for noise and clutter, respectively. The combined noise and clutter covariance is then represented as $\mathbf{R}_{c+n}$. This SNR expression depends on the unknown aspect angle vector $\mathbf{s}$, and thus cannot be maximized directly. Instead, it is chosen to find the filter that maximizes the SNR with the whole column space of the target dictionary. This is done by dropping the dependence on the aspect angle vector, and summing the energy of the inner product of the filter and each column of that dictionary. The new SNR expression is shown in (3.12).

$$\begin{aligned} \mathrm{SNR}_{\mathrm{tot}}^k &= \frac{\|\mathbf{h}_k^* \tilde{\mathbf{\Phi}}_k\|^2}{\mathbf{h}_k^* \mathbf{R}_{c+n} \mathbf{h}_k} \\ &= \frac{\mathbf{h}_k^* \tilde{\mathbf{\Phi}}_k \tilde{\mathbf{\Phi}}_k^* \mathbf{h}_k}{\mathbf{h}_k^* \mathbf{R}_{c+n} \mathbf{h}_k} \end{aligned} \tag{3.12}$$

In the denominator, the expression is expanded as usual, where $\mathbf{R}_n = \mathbb{E}[\mathbf{n}\mathbf{n}^*]$ is the covariance matrix. This is a generalized Rayleigh quotient, of which it is known that the vector $\mathbf{h}$ that maximizes the expression is the eigenvector corresponding to the maximum eigenvalue of $\mathbf{R}_{c+n}^{-1}\tilde{\mathbf{\Phi}}_k\tilde{\mathbf{\Phi}}_k^*$. For the case of white Gaussian noise only, with a variance σ^2 and impulse autocorrelation, this becomes $\mathbf{R}_n = \sigma^2\mathbf{I}$. The SNR expression can then be further simplified to

$$= \frac{1}{\sigma}\frac{\mathbf{h}_k^*\tilde{\mathbf{\Phi}}_k\tilde{\mathbf{\Phi}}_k^*\mathbf{h}_k}{\mathbf{h}_k^*\mathbf{h}_k}, \tag{3.13}$$

which is a standard Rayleigh quotient, maximized for the eigenvector corresponding to the maximum eigenvalue of $\mathbf{\Phi}_k\mathbf{\Phi}_k^*$. It is worth noting the difference between this, and the standard matched filter derivation. For the signal component of a received vector $\mathbf{y}_s$, the SNR expression is similarly found with the maximum eigenvector of $\mathbf{y}_s\mathbf{y}_s^*$. This is an outer product of a vector with itself, and hence can be shown to be a matrix of rank one. This scenario corresponds to having only one eigenvector, which can also be shown to be the original waveform in $\mathbf{y}_s$. This problem differs due to the nature of the inner matrix. With the assumption that the dictionary of responses has a rank greater than one, application of the min-max theorem results in a series of matched filters for this problem corresponding to the eigenvectors of $\mathbf{\Phi}_k\mathbf{\Phi}_k^*$.

This result has a connection with the matrix decomposition technique known as the singular value decomposition (SVD). For completeness, before exploring this, some basic background of this technique will first be given. The SVD is a unique matrix factorization that exists for every real or complex valued matrix $\mathbf{M}$, and is commonly expressed as

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \tag{3.14}$$

where $\mathbf{U} \in \mathbb{C}^{m\times m}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ are unitary matrices with orthonormal columns, and $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ is a matrix with non-negative real entries on the diagonal and with zeroes otherwise. The decomposition is usually ordered from largest to smallest singular values across the diagonal of $\mathbf{\Sigma}$, with the columns of $\mathbf{U}$ and $\mathbf{V}$ reordered accordingly.

The SVD is closely related to an eigenvalue problem involving the correlation matrix $\mathbf{\Phi}\mathbf{\Phi}^*$ that arises in (3.13). To see this, notice

$$\mathbf{\Phi\Phi}^* = \mathbf{U\Sigma V}^*\mathbf{V\Sigma U}^* = \mathbf{U\Sigma^2 U}^* \tag{3.15}$$

and recalling that that $\mathbf{U}$ and $\mathbf{V}$ are unitary, this results in the following eigenvalue problem

$$\mathbf{\Phi\Phi}^*\mathbf{U} = \mathbf{U\Sigma^2} \tag{3.16}$$

Thus, the columns of $\mathbf{U}$ are eigenvectors of the correlation matrix $\mathbf{\Phi\Phi}^*$. Since the singular values are arranged in descending order by magnitude, this results in the singular vectors in $\mathbf{U}$ being ordered by how much correlation they represent in the columns of the original data matrix $\mathbf{\Phi}$. A similar result would occur if the correlation matrix was instead $\mathbf{\Phi}^*\mathbf{\Phi}$, but with the eigenvectors instead being contained in $\mathbf{V}$. Henceforth, when referring to the singular vectors, this is referring to the left singular vectors of the decomposition.

Thus, the k vectors that maximizes (3.13) are the left singular vectors corresponding to the maximum k singular values of $\mathbf{\Phi}$. Additionally, this gives the best k vectors that correlate with the column space of the transfer matrix. This is a result that also decreases the computational cost of finding the eigenvectors, as $\mathbf{\Phi\Phi}^*$ is a much larger matrix in comparison.

After taking the SVD of the transfer matrix, the singular vectors corresponding to the largest singular values will contain the dominantly correlated vectors of the data matrices. In other words, across frequency (left singular vectors), it will obtain a sequence of filters most correlated with the column space of the transfer data.

As this approach results in a series of matched filters, this additional information could be used to improve results of the classification problem, as compared to using just a single matched filter. The test statistic is thus obtained from how strongly the return signal correlates to the most dominant singular vectors of a given target class transfer matrix. Denote the first r left singular vectors for the kth target as $\mathbf{U}_i^r$. The classification decision, D, is then found using

$$D = \max_k \|\tilde{\mathbf{y}}^*\mathbf{U}_k^r\|^2 \tag{3.17}$$

The result is the maximum magnitude of the projection of the received vector for among all target classes. This is shown in Fig. 3.2.

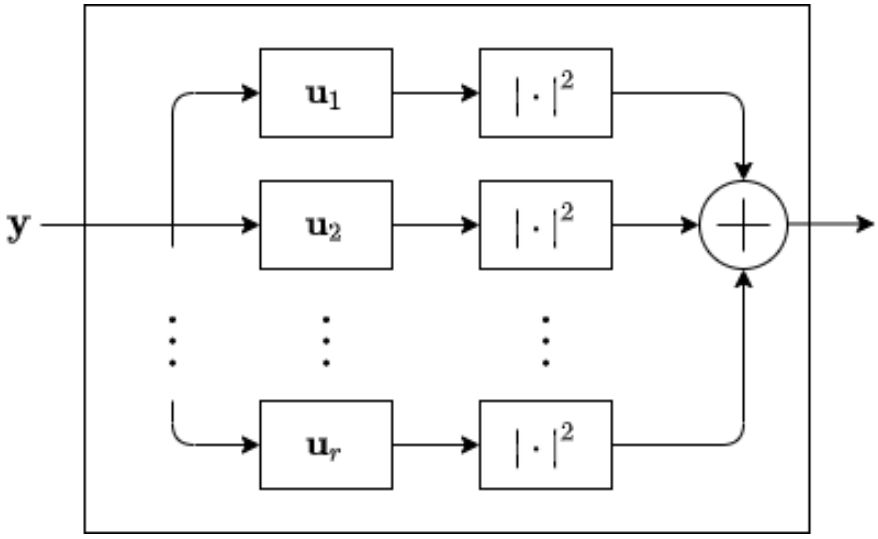


Figure 3.2: Block diagram of a matched filtering block of Fig. 3.1.

3.5 Simulation Results

To test the classification approach, target class and response profiles are generated from CAD models of multiple vehicles. The models used consist of publicly available CAD models of a hummer, car, small propeller airplane, and tank. All targets were assumed to be made of a perfect electric conductor (PEC), except for tires, which were assumed to be made of rubber. These CAD models can be seen in Fig. 3.3, and were imported into electromagnetic field simulation software to generate RCS profiles for each target.

Each target was illuminated with a vertically polarized plane wave. The simulation generated the RCS with 1000 samples over 9-10 GHz at an elevation angle of 70 degrees, and 180 degrees of azimuth sampled at each integer degree around its axis of symmetry, therefore giving the full azimuth of the targets. Each class description thus results in a 1000×180 matrix, for four targets, resulting in 720 total profiles in which a return could originate. The resulting RCS heat maps are shown in Fig. 3.4, where the vertical axis is frequency, the horizontal axis the varying azimuth angles, and magnitudes plotted in decibels referenced to a square meter (dBsm). In these images, zero degrees refers to the view from the front of the vehicle while 180 degrees refers to the view from the rear. The targets are assumed

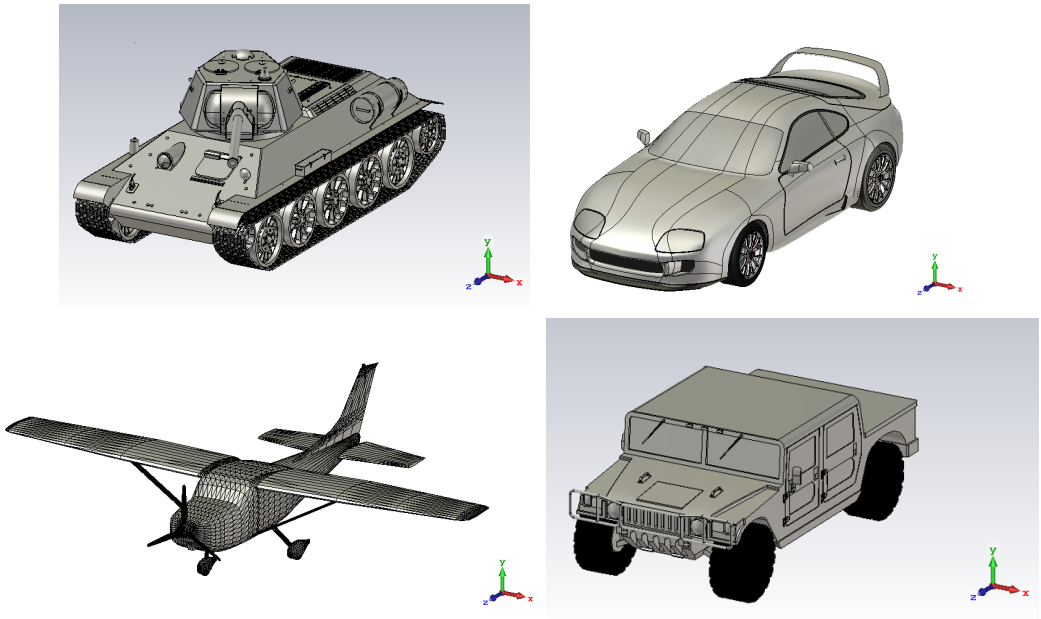


Figure 3.3: Target class CAD model for generating RCS data. Models were obtained from [23]. Clockwise from top left: tank, car. hummer, airplane. Specific CAD models used are available upon request.

to be symmetric on both sides of the vehicle's axis; therefore, the figures are shown only over the range 0–180 degrees.

It is clear that each target has areas of stronger and weaker reflections, and its RCS varies not only with azimuth, but also frequency. This variation can be more clearly seen in an example 45 degree cut from each target in Fig. 3.5. Before using these as filter profiles, each column is normalized in the linear scale, and will be used as the squared magnitude of the frequency response of the target in the signal processing. All subsequent results utilized a linear frequency modulated (LFM) pulse as the transmitted waveform over the entire 9-10 GHz bandwidth.

3.5.1 Sparse Regression Based Classification

Sparse regression based classification performance was simulated using these data sets. The LASSO [22] algorithm was used in MATLAB to solve (3.7) for each SNR value. This solves for the unknown sparse concatenated aspect angle vector $\mathbf{s}'$ that fits the return. The solution is then chosen as the vector element with the largest

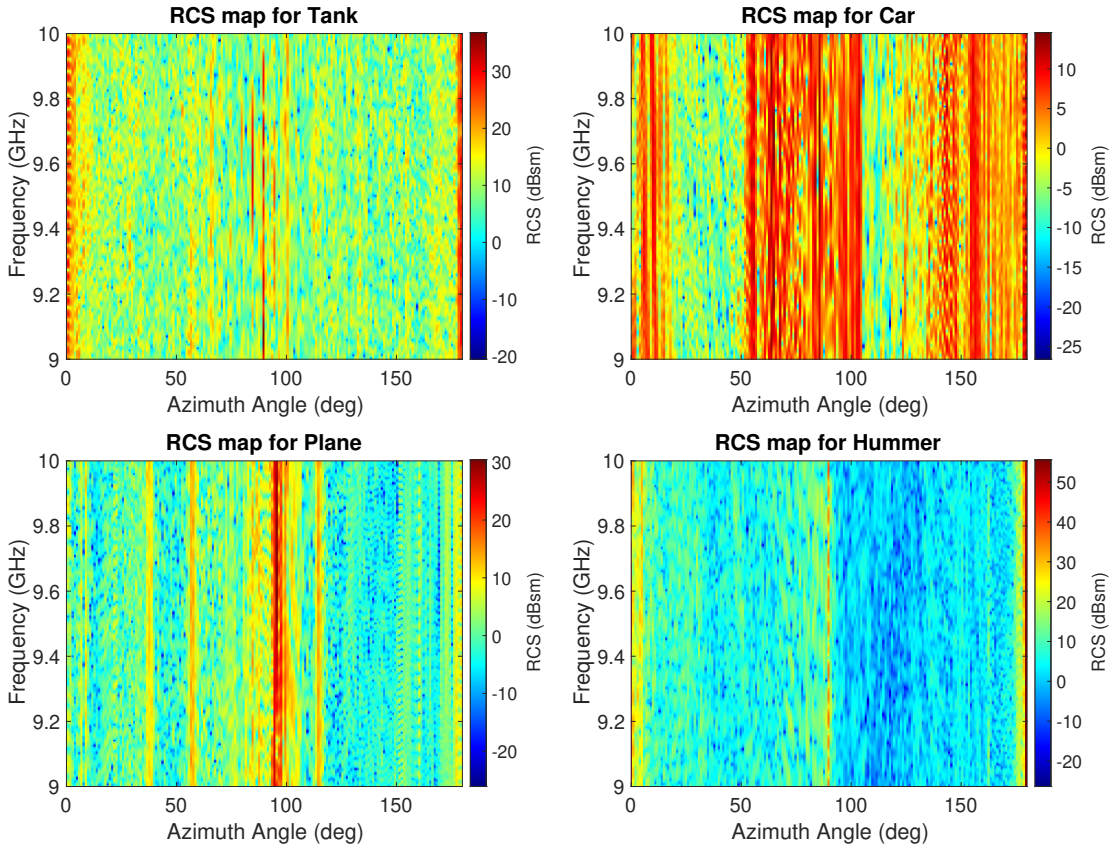


Figure 3.4: RCS profile for CAD models, all at 70 degree incident angle.

magnitude for the λ found to reduce the MSE. For each class, discrete angle, and SNR value, the optimization was reevaluated. For determining the correct class, results are plotted alongside the matched filtering-based approach (shown later in Fig. 3.9 in Section V.B). At each SNR and class, accuracy was obtained by finding the percent of correctly classified angles across all 180 degrees. This was then done for each integer SNR value and class. The filtering results will be further explained in the next subsection. For all SNR values, the sparse regression based approach showed better performance. However, as previously mentioned, there was a large cost in computational speed to achieve this accuracy. This approach could be optimized for speed by choosing a λ value a priori in each case, and solving (3.7) by a different method that does not perform a sweep over this parameter. This was not considered in this chapter, as maximizing accuracy was the goal. Another downside to this approach is that it requires a full characterization of possible

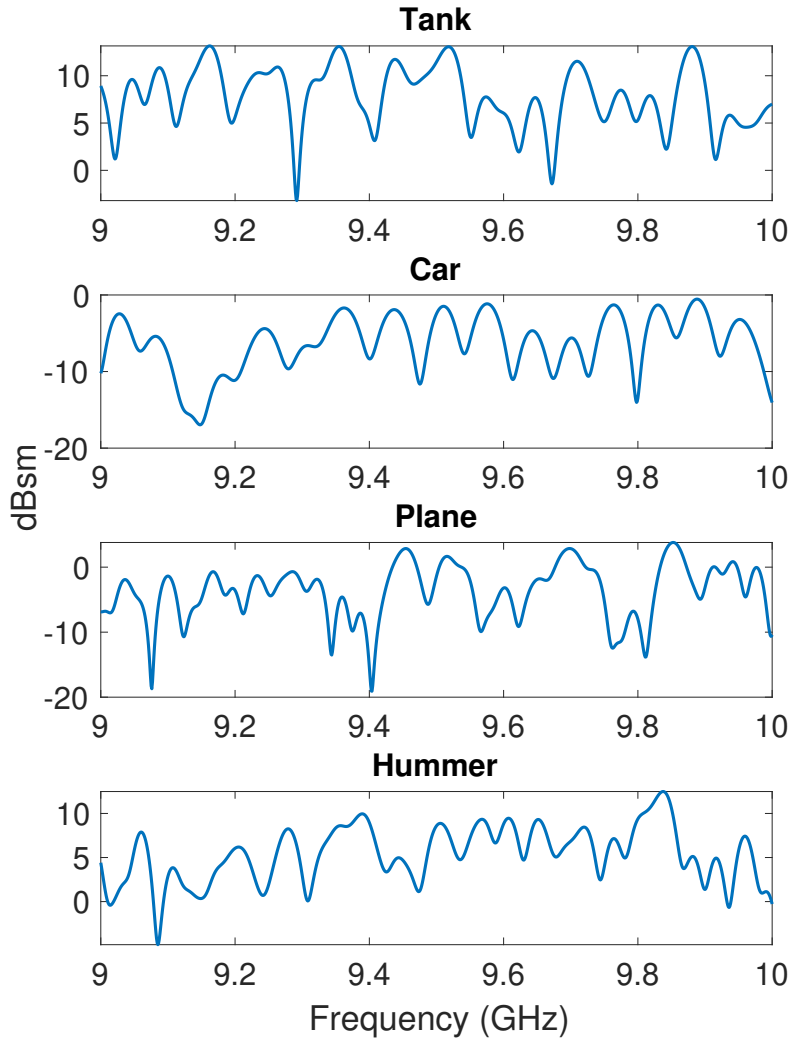


Figure 3.5: Example 45 degree cut of the RCS profiles in Fig. 3.4.

target responses. In practice, there could be variations in the relative positioning of the radar and target that would reduce accuracy.

3.5.2 Matched Filtering Based Classification

The singular value decomposition of each data matrix, in the linear scale, can provide insight into the structure of these profiles. The singular values, ordered by

descending magnitude, corresponds to how much correlation the singular vectors capture in the column space of the matrix. Also, recall that the squared magnitude of these are the eigenvalues of the correlation matrix. By looking at the singular values, an idea of how many dominantly correlated vectors needed to represent the column space of the data can be deduced. For each data matrix, the first 70 singular values are plotted in Fig. 3.6, where the y-axis is plotted in log scale. In each case, the first value is the highest, followed by a drop to the second, and a consistent decrease thereafter.

A related series of plots in Fig. 3.7 shows the cumulative total energy against all of the singular values, plotted for the same first 70. The data corresponding to the car can be the most accurately represented with the least number of modes, while the plane would need the most. Note that 90% of the total energy is captured in the first 47, 34, 63, and 48 singular values for the tank, car, plane, and hummer, respectively. To see the effect of the vectors corresponding to these singular values on the classification problem, each class is tested for different numbers of singular vectors used as filters.

The first test of the approach is testing classification accuracy from varying numbers of the matched filters, or singular vectors, across each azimuth angle. This is started with a relatively noise free environment of 30-dB SNR to verify the idea that the filters can be used for accurate classification. The simulation procedure begins with calculating the received spectral density at each aspect angle, fix the number of filters in the filter bank, compare the filtering output across all classes, and calculate the percentage of correct choices. First, for a fixed number of singular vectors, here chosen to be 25, the confusion matrix for the classification against all four classes is shown in Fig. 3.8. From this, it can be seen that the target that captured the most uncertainty in our simulations was the Hummer, and the one that captured the least was the Plane. Additional results for the multiple hypothesis testing problem against all classes can be seen in Fig. 3.10. In all aforementioned scenarios, for a fixed number of singular vector filters, the classification procedure was performed for a return from each azimuth angle. The final percent correct is the number of correctly classified angle returns for that class, or more explicitly calculated as (correctly classified angles)/(total angles). This procedure is then repeated for an increasing number of filters in Fig. 3.10. Results are shown for varying SNR values in Fig. 3.9. More discussion on this will be given later in this

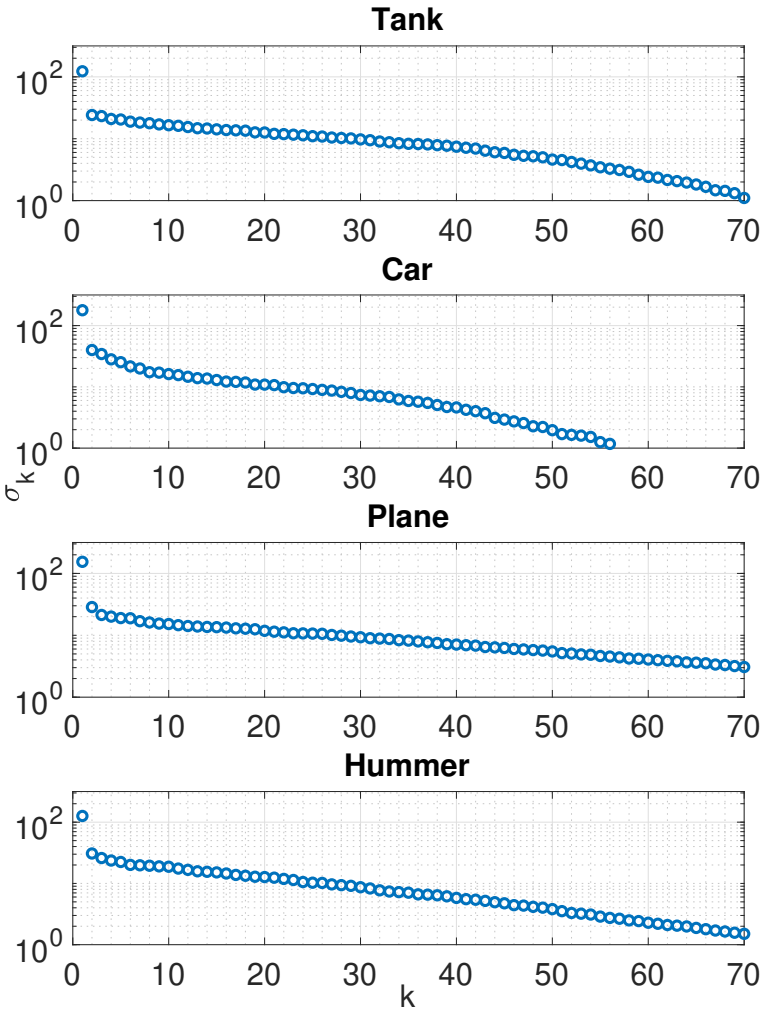


Figure 3.6: First 70 singular values of the normalized RCS data.

chapter.

These results show the general trend in the beginning that as the number of singular vectors used increases, so does the classification accuracy. However, with the exception of the car, the other targets seem to hit a point beyond which increasing the number of singular vectors actually results in decreasing classification accuracy. This can be understood by examining Figs. 3.6 and 3.7. The magnitude of the singular values eventually hit a point where they are very small relative to

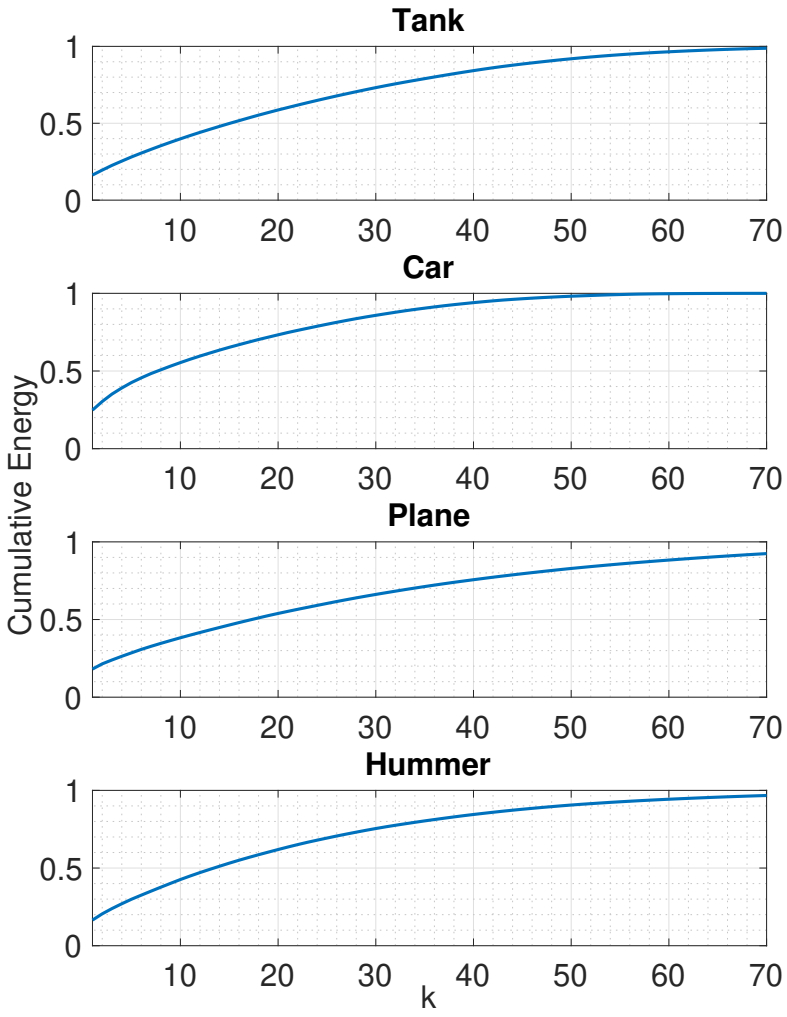


Figure 3.7: Cumulative energy contained in the first k singular values.

the earlier values, and contribute less and less to the cumulative energy contained in all values. There seems to be a point such that each new vector is comparatively minimally correlated to the column space, and would therefore have the potential to add ambiguity to the results. This phenomenon is similar to the well-known concept of diminishing marginal utility, which has origins in economics, or alternatively, the concept of information elasticity, defined for radar systems in [24]. A topic of further study could be to look into this idea more in the domain of target classification.

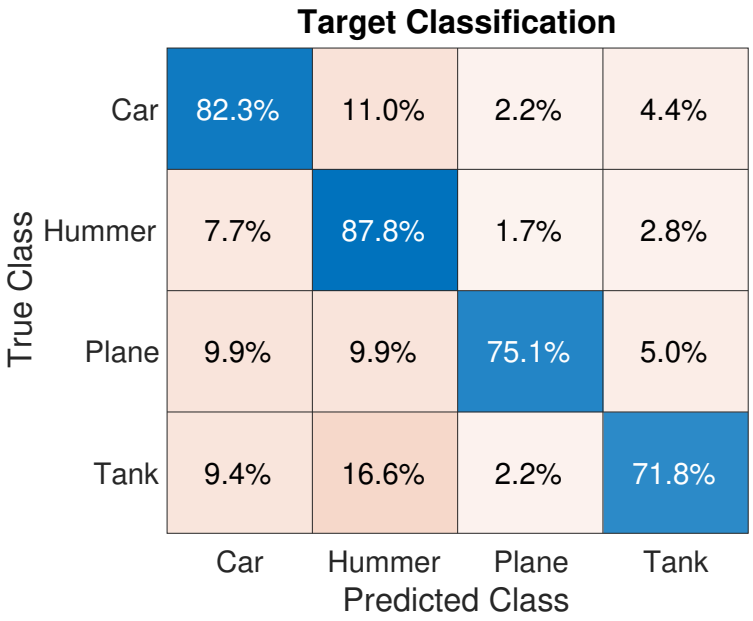


Figure 3.8: Confusion matrix of target classification against all four targets. All scenarios performed at an SNR of 30 dB, and using 25 singular vectors for the matched filtering.

Similarly, classification results were generated using 25 of the singular vectors for each class for varying SNR values, and the results are shown in Fig. 3.9. This shows an expected consistent increase in accuracy for increasing SNR. Additionally, for comparison, this plot also shows classification results using mean sectors of aspect angles. The RCS data were averaged into 25 and 45 sectors. The 25-sector data are used to match the number of singular vectors, while the 45-sector data are used to show results for a smaller section of angles. The inner product of the received signal was then taken with averaged sectors for each class, and the largest correlation was chosen. These results show that the 25 singular vector matched filters outperform the same number of averaged sectors of the data. For 45 mean sectors of the data, the 25 singular vector matched filters also performed better, but the results were closer. However, the number of filters used nearly doubled to achieve the closer, but still lesser, accuracy.

As compared to sparse regression, the filtering gave consistently less percent accuracy across all angles. That said, all filtering results were able to get to at least 70% accuracy as the SNR increased while using the same set of filters for

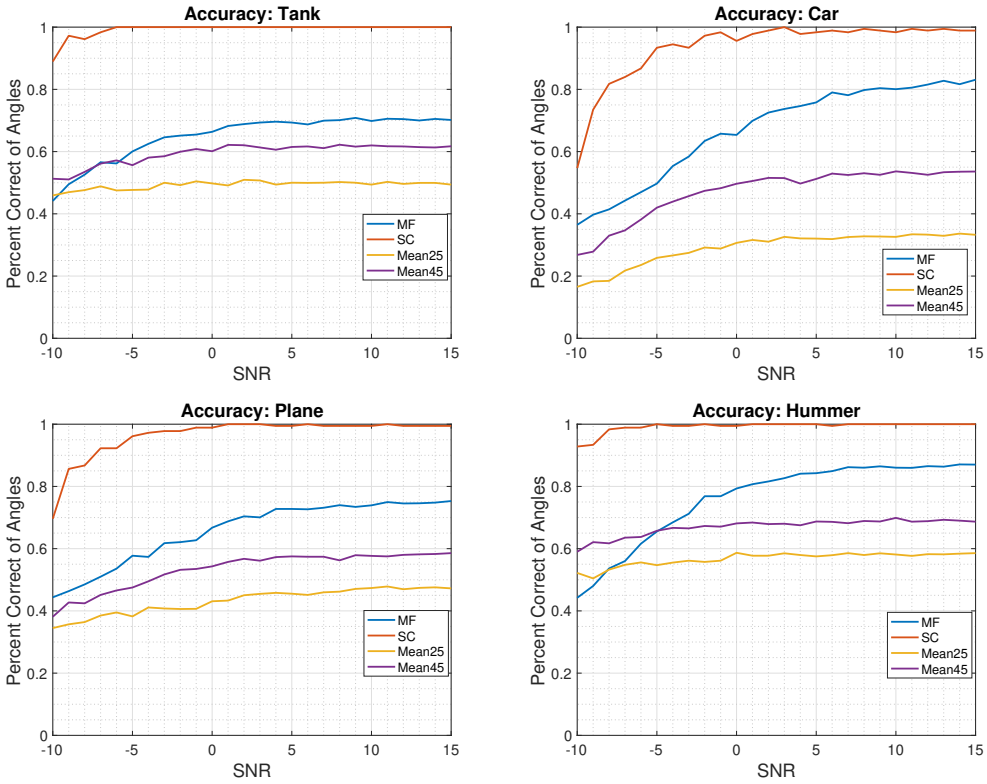


Figure 3.9: Classification accuracy against SNR using 25 singular vectors for matched filtering (MF), sparse classification (SC), and 25 and 45 mean sectors (Mean). Classification was performed against the other three targets. For each SNR value, the classification was performed over all azimuth angles, resulting in the plotted percent correct. All results are from 1 GHz bandwidth waveform and averaged over 10 runs.

every run. In contrast, the sparse regression results were obtained by rerunning the optimization each run. Additionally, the filters were all obtained by determining the dominantly correlated structures in the response dictionaries, and therefore may be more robust to imperfect target response knowledge from slight variations in radar to target look angles. Also, these results are for the multiple hypothesis testing problem, and it would be useful to also study the results of the binary case, in which only two targets are considered. Using the same problem setup as in Fig. 3.10, some examples of the binary problem are shown in Fig. 3.11. Four cases were considered, where the difference between the left and right panels is which of the targets was used as true. Results were shown up to the use of 25 singular vectors,

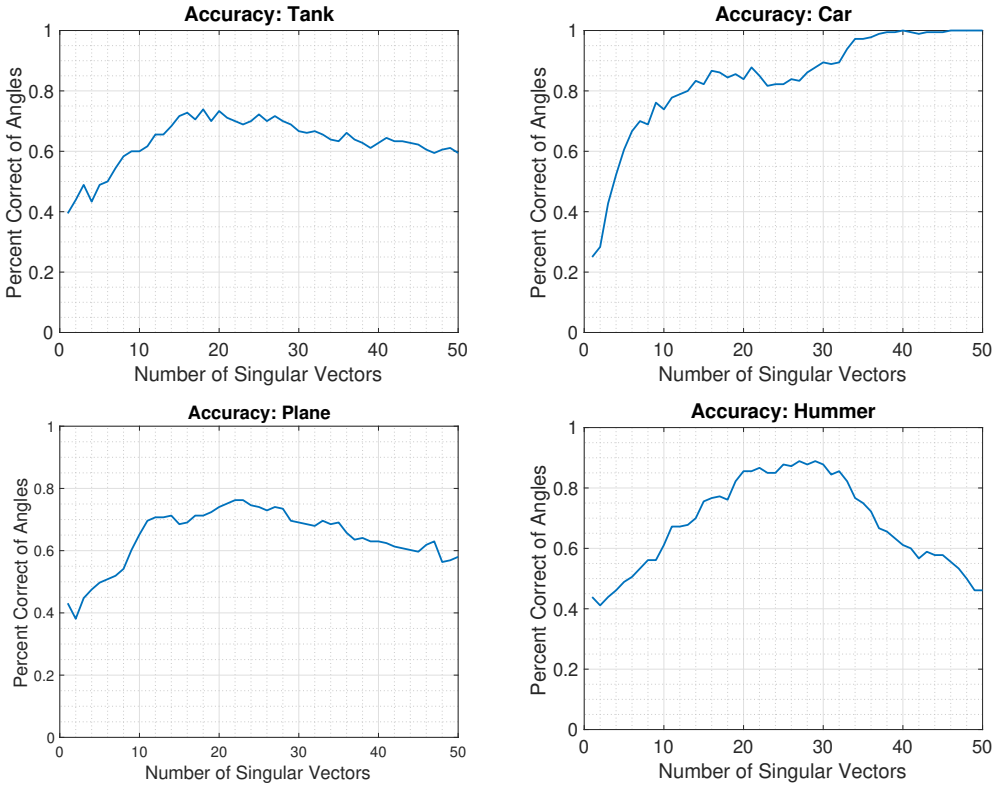


Figure 3.10: Classification accuracy at SNR of 30 dB for test over full 1 GHz bandwidth. Classification was performed against the other three targets. For each singular vector, the classification was performed over all azimuth angles, resulting in the plotted percent correct. All results are from 1 GHz bandwidth waveform.

and all cases showed increasing accuracy for increasing the number of filters.

3.6 Summary

In this chapter, a sparse regression and matched filtering scheme is introduced for radar target classification using simulated dictionaries of target response profiles from varying aspect angles. The sparse regression approach was able to classify both target class and aspect angle with great accuracy, but was computationally much slower in that it required an optimization for each pulse. Additionally, it required a return perfectly matching a response in the target dictionaries.

In contrast, the filtering scheme was designed to be aspect angle invariant by finding filters related to the response dictionaries in a way to avoid angle

dependence. This approach found a set of filters to maximize the SNR at the output of the correct target class. The classification approach would then compare outputs from the filters of each class, and choose the highest magnitude output. These filters were shown to be the eigenvectors of a correlation matrix of the dictionaries corresponding to the maximum eigenvalues. The result was a series of matched filters correlated to the column space of the dictionary matrix. Simulated results for data simulated from target CAD models showed that this could result in reasonably good classification accuracy. Additionally, as this approach only relied on correlations with the data, it is predicted to be more robust to variations in target response than the sparse regression method.

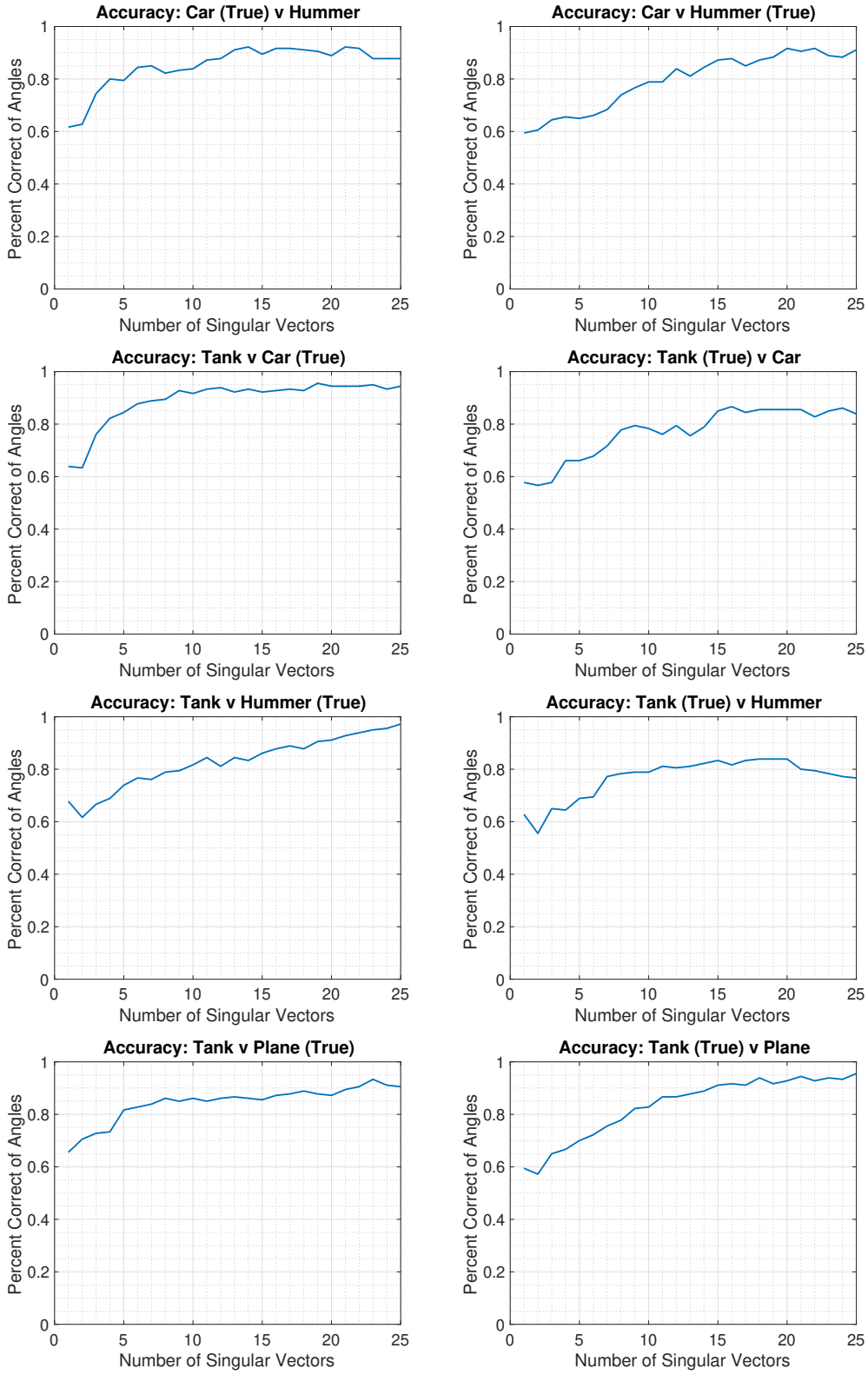


Figure 3.11: Classification at SNR of 30 dB for test. This is a classification accuracy test against one other potential target per scenario. Again, tested over all azimuth angles. All results are from 1 GHz bandwidth waveform.

CHAPTER 4

SPECTRAL SHAPING NOISE WAVEFORM FILTER

4.1 INTRODUCTION

In the study of cognitive and adaptive radar, the system is meant to take advantage of the closed loop feedback between the receiver and the transmitter to optimize performance. An important ability of a cognitive radar is to be able to adapt the transmitted waveform to varying operational conditions, radio environment knowledge, or task performance optimization. This could include parameter selection for higher level degrees of freedom, such as pulse duration, bandwidth, chirp rate, etc [7]. A step further in the cognitive design might include pulse-to-pulse design to optimize its spectral characteristics for different purposes. This could include radar functions such as matched illumination [25,26] for target identification and classification, increasing the signal to interference plus noise ratio (SINR) or signal to clutter ratio (SCR) for improved detection [6], notching to avoid in-band interference [27], or numerous other possibilities.

Noise radar technology offers a number of advantages for waveform design [28]. Some advantages include low ambiguity in both time and Doppler giving rise to a thumbtack like ambiguity function, being robust in the presence of interference and jamming, ease of wideband waveform generation, and an even distribution

of energy across the signal bandwidth. The structure of noise-based waveforms is such that every generated pulse is aperiodic and unique, which gives rise to a number of benefits. Due to the need for the transmitting radar to maintain a copy of the waveform to correlate the returns, it becomes difficult for hostile receivers or jammers to detect and disrupt operation. These characteristics make noise waveform radar a desirable candidate for tactical radar scenarios. The ability to adapt and efficiently modify the noise waveform's characteristics and degrees of freedom will further increase its usefulness and versatility.

Previously proposed was spectrum shaping by sending a wideband waveform through a perfect reconstruction filter bank, where the gain of each subband can be scaled to a desired level [29]. This chapter attempts to address shaping the in-band energy of a noise radar pulse in a more general way that is amenable to different optimal spectrum shaping methods in literature. The outline of this chapter is to first introduce the spectral shaping strategy considered in [6], then demonstrate how to derive a filter to shape a noise radar pulse to the desired spectrum. This is shown in the context of an adaptive transmit block of a cognitive radar system block diagram that can often be found in literature. In other words, the way in which the information is found and given for waveform synthesis to the adaptive transmitter will be considered a separate subsystem not explicitly considered in this chapter. Rather, the design of the adaptive transmit block would be meant to be a step towards a versatile subsystem that could be used in conjunction with various knowledge systems. This may include the cognitive radio concept of dynamic spectrum access [7], clutter maps, or other radio environment sources and models.

4.2 Optimal Spectrum Design

There have been various approaches in literature that have discussed how to approach the problem of optimizing a waveform's energy spectrum. Common metrics for optimization include maximizing the SINR of the received signal or the mutual information between the target and transmitted signal [25, 26]. Some constraints on the design often includes the total energy of the waveform, maintaining a constant modulus, or maintaining desirable autocorrelation range sidelobes. This work will be focused on using the procedure described in [6] for optimizing a waveform in the presence of additive colored noise (ACN). An outline of the approach will be given

and the design of a noise radar waveform will be shown that matches the optimal spectrum.

4.2.1 Optimizing Spectrum in Additive Colored Noise

In the presence of additive colored noise, it was shown that the optimal waveform, denoted as $\mathbf{s}_{opt}$, is the one that maximizes the energy in the whitened target echo [6]. Outlining the procedure, this was derived by first considering the returned signal, and the receiver that maximizes the SINR in the ACN case. First, the transmitted signal, $\mathbf{s}$, interacts with the target. This is represented by the target transfer matrix $\mathbf{H}_T$, which is a lower triangular matrix containing discrete samples of some complex target impulse response. This signal subsequently corrupted with ACN, a vector denoted by $\mathbf{n}$, is then reflected back to the radar . The optimal receiver for this was shown to be a whitening filter followed by a matched filter. If the total interference covariance matrix, $\mathbf{R}$, is known, then the whitening filter is given by

$$\mathbf{H}_w = \mathbf{R}^{-1/2} \tag{4.1}$$

and the received waveform, after going through the whitening filter, is a vector denoted as $\mathbf{z}$ given by

$$\mathbf{z} = \mathbf{H}_w\mathbf{H}_T\mathbf{s} + \mathbf{H}_w\mathbf{n} \tag{4.2}$$

We denote

$$\mathbf{H} \triangleq \mathbf{H}_w\mathbf{H}_T \tag{4.3}$$

The function can be defined to maximize the SINR in the receiver that explicitly depends on the transmit waveforms as

$$\max_{\{\mathbf{s}\}} |\mathbf{s}^T(\mathbf{H}^T\mathbf{H})\mathbf{s}| \tag{4.4}$$

Invoking Cauchy-Schwarz theorem and the condition for equality, as equation (4.4) can be seen as the inner product between $\mathbf{s}$ and $(\mathbf{H}^T\mathbf{H})\mathbf{s}$, the maximum can be expressed as in equation (4.5). The optimal waveform, $\mathbf{s}_{\text{opt}}$, can now be seen to be the eigenvector associated with the maximum eigenvalue, $\lambda_{\max}$, of $\mathbf{H}^T\mathbf{H}$.

$$(\mathbf{H}^T\mathbf{H})\mathbf{s}_{\text{opt}} = \lambda_{\max}\mathbf{s}_{\text{opt}} \tag{4.5}$$

This procedure could result in a rather narrowband waveform that allocates energy in a narrow region or regions where the noise energy is low. Such a narrowband signal may not be desired as this would cause negative effects on other radar performance metrics, such as the range resolution. This can be mitigated by placing a constraint on the optimized waveform which allows the waveform or energy spectral density to deviate in a predefined way, and allow for the design of a wideband waveform that will still achieve SINR gains by shaping the in-band energy [6]. The constrained optimization for the radar pulse would then have the form of equation (4.6) where some maximum deviation would be allowed.

$$\begin{aligned} &\max_{\{s\}} |\mathbf{s}^T(\mathbf{H}^T\mathbf{H})\mathbf{s}| \\ &\text{subject to} : ||\mathbf{s} - \mathbf{s}_{\text{noise}}|| \leq \delta \end{aligned} \tag{4.6}$$

Exact specification of this similarity parameter would be situation or task dependent, so a heuristic approach to achieving a desired spectrum is used where a small number of range bins are used in the calculation. Simulations have shown that an optimal spectrum obtained this way corresponds to an energy spectral density that deviates in a more modest way. Further discussion on similarity constraints in the context of linear frequency modulated waveforms can be found in [6].

An example of finding an optimum for a given realization of ACN is shown in Fig. 4.1. Colored noise was generated by sending white noise through a complex-valued tapped delay line representing the channel interference characteristics. The target transfer function was considered to be a unit impulse at the origin. The line denoted with *Opt* was calculated and will be considered the desired spectral mask. This represents the spectral distribution for which the noise radar waveform will be shaped to achieve an SINR gain.

A limitation to this procedure is the assumption that the interference covariance matrix is known. In practice, this and the channel characteristics would need to be estimated in real time, and the optimal spectrum be recalculated. For the purposes of demonstrating the shaping procedure on noise radar pulse generation, it will be assumed that the interference characteristics are known. Additionally, $\mathbf{s}_{\text{opt}}$ obtained in this manner does not make for a practical waveform, as it does not conform to a

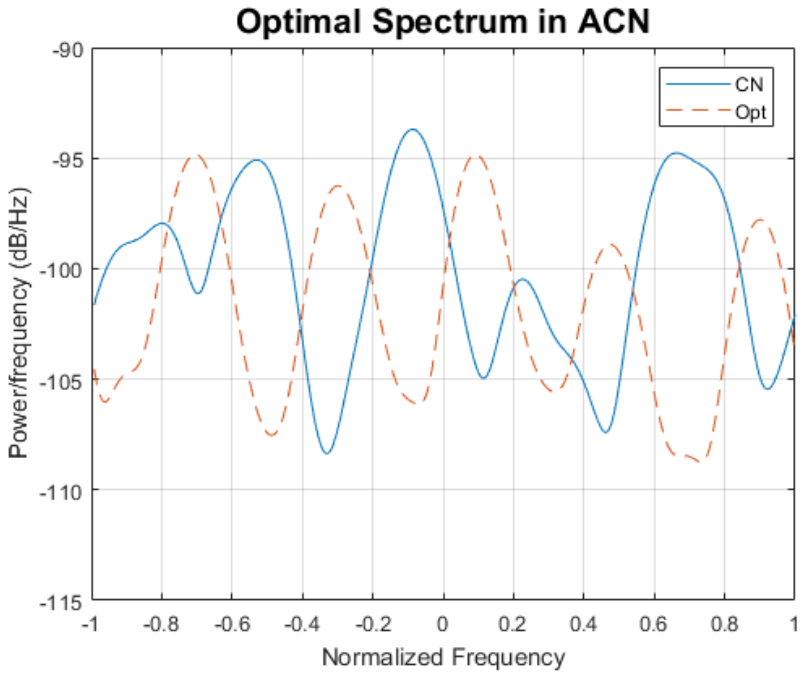


Figure 4.1: Realization of additive colored noise (ACN) and derived optimal spectrum.

constant modulus constraint. A waveform exhibiting a constant modulus in the time domain would be of practical use. An additional approach considered in [6] to achieve a constant modulus waveform is to use the method of stationary phase to model a nonlinear frequency modulated (NLFM) pulse to match the derived spectrum. This chapter's work is an alternate approach utilizing noise waveforms.

4.3 Spectral Shaping

Given an optimized spectral mask discussed in the last section, the next step involves deriving a filter to shape the noise pulse to the mask. Using concepts typically applied to estimating the parameters of a stochastic process, a filter can be derived to generate a signal with the desired energy spectral density from a white noise input. In this section, first a description of the general autoregressive moving average (ARMA) model will be given, followed by a focus to all-pole modeling with an autoregressive (AR) model. A further description of these models and the

parameter estimation can be found in [30].

4.3.1 Autoregressive Moving Average Model

In stochastic random process modeling, parametric models can be derived when given a statistical characterization of the random process. The most general form is known as an ARMA process with p poles and q zeros, denoted as ARMA(p,q). The z-domain expression for the causal linear shift invariant filter is expressed as

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q(k) z^{-k}}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \tag{4.7}$$

where the $a_p(k)$ and $b_q(k)$ are the filter coefficients to be found. These filter coefficients correspond to p poles and q zeros in the system transfer function. Given time series samples for the random process to be modeled, a statistical description can be found, the filter coefficients can be calculated, and the desired process can be generated by filtering white noise. In practice the time series data would be used to calculate an autocorrelation sequence, which would be used to calculate both the AR and MA coefficients. This chapter deviates in the approach used to estimate the autocorrelation. In many cases, all-pole modeling using just an AR model is of interest due to the relative ease with which the coefficients can be obtained by the Yule-Walker equations, and the efficient algorithm that can be used in their calculation known as the Levinson-Durbin recursion. This type of model will be of primary interest throughout the rest of this chapter, although the more general model could be used.

4.3.2 Autoregressive Model

When $q = 0$ in (4.7), the process reduces to an autoregressive process of order p. This leaves just one coefficient in the numerator, and the filter transfer function for this system becomes

$$H(z) = \frac{b_q(0)}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \tag{4.8}$$

The coefficients $a_p(k)$ can be solved using what is known as the Yule-Walker equations. These can be expressed as

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = |b(0)|^2 \delta(k) \qquad k \geq 0 \tag{4.9}$$

where the $r_x(k)$ is the statistical autocorrelation of the process at a lag of k. Using the conjugate symmetry of $r_x(k)$, the Yule-Walker equations can be written in matrix form as

$$\begin{bmatrix} r_x(0) & r_x^*(1) & r_x^*(2) & \dots & r_x^*(p-1) \\ r_x(1) & r_x(0) & r_x^*(1) & \dots & r_x^*(p-2) \\ r_x(2) & r_x(1) & r_x(0) & \dots & r_x^*(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_x(p-1) & r_x(p-2) & r_x(p-3) & \dots & r_x(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ a_p(3) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \vdots \\ r_x(p) \end{bmatrix} \tag{4.10}$$

The matrix of the p autocorrelation values is a Toeplitz matrix. Linear equations of this form involving a Toeplitz matrix allow for the use of an algorithm, known as the Levinson-Durbin recursion, to efficiently solve for the coefficients [30]. The $b(0)$ coefficient can then be solved for by setting $k = 0$ in the Yule-Walker equation as shown in (4.11).

$$|b(0)|^2 = r_x(0) + \sum_{k=1}^{p} a_p(k) r_x(k) \tag{4.11}$$

What remains in the derivation of the pulse shaping filter is to find the autocorrelation values given the desired spectral mask as discussed in the previous section.

4.3.3 Autocorrelation Estimate

The Fourier transform of an autocorrelation sequence is known to give the power spectral density of a random process. The discrete time version of this is shown in (4.12).

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-jk\omega} \tag{4.12}$$

Using this relation and the optimal spectrum previously derived, the autocorrelation can be found by taking the Inverse Fast Fourier Transform (IFFT) of the desired spectral mask. It is clear that the inverse transform would need to be defined over an infinite interval. In practice, an estimate of the autocorrelation sequence would be calculated by taking a finite number of power spectrum values. This would correspond to the autocorrelation estimate associated with the periodogram.

4.3.4 Model Order Selection

Selecting the model order p of the autoregressive model could depend on a number of factors. This could include the needed accuracy of the estimate, hardware limitations, and latency requirements. A model order could be selected a priori known to achieve desired results given knowledge of a given environment or target, or could be selected dynamically by optimizing some cost function such as a log-likelihood function that penalizes large order models. In the situation where the desired energy spectral density is known a priori, it is mentioned in [31] to choose an autoregressive model order $p > 2M$, where M is the number of peaks. In the case of the interference shown in Fig. 4.1, this will result in $p > 10$.

4.4 Waveform Model

The needed tools have been discussed in the previous section, and now what is left is to put this in the context of an adaptive transmitter block of a cognitive radar. First, it needs to be assumed that there is feedback between an intelligent receiver and the transmitter. This feedback could include channel characteristics, details from an environmental database, or any number of inputs to increase the radar's situational awareness. The information of importance in this chapter is the information needed to shape the waveform's spectrum to optimize performance at a desired task. As optimizing for SINR in the presence of additive colored noise is considered, the relevant information to be relayed to the adaptive transmit block would be the channel noise characteristics. A general form of an adaptive transmit block that is considered in this chapter can be seen in Fig. 4.2.

It is assumed that a noise waveform generator is present and fed into this block to be further processed. The first stage includes obtaining the optimal spectrum

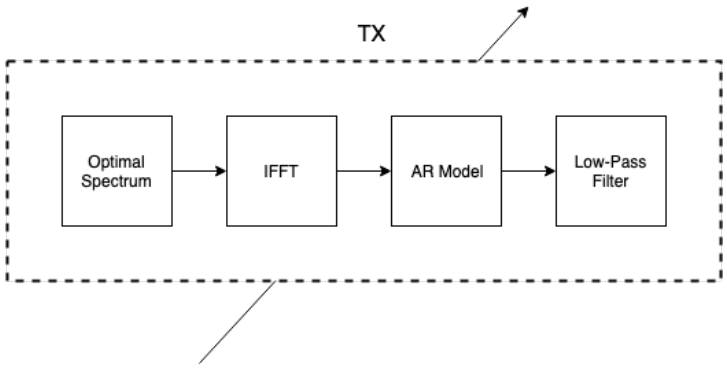


Figure 4.2: Adaptive transmitter block of adaptive noise radar pulse shaping system.

for the desired task under the necessary constraints. Next, the IFFT of the optimal spectrum is calculated to obtain an estimate of the autocorrelation sequence. The estimated autocorrelation is then used in the Yule-Walker equations to calculate the autoregressive model parameters for some model order. This filter is then applied to the input white noise to shape its spectrum. Finally, the waveform is low-pass filtered to the desired bandwidth.

4.5 Simulation

To demonstrate the performance gain that can be achieved by pulse shaping in the proposed manner, different interference scenarios were constructed where the levels of interference varies within a desired pulse bandwidth. A noise radar pulse is designed and shaped to this interference at baseband. To begin, a noise sequence is generated by a Gaussian random number generator. Using the sampling frequency and pulse duration parameters defined in Table 4.1, this results in $40,000$ samples of noise. After pulse shaping, this sequence is low-pass filtered to achieve the specified bandwidth.

The pulse is to be shaped within the presence of two different channel noise and interference scenarios, each described by some complex tapped delay line representing the colored noise (CN) channel characteristics. A different Gaussian noise sequence was generated and filtered to give the interference. In each case, there are regions within the pulse bandwidth where the interference is high, and

Table 4.1: Radar pulse parameters.

Bandwidth: B	1 GHz
Pulse Duration: τ	20 μs
Sampling Frequency: f_s	2 GHz

an SINR gain could be achieved by proper spectral shaping and energy allocation. These gains are quantified as in (4.13), where $\text{SINR}_{\text{sopt}}$ is for the shaped noise pulse, and SINR_{s} is for the unshaped low-pass filtered noise pulse.

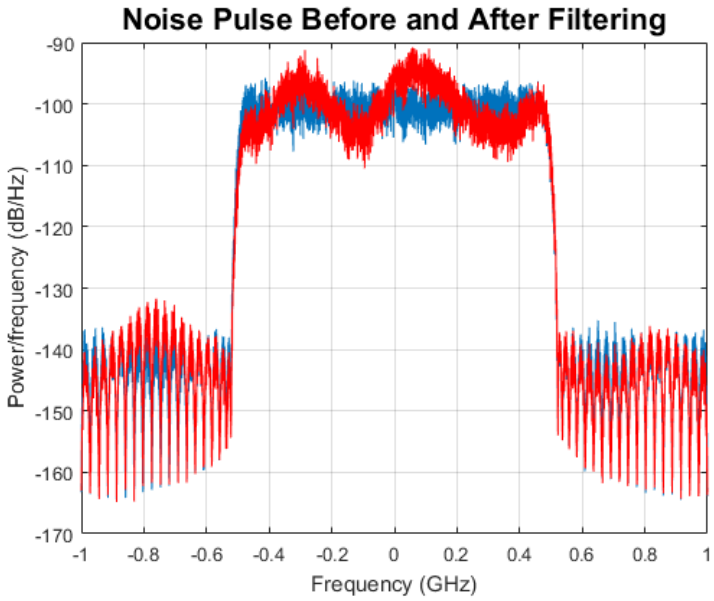
$$\text{SINR}_{\text{gain}} = \frac{\text{SINR}_{\text{sopt}}}{\text{SINR}_{\text{s}}} \tag{4.13}$$

An optimal spectrum was calculated for each scenario as described in Section 4.2. From the optimal spectral mask, the autocorrelation sequence was estimated as described in Section 4.3. These autocorrelation samples were then used to calculate the autoregressive coefficients, where the model order was fixed to the same value in each scenario. These AR coefficients were then used in the filtering of the generated waveform noise samples to achieve the desired spectrum. The final stage of the waveform design is low-pass filtering the pulse to the desired bandwidth.

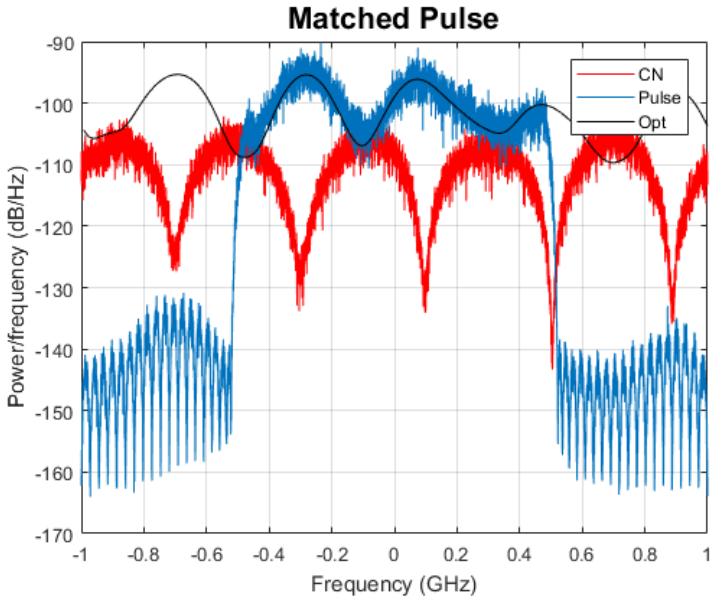
The final pulse energy spectral density is shown in Fig. 4.3 and 4.4 for each interference scenario. It is clear that the pulse conforms to the spectral mask labelled *Opt*, and allocates more energy where the noise power is low. The SINR gain relative to the unshaped noise pulse is shown in Table 4.2, where it is clear that matching gains can be achieved by pulse shaping.

Table 4.2: SINR gain for each ACN case shown in Fig. 4.3 and Fig. 4.4 respectively.

Scenario 1	1.46 dB
Scenario 2	4.91 dB

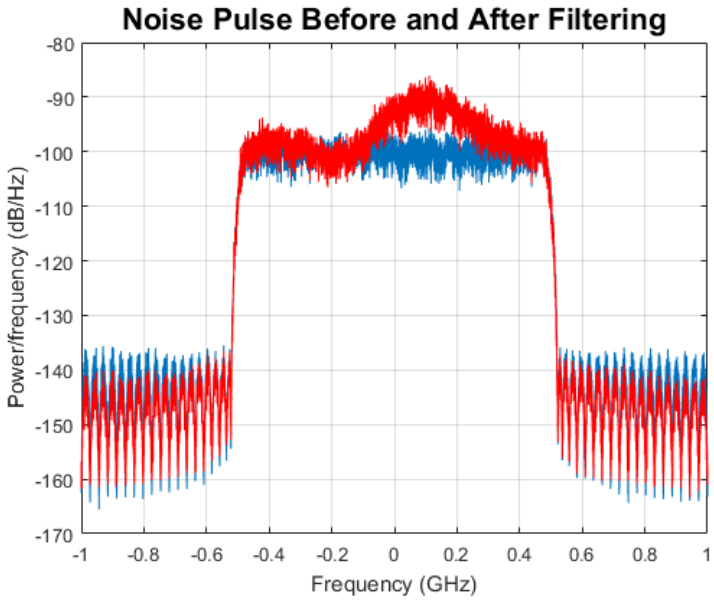


(a)

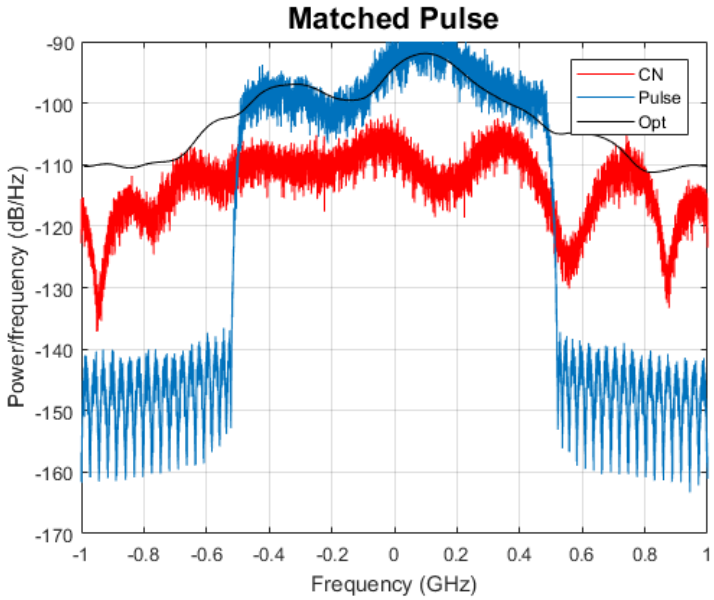


(b)

Figure 4.3: Scenario 1: (a) Overlay of waveform energy spectrum before and after pulse shaping. (b) Matched pulse to the optimal spectrum in the presence of colored noise.



(a)



(b)

Figure 4.4: Scenario 2: (a) Overlay of waveform energy spectrum before and after pulse shaping. (b) Matched pulse to the optimal spectrum in the presence of colored noise.

4.6 Summary

In cognitive and adaptive radar, being able to modify the in-band spectral energy can provide benefits in such tasks as matched illumination, increasing signal to interference plus noise ratio in additive colored noise, and notching in-band interference. Noise radar provides benefits such as LPI/LPD characteristics, wideband waveform generation, and a thumbtack-like ambiguity function. In this chapter, a method was demonstrated that can shape an otherwise flat noise radar pulse spectrum to have desired characteristics.

Future work could include providing additional constraints to control the autocorrelation and range sidelobe behavior, optimizing the algorithm for spectrum notching, and experimentally demonstrating the gains that can be provided by pulse shaping. Additionally, providing the theory for the adaptive transmitter to take advantage of other radar system degrees of freedom would allow for greater versatility. This may include bandwidth, pulse duration, pulse repetition frequency, etc.

CHAPTER 5

KOOPMAN OPERATOR AND DYNAMIC MODE DECOMPOSITION

5.1 Introduction

Nonlinear systems theory is a branch of mathematics that seeks to understand and model systems governed by nonlinear equations. Study of such systems often revolves around analyzing the orbits of the associated system. Various difficulties arise in such systems that can make estimation, prediction, and control of the process more challenging than in linear systems. An operator-based approach to such problems will be discussed in this chapter, which was first introduced in 1931 by Bernard Koopman, thus bearing his name. This concept has gained rapidly in popularity in recent years due to increasing amounts of data and new approximation methods. This chapter will first outline the basic properties and definitions of the Koopman operator and its approximation, its application to state observer design, and ultimately a framework for its application to anomaly detection in nonlinear systems.

The Koopman operator is an approach for decomposing complex systems into simpler structures via the associated spectral decomposition. Another, perhaps historically more common operator theoretic approach to dynamical systems is that of the Perron-Frobenius operator. Rather than looking at the evolution of

points, or orbits, this operator describes the time evolution of densities in phase space. In other words, it is the push-forward for density functions. It is formally related to the operator of interest in that it can be shown to be formally adjoint to the Koopman operator in appropriately defined function spaces [32]. Rather than looking at the evolution of densities, as in the Perron-Frobenius operator, the Koopman operator describes the evolution of observables of the system. In effect, it lifts the dynamics of a finite-dimensional nonlinear dynamical system into an infinite dimensional linear operator acting on measurement functions of that system.

The Koopman operator was introduced in [33] in a measure preserving setting, where it is a composition operator acting on the Hilbert space of square-integrable functions, or L^2 space. Formally, it is the pull-back operator on the space of scalar observable functions. Bernard Koopman realized that this linear transformation, now bearing his name, is unitary for Hamiltonian dynamical systems. This unitarity implies the unchanging nature of the inner product of any two measurement, or obervable, functions through application of the Koopman operator. For Hamiltonian systems, this is related to the associated phase-space volume preserving property [34]. This observation lead to work by John von Neumann and the proof for the mean ergodic theorem [35], and in Koopman and von Neumann's collaboration to introduce the notion of the spectrum of a dynamical system [36]. That is, the connection to the associated Koopman operator of a dynamical system, as well as the connection between the Koopman spectrum, and the continuous part of the spectrum associated with chaotic systems. Renewed interest in the operator began in the mid 2000's papers [37, 38] by Mezić and collaborators, where the notion of a system's Koopman modes was introduced. This led to early applications of the Koopman mode decomposition to the field of fluid dynamics [39]. The advantages were twofold, it made the connection between the state space dynamics and measurements in the physical domain clear, and, perhaps most importantly, it was completely data-driven. This same work made the connection between the Koopman operator, and its now main method of approximation known as Dynamic Mode Decomposition [40]. Since this time, many new applications, extensions, variants, and analysis have been published in this surging field of study.

As mentionded, the renewed interest of the operator in recent years has emerged from work out of Mezić and collaborators [37], where new connections and advance-

ments in data driven techniques have been made for approximating the elements of its spectral decomposition. Beyond the interesting theoretical advancements and results, this has led to practical applications of the operator in fields such as fluid mechanics [41], power systems [42], EEG [43], and a growing list of others. A recent review paper by Brunton et. al., [44], provides an excellent overview of the state of applied Koopman theory and its various branches. The paper [45] goes into the relation between the Perron-Frobenius and Koopman operators, and the practical considerations for numerical approximations of each. Applied aspects of the operator are also discussed in the recent book [34], and another recent book primarily focused on the Koopman operator in systems and controls [46].

Central to the Koopman operator is this idea of measurement functions, or observables. The power of the Koopman operator is that it is a linear operator in the space of observables although it is an infinite-dimensional operator. Practical applications would involve the problem of approximating their spectral components, i.e. eigenvalues, eigenfunctions, and Koopman modes, in some finite-dimensional invariant eigenspace spanning the usable observables. For general dynamical systems, representing these Koopman eigenfunctions remains an unsolved challenge. For a given dynamical system, the problem is to find an invariant subspace spanned by a finite set of these eigenfunctions. This results by taking advantage of the low-rank structure exhibited by some high dimensional systems. This results in a finite-dimensional approximation of the Koopman operator when limited to this subspace. The practical utility lies in this linear representation of the measurements of the possibly nonlinear dynamics, as our understanding and toolbox for handling linear systems far exceeds that of nonlinear systems. Additionally, finding linear representations that are valid far away from fixed points and periodic orbits is a goal that further complicates this problem.

The goal of this chapter is to set up a framework of using the operator for anomaly detection on nonlinear systems, and specifically with the goal of an outline for systems governed by stochastic differential equations. Anomaly detection can have a number of different definitions depending on the field of study. The focus here is on anomaly detection for signal processing of time series data [47]. This typically involves hypothesis testing on some running distance function defined on the data. For example, this could be a test on the difference between the predicted and measured state of a system, which indeed is the approach taken here. The

problem then becomes finding an appropriate predicted state. In linear systems, this can be done using a Kalman filtering framework, as this is known to be an optimal estimate in a mean-square linear sense, but can present difficulties in nonlinear systems unless restricted to small-signal perturbations from the operating point. There exists extensions to systems modeled with nonlinear dynamics, namely the Extended Kalman Filter and its different modifications, but this method relies upon linearization and is not guaranteed to converge. Additionally, one would need a model of the dynamics in this setting. Instead, by using data-driven approaches and relying on the linearity of the Koopman operator, one could use linear tools on measurement functions of the state for purposes of anomaly detection.

The outline of this chapter is to first go over the basic properties and definitions of the Koopman operator. Then, an overview of a specific approach to the finite-dimensional approximation based on a technique called dynamic mode decomposition and time delay embedding is reviewed. With this background in place, the chapter then goes into a recent extension of the Koopman operator to systems governed by random dynamical systems and specifically stochastic differential equations. Finally, the chapter finishes by getting into conclusions and later applications of these tools.

5.2 Mathematical Preliminaries and Background

This section will go over the basic background properties and definitions related the Koopman operator, and briefly touch on its connection to the Perron-Frobenius operator. This background can be found, and more thoroughly reviewed, in [32], [44], and [48]. The following are some relevant definitions and concepts which are used to describe subsequent material.

Definition 5.2.1. Let T be a linear transformation whose domain and range are contained in a complex Banach space X. The set of all λ such that the range of $(\lambda\mathbf{I} - \mathbf{T})$ is dense in X and such that $(\lambda\mathbf{I} - \mathbf{T})$ has a continuous inverse defined on its range is said to be the **resolvent set** of T, and is denoted by $\rho(\mathbf{T})$. The spectrum $\sigma(\mathbf{T})$ of an operator is the set of all complex numbers that are not in the resolvent set. This spectrum can be further classified into subsets of complex numbers λ that fail to be in the resolvent set.

The **point spectrum**, $\sigma_p(\mathbf{T})$, of a linear transformation $\mathbf{T}$ is the subset of all λ such that $(\lambda\mathbf{I} - \mathbf{T})$ is non-injective. This coincides with normal eigenvalues in a finite-dimensional spectrum.

The **continuous spectrum**, $\sigma_c(\mathbf{T})$, of a linear transformation $\mathbf{T}$ is the subset of all λ such that $(\lambda\mathbf{I} - \mathbf{T})$ has its range dense in X, is injective, and for which the inverse defined on the range is not continuous.

The **residual spectrum**, $\sigma_r(\mathbf{T})$, of a linear transformation $\mathbf{T}$ is the subset of all λ such that $(\lambda\mathbf{I} - \mathbf{T})$ is injective, but does not have its range dense in X.

Definition 5.2.2. A **dynamical system** $\{S_t\}_{t\in\mathbb{R}}$ on a phase space X is a family of transformations $S_t : X \to X, t \in \mathbb{R}$, satisfying

$S_0(x) = x$ for all $x \in X$,

$S_t(S_{t'}(x)) = S_{t+t'}(x)$ for all $x \in X$, with $t, t' \in \mathbb{R}$, and

The mapping $(t, x) \to S_t(x)$ from $\mathbb{R} \times X$ into X is continuous.

5.2.1 The Perron-Frobenius Operator

The Perron-Frobenius (PF) operator is a linear representation of nonlinear dynamics with origins in statistical physics. The PF operator evolves measures with support on the domain of the dynamical system. This is in contrast to the Koopman operator, which will be seen shortly to propogate measurement functions on the domain of the dynamics. However, the PF and Koopman operators are closely related as being formally adjoint in appropriately defined function spaces. Basic notions and definitions will be outlined here for completeness, but for a more detailed and general introduction to the topic see [32].

Definition 5.2.3. Let $(X, \mathcal{A}, \mu)$ be a measure space. A transformation $S : X \to X$ is **measurable** if

$$S^{-1}(A) \in \mathcal{A} \quad \text{for all } A \in \mathcal{A}$$

Definition 5.2.4. A measurable transformation $S : X \to X$ on a measure space $(X, \mathcal{A}, \mu)$ is **nonsingular** if $\mu(S^{-1}(A)) = 0$ for all $A \in \mathcal{A}$ such that $\mu(A) = 0$.

Definition 5.2.5. Let $(X, \mathcal{A}, \mu)$ be a measure space. If $S : X \to X$ is a nonsingular transformation the unique operator $\mathcal{P} : L^1 \to L^1$ defined by

$$\int_A \mathcal{P}f(x)d\mu = \int_{S^{-1}(A)} f(x)d\mu$$

is called the **Perron-Frobenius operator** corresponding to S.

5.2.2 The Koopman Operator

Definition 5.2.6. Let $(X, \mathcal{A}, \mu)$ be a measure space, $S : X \to X$ a nonsingular transformation, and $f \in L^\infty$. the operator $\mathcal{K} : L^\infty \to L^\infty$ defined by

$$\mathcal{K}f(x) = f(S(x))$$

is called the **Koopman operator** with respect to S.

By defining the Koopman and Perron-Frobenius operators, respectively, on L^∞ and L^1, or both on L^2, it can be shown that they are formally adjoint to one another. That is, for every $f \in L^1, g \in L^\infty$,

$$\langle \mathcal{P}^t f, g\rangle = \langle f, \mathcal{K}^t g\rangle, \text{ where } \langle f, g\rangle = \int_X \bar{f}(x)g(x)d\mu \tag{5.1}$$

Rather than advancing the state of the system, as in a flow map, the Koopman operator advances measurement functions of the state along the flow of the dynamics. The advantage of this is that often there will not be access to the direct state of the system, but instead some set of measurements on the state space.

Definition 5.2.7. A measurement function, or more commonly called *observable*, is any measurable map $g : \mathbf{M} \to \mathbb{C}$. These are typically contained in an infinite-dimensional Hilbert space given by the Lebesgue square-integrable functions on $\mathbf{M}$, though other choices of measure spaces are also valid.

First, the flow map operator advances initial conditions along the trajectory forward a time t as

$$\mathbf{x}(t) = \mathbf{F}_t(\mathbf{x}(0)), \tag{5.2}$$

Now, the Koopman operator $\mathcal{K}_t$ is defined as an infinite-dimensional linear operator that acts on measurement functions g

$$\mathcal{K}_t g = g \circ \mathbf{F}_t \tag{5.3}$$

For discrete time systems with timestep Δt, this becomes

$$\mathcal{K}_{\Delta t} g(\mathbf{x}_k) = g(\mathbf{F}_{\Delta t}(\mathbf{x}_k)) = g(\mathbf{x}_{k+1}) \tag{5.4}$$

The discrete time case will be the focus going forward, so henceforth the subscript will be dropped. The linearity of the Koopman operator is inherited from the linearity of the function spaces

$$\begin{aligned} \mathcal{K}(\alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x})) &= \alpha_1 g_1(\mathbf{F}_\mathbf{t}(\mathbf{x})) + \alpha_2 g_2(\mathbf{F}_\mathbf{t}(\mathbf{x})) \\ &= \alpha_1 g_1 \mathcal{K}(\mathbf{x}) + \alpha_2 \mathcal{K} g_2(\mathbf{x}) \end{aligned} \tag{5.5}$$

An eigenfunction $\phi(\mathbf{x})$ of the Koopman operator with eigenvalue λ satisfies

$$\phi(\mathbf{x}_{k+1}) = \mathcal{K}\phi(\mathbf{x}_k) = \lambda\phi(\mathbf{x}_k). \tag{5.6}$$

Being an infinite-dimensional operator, it may also contain a continuous or residual spectral. For the purposes of this chapter in finite-dimensional applications the point spectrum, coinciding with eigenvalues, will suffice.

5.2.3 Koopman Mode Decomposition

In cases where one has multiple measurements of a system, or in extreme cases the entire state measurement, the measurements will be arranged in a vector $\mathbf{g}(\mathbf{x})$.

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_p(\mathbf{x}) \end{bmatrix} \tag{5.7}$$

This can be expanded in terms of the Hilbert space basis constructed of Koopman eigenfunctions $\phi_j(\mathbf{x})$ as

$$g_i(\mathbf{x}) = \sum_{j=1}^{\infty} v_{ij} \phi_j(\mathbf{x}), \tag{5.8}$$

and for the vector measurement case as

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_p(\mathbf{x}) \end{bmatrix} = \sum_{j=1}^{\infty} \phi_j(\mathbf{x})\mathbf{v}_j. \tag{5.9}$$

The $\mathbf{v}$'s are known as the *Koopman modes*, where $\mathbf{v}_j$ is the j-th mode associated with eigenfunction ϕ_j. With this decomposition, the measurement dynamics are given as follows

$$\begin{aligned} \mathbf{g}(\mathbf{x}_k) = \mathcal{K}_{\Delta t}^k \mathbf{g}(\mathbf{x}_0) &= \mathcal{K}_{\Delta t}^k \sum_{j=0}^{\infty} \phi_j(\mathbf{x}_0)\mathbf{v}_j \\ &= \sum_{j=0}^{\infty} \mathcal{K}_{\Delta t}^k \phi_j(\mathbf{x}_0)\mathbf{v}_j \\ &= \sum_{j=0}^{\infty} \lambda_j^k \phi_j(\mathbf{x}_0)\mathbf{v}_j \end{aligned} \tag{5.10}$$

The triple $(\lambda_j, \phi_j, \mathbf{v}_j)$ is known as the *Koopman mode decomposition*. This was introduced in [37] by Mezic. The continuous time analog would follow in the expected way, with the Koopman modes evolving according to the complex exponentials associated with the complex Koopman eigenvalue.

5.2.4 Invariant Eigenspaces

Important is the concept of Koopman invariant eigenspaces. This allows for the approximation of the evolution on this invariant subspace spanned by a finite dimensional set of measurement functions. The invariant subspace is defined by a spanning set of functions $\{g_1, g_2, \ldots, g_p\}$, such that a function g in this subspace expressed as

$$g = \alpha_1 g_1 + \alpha_2 g_2 + \ldots + \alpha_p g_p \tag{5.11}$$

remains in the subspace after action of the Koopman operator.

$$\mathcal{K}g = \beta_1 g_1 + \beta_2 g_2 + \ldots + \beta_p g_p \tag{5.12}$$

This is important and useful in that this allows for a finite-dimensional representation of the operator, as a matrix, spanned by a finite set of functions.

A central challenge of modern Koopman theory is in identifying these subspaces and estimating the corresponding eigenfunctions. These eigenfunctions then define a space in which the system dynamics evolve linearly.

5.2.5 Koopman Spectrum Beyond Eigenvalues

For chaotic dynamical systems, the linear Koopman expansion does not hold. For such systems, the space of observables is not spanned by the eigenfunctions of the Koopman operator. It then follows that the system dynamics can not be decomposed according the the exponentials and sinusoids of the Koopman eigenvalues [35, 44]. Such cases are usually associated with a continuous spectrum of the Koopman operator. See [35, 44] for a further review of this phenomena, its characterization, and how it can be handled. For the purposes of this chapter, the point spectrum will suffice, if not exactly, then as an approximation. A finite dimensioal approximation variant of the Koopman operator suitable for some chaotic system is introduced in [49], and will be highlighted in the next section.

5.3 Finite Dimensional Approximation

A technique known as dynamic mode decomposition (DMD) was developed in the field of fluids [40] was first shown to be connected to the Koopman mode decomposition in [39]. It can be formulated as an algorithm to advance high-dimensional measurements forward, which acts to approximate the Koopman operator restricted to full state measurements. Convergence of the DMD under ergodicity assumptions of the underlying dynamical system to the spectral properties of the Koopman operator was proved in [50]. This section will first go over the basics of the algorithm. Then, an outline of a specific variation of the algorithm based on the idea of time delay observables that provides a number of advantages will be the focus.

5.3.1 Dynamic Mode Decomposition

This section will present what is known as the exact DMD [51], which provides a precise mathematical formulation in which to build upon. Given a measurement matrix $\mathbf{X}$, where the columns represent time snapshots of state measurements, a

new matrix $\mathbf{X}'$ can be constructed in which each column is some time delay of $\mathbf{X}$. Each column of this matrix could be defined as $x_j = x(t_j)$, at some time j. Then these matrices could take the form of (5.13).

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{m-1} \\ | & | & & | \end{bmatrix}, \quad \mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathbf{x}_2 & \mathbf{x}_3 & \dots & \mathbf{x}_m \\ | & | & & | \end{bmatrix} \tag{5.13}$$

The goal of dynamic mode decomposition is then to find an $\mathbf{A}$, such that the dynamics of the time evolution can be modeled as in (5.14).

$$\mathbf{X}' \approx \mathbf{A}\mathbf{X} \tag{5.14}$$

The best fit linear approximation of $\mathbf{A}$ is defined as

$$\begin{aligned} \mathbf{A} &= \arg\min_{\mathbf{A}} \|\mathbf{X}' - \mathbf{A}\mathbf{X}\|_F \\ &= \mathbf{X}'\mathbf{X}^{\dagger} \end{aligned} \tag{5.15}$$

where the symbol, $\dagger$, denotes the Moore-Penrose pseudo inverse, and $\|\cdot\|_F$ is the Frobenius norm. The solution shown in (5.14) is the exact, but would be inefficient in practice as it could depend on very large, high dimensional data. The structure of the $n \times m$ data matrices $\mathbf{X}$ and $\mathbf{X}'$ is such that they often have far more rows than columns, i.e. $m \ll n$. This will result in at most m nonzero eigenvalues and non-trivial eigenvectors. The effective rank of the data matrices and the corresponding operator $\mathbf{A}$ will in practice often be even lower, i.e. some rank $r < m$. The singular value decomposition allows for a rank r truncation of the data, where r encompasses the dominant modes of the system. Literature also suggests that this method is more robust to noise and numerical errors [34]. For notation, $\mathbf{\Phi}$ denotes the DMD modes, and $\mathbf{X}^*$ denotes the conjugate transpose of the matrix. Starting with the singular value decomposition of $\mathbf{X}$ and rank r truncation, the problem is then expressed as

$$\begin{aligned} \mathbf{X} &= \mathbf{U_r}\mathbf{\Sigma_r}\mathbf{V_r}^* \\ \mathbf{A} &= \mathbf{X}'\mathbf{X}^{\dagger} = \mathbf{X}'\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^* \end{aligned} \tag{5.16}$$

Since the algorithm is interested in a low rank embedding space, taking the matrix $\mathbf{A}$, which is some large $n \times n$ matrix, this can be reduced down to some lower

rank $r \times r$, where $r << n$. This matrix will be denoted by $\tilde{A}$. This is beneficial because the reduced matrix has the same nonzero eigenvalues as the full-sized data matrix [34]. Working with this space in the singular value decomposition becomes

$$\tilde{\mathbf{A}} = \mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{U}^* \mathbf{X}' \mathbf{V} \mathbf{\Sigma}^{-1} \tag{5.17}$$

A spectral decomposition of $\mathbf{A}$ can be computed from the much smaller $\tilde{\mathbf{A}}$ as in (5.18).

$$\tilde{\mathbf{A}} \mathbf{W} = \mathbf{W} \mathbf{\Lambda}. \tag{5.18}$$

The columns of $\mathbf{W}$ are the eigenvectors of $\tilde{\mathbf{A}}$, and the diagonal matrix $\mathbf{\Lambda}$ are the DMD eigenvalues. The high-dimensional DMD modes, denoted with $\mathbf{\Phi}$, are the eigenvectors of the full matrix $\mathbf{A}$. These can be reconstructed using this decomposition and the original shifted data matrix by

$$\mathbf{\Phi} = \mathbf{X}' \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{W} \tag{5.19}$$

There have been many improvements and variations of this algorithm in literature, such as taking advantage of work and principles in sparse representations, algorithmic accelerations, multiresolution analysis, and the use of delay coordinates to name a few. See [34, 44] and references therein for more discussion on the varied work in this field. Of particular interest is the use of delay measurements, which can provide advantages in many practical settings. This will be the focus of the next section specifically in the context of the Koopman operator.

5.3.1.1 Exact DMD

In summary, this above formulation outlines what is known as exact DMD. The algorithm will be outiined here. Given a measurement matrix $\mathbf{X}$, where the columns represent time snapshots of state measurements, a new matrix $\mathbf{X}'$ can be constructed in which each column is some time delay of $\mathbf{X}$. Each column of this matrix could be defined as $x_j = x(t_j)$, at some time j.

1. Take the SVD of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*$$

with rank-r truncation

$$\mathbf{X} \approx \tilde{\mathbf{U}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{V}}^*$$

2. Obtain the $r \times r$ matrix $\tilde{\mathbf{A}}$ by projecting $\mathbf{A}$ onto $\tilde{\mathbf{U}}$

$$\tilde{\mathbf{A}} = \tilde{\mathbf{U}}^*\mathbf{A}\tilde{\mathbf{U}} = \tilde{\mathbf{U}}^*\mathbf{X}'\tilde{\mathbf{V}}\tilde{\mathbf{\Sigma}}^{-1}$$

3. Compute the eigendecomposition of $\tilde{\mathbf{A}}$

$$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda}$$

4. Reconstruct full dimensional eigenvectors of $\mathbf{A}$, given by the columns of $\mathbf{\Phi}$. These are the DMD modes.

$$\mathbf{\Phi} = \mathbf{X}'\tilde{\mathbf{V}}\tilde{\mathbf{\Sigma}}^{-1}\mathbf{W}$$

5.3.1.2 Spectral Decomposition and DMD

Once the DMD modes and eigenvalues are computed, it is possible to represent the system state in terms of the DMD expansion:

$$\mathbf{x}_k = \sum_{j=1}^{r} \phi_j \lambda_j^{k-1} b_j = \mathbf{\Phi}\mathbf{\Lambda}^{k-1}\mathbf{b} \tag{5.20}$$

where ϕ_j are eigenvectors of $\mathbf{A}$ (DMD modes), λ_j are eigenvalues of $\mathbf{A}$ (DMD eigenvalues), and b_j are the mode amplitudes. The amplitudes in $\mathbf{b}$ are given by

$$\mathbf{b} = \mathbf{\Phi}^{\dagger}\mathbf{x}_1$$

A continuous time version of the spectral expansion is done by first converting the eigenvalues to their continuous version. This is done using $\omega = \log(\lambda)/\Delta t$, and with the following:

$$\mathbf{x}(t) = \sum_{j=1}^{r} \phi_j e^{\omega_j t} b_j = \mathbf{\Phi}\exp(\mathbf{\Omega}t)\mathbf{b} \tag{5.21}$$

5.3.2 Hankel DMD and Time-Delayed Observables

Application of exact DMD to approximating the Koopman operator does not typically lead to good results, as many real systems are nonlinear, and the linear operator estimated by DMD will not capture this effectively [52]. The quality of the approximation using DMD to approximate the Koopman operator depends heavily on the choice of observables used in the algorithm. In exact DMD, the algorithm was shown using the exact state of the system. One extension proposed for application to the Koopman operator is known as the extended DMD [53]. This involves including nonlinear function of the state variables in the data matrices, and finding the linear operator that approximates the time evolution. This requires a prior selection of the Koopman-invariant subspace, i.e. observables, for effective operation. To date there does not appear to be a systematic way to select these measurement functions. A variant of DMD that provides a more general framework for obtaining effective functions is based on time-delay embedding.

Delay embedding has had application in nonlinear systems theory for some time due to the seminal embedding theorem by Takens [54]. This establishes the conditions in which the attractor of the original system can be reconstructed using a lower dimensional measurement of a higher dimensional system. In many situations having access to the full-state of a system is not practical, and characterization would need to be performed with incomplete measurements. Instead, by taking a single, or possibly a few, measurements functions and augmenting it with its past history of some time lag can result in a new observable. The size of the embedding dimension can be estimated by taking the SVD of a Hankel matrix of the measurement trajectory data as in (5.22), and truncating from some threshold of the singular values.

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{x}_1) & g(\mathbf{x}_2) & \dots & g(\mathbf{x}_M) \\ g(\mathbf{x}_2) & g(\mathbf{x}_3) & \dots & g(\mathbf{x}_{M+1}) \\ \vdots & \vdots & \ddots & \vdots \\ g(\mathbf{x}_N) & g(\mathbf{x}_{N+1}) & \dots & g(\mathbf{x}_{N+M-1}) \end{bmatrix} \tag{5.22}$$

The Hankel matrix can be connected to the Koopman operator by rewriting it using the properties of it describing the time evolution of the observables:

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{x}_1) & \mathcal{K}g(\mathbf{x}_1) & \dots & \mathcal{K}^{M-1}g(\mathbf{x}_1) \\ \mathcal{K}g(\mathbf{x}_1) & \mathcal{K}^2g(\mathbf{x}_1) & \dots & \mathcal{K}^{M}g(\mathbf{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{K}^{N-1}g(\mathbf{x}_1) & \mathcal{K}^{N}g(\mathbf{x}_1) & \dots & \mathcal{K}^{M+N-2}g(\mathbf{x}_1) \end{bmatrix} \tag{5.23}$$

The connection between Koopman theory and chaotic systems was studied in [49], with the approach being referred to as the Hankel Alternative View of Koopman (HAVOK) analysis. The properties of these HAVOK models were studied in [50] for ergodic systems, and under these conditions these models converge to the true Koopman eigenvalues and eigenfunctions. An extension of this method to convolutional coordinates was shown in [52].

5.3.2.1 Example: Lorenz System

A primary example of a nonlinear system exhibiting chaos in literature is that of the Lorenz system [55]. The Lorenz system is given by the three coupled nonlinear differential equations

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= xy - \beta z, \end{aligned} \tag{5.24}$$

and with original parameter values $\sigma = 10, \rho = 28,$ and $\beta = 8/3$. The Lorenz attractor has been proven to have the mixing property [56], which is stronger than ergodicity [57]. This has been studied in a number of papers related to DMD, including [49] which will be outlined in the next section, but the simulations of this dissertation will follow the sampling structure and Hankel matrix construction of [50]. The observable will be the x variable, sampled every 0.01 seconds, and formed into a tall $m \times n$ Hankel matrix with $m = 10000$ and $n = 500$. The resulting attractor and a snapshot of the x-component can be seen in Fig. 5.1.

In this case, after applying the SVD to the Hankel matrix, the left singular vectors will provide an orthonormal basis spanning the column space of the matrix, the singular values will provide the weighted contribution of the basis elements, and the right singular vectors will be the principal coordinates. The first six left singular vectors are shown as a colormap on the attractor in Fig. 5.2. The associated right

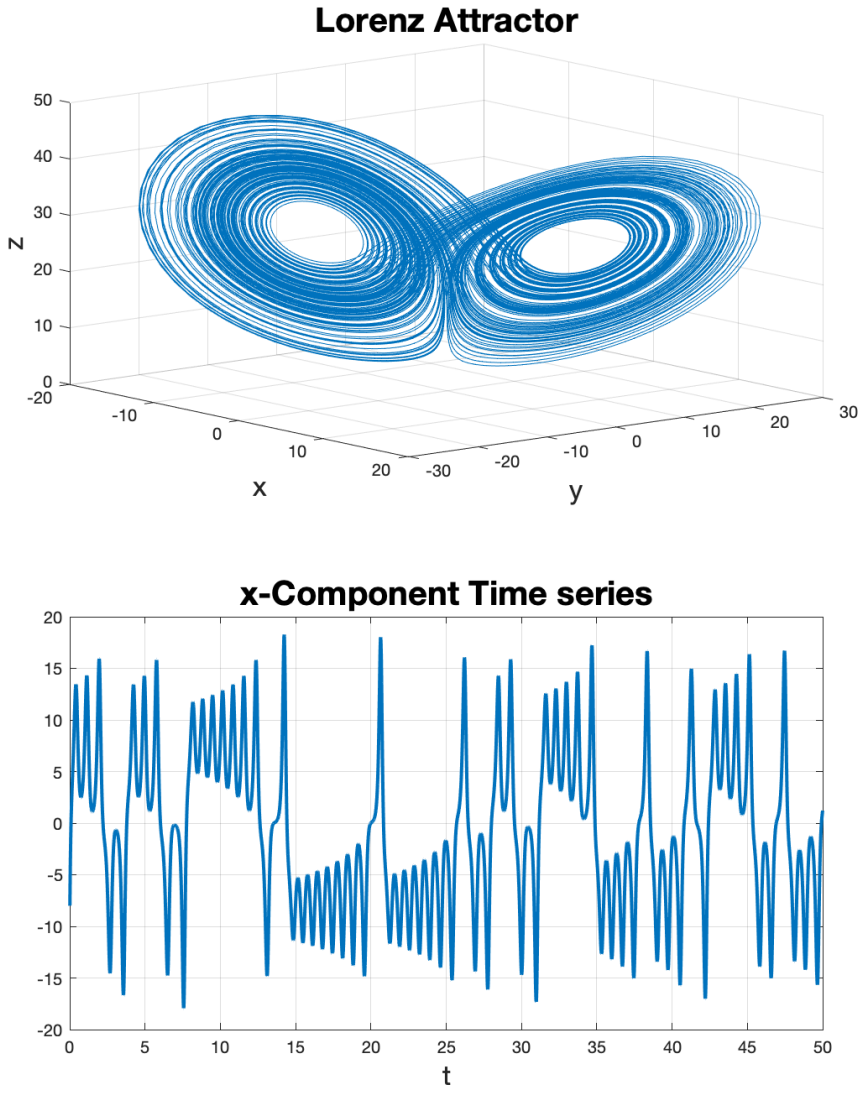


Figure 5.1: Lorenz attractor and x component time series for parameters $\sigma = 10, \rho = 28,$ and $\beta = 8/3$.

singular vectors are shown in Fig. 5.3.

The sinusoids in the principal coordinates in Fig. 5.3 show the linearity of their evolution. Despite this, there is no finite-dimensional linear system that captures this evolution. It is stated in [50] that the reason for this is that for mixing

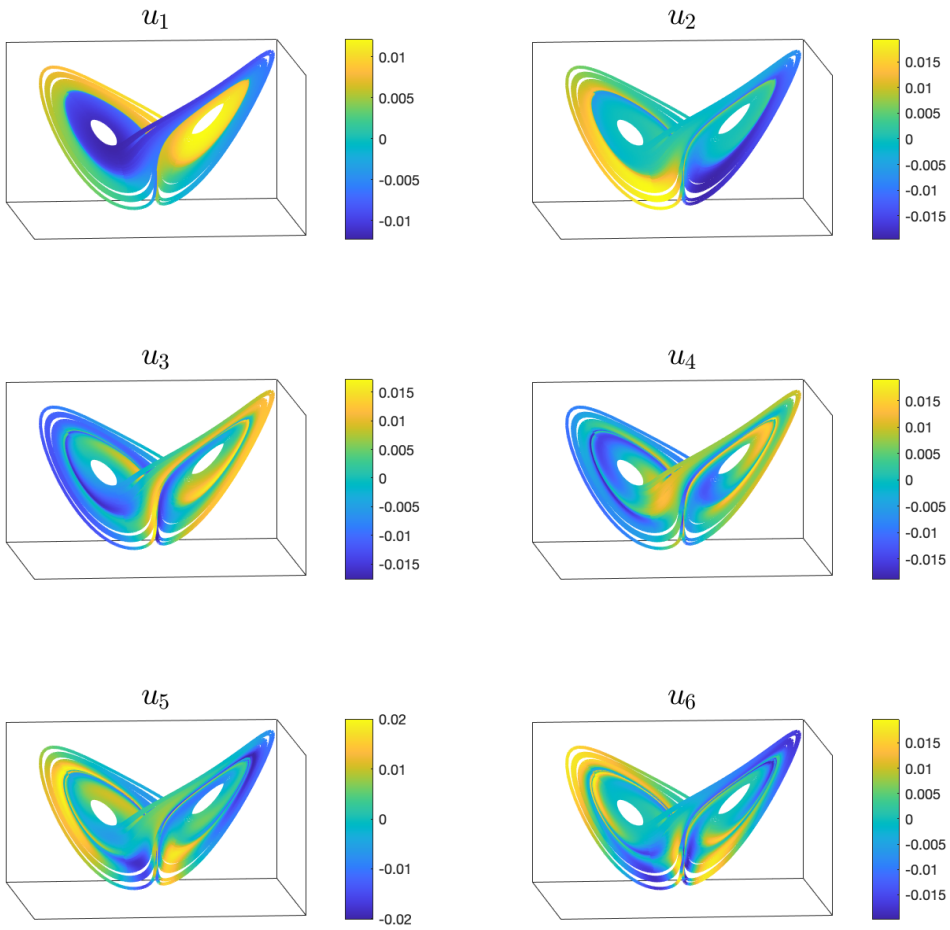


Figure 5.2: Colormap of left singular vectors in the Lorenz system.

attractors, such as Lorenz, there is only one discrete eigenvalue at $\lambda = 1$, which is associated with an almost everywhere constant eigenfunction on the attractor. For this reason, the Koopman operator cannot have a finite-dimensional invariant subspace other than one that is spanned by these constant functions. A new approach, outlined in the next section, proposes a method of creating a linear representation of the Lorenz system and several other chaotic systems via a low dimensional linear system with intermittent forcing.

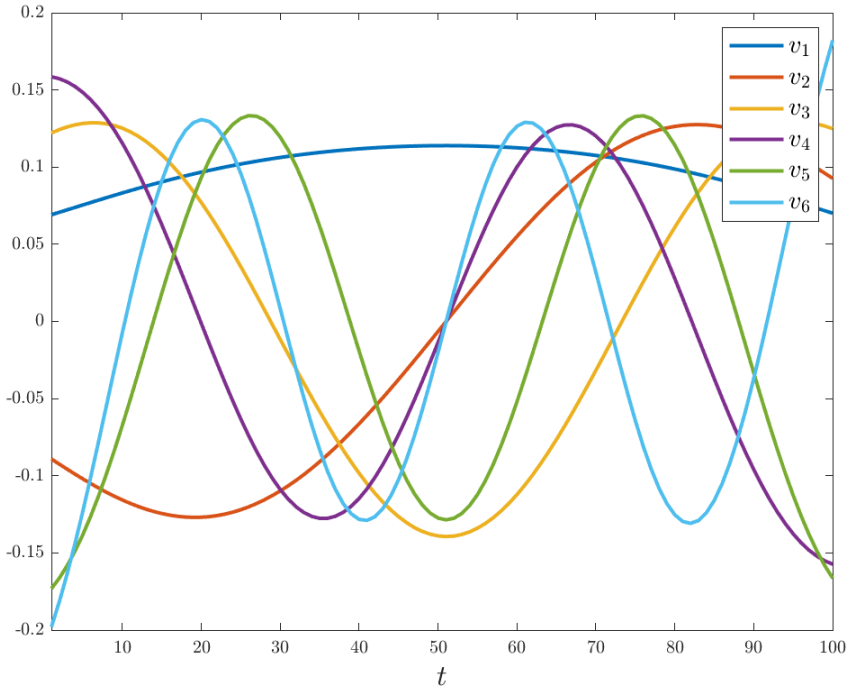


Figure 5.3: Right singular vectors of Hankel DMD.

5.4 Stochastic Koopman Operator

Extending and generalizing the Koopman operator framework to random dynamical systems and stochastic systems was discussed in [58]. Under the definitions and terminology in [59], and used in [58], the authors defined the stochastic Koopman operator for different classes of discrete and continuous-time random dynamical systems, and explored some of its spectral properties, its generators, and touched on its numerical approximation. A recent follow up paper on the robust approximation of the stochastic Koopman operator [60] introduces a Krylov subspace method (Hankel DMD) adapted to these random systems.

This section first introduces the formal definitions and properties of random dynamical systems and the stochastic Koopman operator. Next, specifically go over the relevant results related to systems governed by stochastic differential equations. These will be the specific systems of interest going forward. The next section will outline the Hankel based dynamic mode decomposition algorithm using time delay

observables for the finite dimensional approximation of the Koopman operator, and its random extension from [60].

5.4.1 Random Dynamical Systems

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathbb{T}$ a semigroup (which can be thought of as time). Suppose that $\theta = (\theta(t))_{t\in\mathbb{T}}$ is a group of measurable transformations from the probability space preserving the measure P, such that the map $(t, \omega) \to \theta(t)\omega$ is measurable. The quadruple $(\Omega, \mathcal{F}, P, (\theta(t))_{t\in\mathbb{T}})$ is what is termed a metric dynamical system. A random dynamical system (RDS) on the measurable space is defined as a measurable map $T : \mathbb{T} \times \Omega \times M \to M$ that satisfies certain cocycle properties related to $\theta(\cdot)$. See [59, Ch. 1] for more details on this construction.

Definition 5.4.1. The stochastic Koopman operator $\mathcal{K}^t$ associated with the RDS $T^t_\omega := T(\omega, t, \cdot) : M \to M$ is defined on a space of functions (observables) $f : M \to \mathbb{C}$ for which the functional

$$\mathcal{K}^t f(x) = \mathbb{E}[f \circ T^t_\omega(x)] = \int_\Omega f \circ T^t_\omega(x) dP, \quad \mathbf{x} \in M \tag{5.25}$$

exists. The family of operators $(\mathcal{K}^t)_{t\in\mathbb{T}}$ are referred to as the stochastic Koopman operator family.

Definition 5.4.2. The observables $\phi^t : M \to \mathbb{C}$ that satisfy equation

$$\mathcal{K}^t \phi^t(\mathbf{X}) = \lambda^S(t)\phi^t(\mathbf{x}) \tag{5.26}$$

are called the eigenfunctions of the stochastic Koopman operator, while the associated values $\lambda^S(t)$ are called the stochastic Koopman eigenvalues.

5.4.1.1 RDS Generated by Stochastic Differential Equations

The specific RDS of interest are those generated by stochastic differential equations. Reasons for this will be for anomaly detection for certain systems that will be further discussed in future work. Again, notationally as defined in [58], let $\mathbb{T} = \mathbb{R}^+$ and $M = \mathbb{R}^d$. Suppose that the stochastic process $X_t(\omega), t \in [0, T], \omega \in \Omega$ is obtained as a solution of the nonautonomous SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{5.27}$$

where b and σ are L^2 measurable. This has the solution, as defined in terms of the Itô integral [61] as

$$X_t(\omega) = X_0(\omega) + \int_{t_0}^{t} b(s, X_s(\omega))ds + \int_{t_0}^{t} \sigma(s, X_s(\omega))dB_s \tag{5.28}$$

Here, B_t denotes an r-dimensional Wiener process,$B_t = (B_t^1, \ldots, B_t^r)$, or Brownian motion process with the standard properties , i.e., $\mathbb{E}(B_t^i) = 0, \mathbb{E}(B_t^i B_s^j) = \min\{t, s\}\delta_{ij}$. The probability space considered for the process is identified with $\Omega = \mathcal{C}_0(\mathbb{R}^+, \mathbb{R}^r)$, which they define as the space of continuous functions satisfying $\omega(t_0) = 0$. The transformations $\theta(t)$ defined by the Wiener shifts

$$\theta(t)\omega(\tau) = W_{t+\tau}(\omega) - W_t(\omega), \quad \omega \in \Omega, \tag{5.29}$$

$(\Omega, \mathcal{F}, \mathbb{P}, (\theta(t))_{t\in\mathbb{T}})$ is said to become a metric dynamical system. See [59, App. A] for more information on this. It is a driving dynamical system for two parameter random dynamical system $\phi(t, t_0, \omega)$, for initial condition $X_{t_0}(\omega) = \mathbf{x}$, given by

$$\phi(t, t_0, \omega)\mathbf{x} = X_t(\omega) \tag{5.30}$$

The Koopman operator, $\mathcal{K}^{t,t_0}$, for this system is defined by

$$\mathcal{K}^{t,t_0} f(\mathbf{x}) = \mathbb{E}[f(\phi(t, t_0, \omega)\mathbf{x}] \tag{5.31}$$

With the two parameter family of Koopman operators, the eigenfunctions $\phi^{t,t_0} : M \to \mathbb{C}$ and eigenvalues $\lambda^S(t, t_0)$ of the Koopman operator defined on a finite-time interval satisfy

$$\mathcal{K}^{t,t_0}\phi^{t,t_0}(\mathbf{x}) = \lambda^S(t, t_0)\phi^{t,t_0}(\mathbf{x}) \tag{5.32}$$

Further closed form analysis in [58] was performed for SDEs characterized by just additive noise, and also for ones with multiplicative noise. It remains to understand the numerical approximation of the stochastic Koopman operator in this case. For this, a follow up paper [60], and references therein, discuss the state of this problem and proposes an approach to the handling of stochastic

systems, and these same systems with noisy observables. Convergence properties were studied for DMD algorithms applied to random dynamical systems, and the new DMD algorithm was shown to converge to the spectrum of the stochastic Koopman operator for both random dynamics and noisy observables. Importantly, the algorithm was specialized to the case of time-delay observables, and it was shown that a single observable's trajectory could generate an approximation of the stochastic Koopman operator. As mentioned previously, this alleviates the need to choose a basis through functions of the measurements, as well as being suited to the case of limited measurement capability. Despite the recent published works studying the application of Koopman operator theory and DMD to SDE models, there are still many open questions for best practices in this case. This dissertation presents initial results using time-delay observables and exact DMD. Future work will include a more in depth study to optimizing the application of these techniques to the relevant SDE models.

CHAPTER 6

DATA-DRIVEN MODELING OF SEA CLUTTER STOCHASTIC DIFFERENTIAL EQUATIONS

6.1 Introduction

Radar object detection in the presence of clutter involves signal processing and algorithm design that can account for that additional background backscatter. For a moving target in the presence of a near stationary scene, successive observations and filtering schemes can be used to try and extract the target of interest. This is because the Doppler spectrum of the target can be effectively separated from the background. When considering radar observation in the presence of sea clutter, the difficulty is increased by the background clutter also being a dynamical system. Trying to detect a small floating target in this environment necessitates some kind of accounting for these dynamics beyond common clutter algorithms for successfully completing the radar task. This is because it is known that sea clutter typically decorrelates over relatively short time scales. Therefore, comparison between successive measurements becomes an issue due to the changes temporally and spatially. The objective of this chapter is to take data generated from stochastic differential equation (SDE) models of sea clutter, and apply recent advances is

data driven modeling of dynamical systems to generate a model of the dynamical background from measurements. An approach for taking advantage of this model for target detection through anomaly detection is also presented.

Statistical models are commonly used to describe sea clutter. These provide a probability density model of sea clutter that describes the amplitude distribution of the returns. This is a limitation as it does not take any time dependency into account, with the exception of possibly incorporating decorrelation times. The nature of the sea surface, namely the short capillary and wind waves, and longer gravity waves, was proposed to be captured by a compound K-distribution [7,62,63]. These experimental studies showed that over shorter time periods (order of a few hundred milliseconds) the surface amplitude can be fit to a Rayleigh distribution, and over a longer time period can be fit to a Chi, or root-gamma, distribution. Then, the compound K distribution is modeled as a product of a Rayleigh and root-gamma distributed terms. Though useful, this result still does not effectively model the temporal characteristics of what was shown to be a nonlinear dynamical system [7]. However, there does exist a physically motivated derivation that gives a dynamical system model that accounts both for this distribution behavior, and time-varying characteristics.

A model that can take this time dependency into account comes from Field's model of scattering from random media [64]. This approach begins by assuming that the received signal is a sum of contributions from dynamic scatterers of a random population. Taking the limit for very large populations of scatterers results in stochastic differential equation models of the various observables of the system. This approach has been shown to effectively model K-distributed scattering, while also modeling the temporal dynamics of the process. Additionally, this approach has been shown to be experimentally valid [65]. This model involves certain parameters that would need to be estimated to apply it to a real scenario. While work on this has been proposed, such as in [66], a different and perhaps more general approach is taken in this chapter. Instead of estimating the parameters of some equation of an underlying process, a model of the dynamics can instead be built directly from measurements of the system. Often in recent literature, machine learning techniques are used for modeling systems. Inherent in these methods is the need for large amounts of prior data and computation to train a system. Though such methods have been proposed, training a model for application on a constantly

changing sea state could prove to be challenging. Instead, the approach is to look at other data driven methods, that have also seen a rapid recent growth in research output, and is perhaps more amenable to real time application. The motivation is to find a data-driven model effective in modeling a system locally in both space and time.

This more recent approach to modeling dynamical systems, and nonlinear dynamical systems in particular, comes from applying newer data-driven advancements to an older result in linear operator theory. Specifically, the Koopman operator was shown in the 1930's for the case of a Hilbert space of square integrable functions to be able to lift finite dimensional nonlinear dynamics into an infinite dimensional linear operator acting on observables of the system [33]. It wasn't until more recent years when advancements in data-driven methods in the fluid dynamics community [37], that methods became popularized for estimating a finite dimensional representation of the Koopman operator, when the dynamics are contained, or nearly contained, in some invariant subspace. This family of data-driven methods is referred to as dynamic mode decomposition (DMD), and work continues on understanding and improving this method under various conditions to this day [40, 53, 67]. In the simplest case, DMD estimates the best-fit linear operator that pushes measurements of a system forward in time. Resulting spectral elements known as the DMD modes and eigenvalues were shown to be related to the spectral properties of the Koopman operator [40], thus leading to the proliferation of research on this topic in recent years.

A specific approach to modeling Koopman dynamics in the form of a Kalman like observer was shown in [68]. This takes the output of the Koopman decomposition, and puts the system into a linear observer form similar to a classic state space model as used in Kalman filtering applications. This connects the model free methods of the Koopman decomposition to some of the model driven techniques in state space estimation, and specifically as applied to nonlinear systems. This is a different approach to than that in the nonlinear observer approach of the Extended Kalman Filter (EKF), which relies on linearization, some understanding of the underlying model, and additive Gaussian noise. This is in contrast to the goals of this work, which utilizes the Koopman framework and observer form from [68] to discover a model from measurements of the system, and adapt it to traditional estimation techniques. For further discussion on the EKF and Koopman observer

model, see [68].

In this chapter, it is proposed to take advantage of these methods for modeling stochastic dynamical systems of radar backscatter from the sea surface. The discovered model is compared against the equation generated synthetic observables of the system. It needs to be highlighted that this chapter is not the first to make this connection, as a recent work [69] has separately appeared to be the first to demonstrate the usefulness of DMD in application to the IPIX radar sea clutter data set, and shows some properties of their DMD modes and reconstruction. In this work, the problem has not been approached from existing data sets, but from stochastic dynamical system models of sea clutter scattering, and demonstrate an approach for using these methods in target detection. These methods can potentially provide a more general and flexible foundation for algorithm development. In this chapter, the SDE model of K-distributed scattering is reviewed, as well as the DMD algorithm and the Koopman operator. An initial approach for an anomaly detection framework for small target detection in the presence of sea clutter is proposed.

6.1.1 Related Work

In [64], an extension of the random walk model for K-distributed noise (see [70]) gave a continuous model in the form of a set of stochastic differential equations accounting for the dynamics of the scattering amplitude and associated parameters. Additional work from Field and collaborators [65, 71, 72] discusses various aspects of this model, with a self contained summary found in their associated book [8]. This model was considered in the scenario of additive noise in [73]. Other work along these lines has looked into methods for estimating the parameters [66], and analysis of the associated Fokker-Planck equations of the sea clutter SDEs [74]. An overview of the nonlinear dynamics of sea clutter is discussed extensively in [7], and in relation to the SDE model in [75].

Other sea clutter related work has considered the target detection problem in the context of anomaly detection [76–78]. In [76], an anomaly detection based approach for sea surface small target detection is outlined with a K-nearest neighbor classifier. A convex hull learning algorithm is used on singular spectrum analysis based features in [78]. An adaptive filtering scheme was proposed in [79] to suppress

the clutter, and preserve the target spectrum. In [80], various clutter suppression techniques including root cycle cancellation, singular value decomposition based methods, empirical mode decomposition, and wavelet weighted construction are compared. Low flying targets in the presence of sea clutter are addressed with a filtering scheme in [81]. Approaches for waveform design in the sea clutter regime are proposed in [82, 83]. A distribution model of sea clutter is the most common approach among sea clutter based radar literature.

Anomaly detection in nonlinear system has been studied and found potential application in numerous fields. A review on general anomaly detection in time series data can be found in [47]. In uncertain dynamical systems, [84] developed an anomaly detection algorithm utilizing symbolic time series analysis. The technique was built upon spectral analysis of ergodic sequences of measure-preserving transformations. On systems modeled by stochastic differential equations, [85] proposed an anomaly detection framework that relies on estimating parameters of an SDE model. A Koopman operator theoretic perspective on anomaly detection was proposed in [86], where the monitored metric was the maximum deviation in the linear Koopman coordinates. Here, a data-driven Koopman-based approach that builds upon the linear estimation approach of Kalman filtering will be used for detecting anomalous behavior in sea clutter data.

6.2 Stochastic Model of Sea Clutter

In this section, the fundamentals Field's model of the K-distributed SDE scattering model is summarized (see [64, 65] for review). For a better understanding of SDEs or the Itô calculus see [61], or for a shorter overview [8, Ch. 2].

An SDE is an equation in which at least one of the terms is a stochastic process, with a typical equation taking the following form

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \tag{6.1}$$

where μ is referred to as the drift, σ as the volatility or diffusion, and W_t is a Wiener process, also called Brownian motion. Technical details can be found in [61], and in this chapter will avoid measure theoretic discussions and terminology.

6.2.1 Rayleigh Scattering

Field's model of electromagnetic scattering from random media represents the complex electric field scattered from the sea surface as a sum of the contributions of the individual scatterers. This results in a random walk model in the complex plane:

$$\mathcal{E}_t^{(N)} = \sum_{k=1}^{N} \exp\left[i\varphi_t^{(k)}\right] \tag{6.2}$$

with constant number of scatterers N, and phase of the kth scatterer $\varphi_t^{(k)}$. These component phases form a stochastic process, or more specifically a diffusion model, that develop as a collection of displaced Wiener processes $\{W_t^{(k)}\}$ on a suitable time scale

$$\varphi_t^{(k)} = \Delta^{(k)} + \mathcal{B}^{1/2} W_t^{(k)} \tag{6.3}$$

with random initialization $\{\Delta^{(j)}\}$. These are a set of independent random variables uniformly distributed on the interval $[0, 2\pi)$. The resulting phase dynamics follow

$$d\varphi_t^{(k)} = \mathcal{B}^{1/2} W_t^{(k)} \tag{6.4}$$

Using Itò calculus, it was shown that the amplitude process $\psi_t = \lim_{N\to\infty}[\mathcal{E}_t^{(N)}/N^{1/2}]$ gives

$$d\psi_t = -\frac{1}{2}\mathcal{B}\gamma_t dt + (\mathcal{B}x)^{1/2} d\xi_t \tag{6.5}$$

where the continuous random variable x is the average scattering power coming from the large population limit $x = \lim_{N\to\infty}[N/\bar{N}]$.

6.2.2 K-scattering

Introducing a fluctuating number of steps $N \to N_t$ into (6.2) leads towards the desired K-distributed process. The resulting normalized amplitude has the following representation

$$\begin{aligned}\psi_t &= \lim_{N_t\to\infty}[(N/\bar{N})^{1/2}(\mathcal{E}_t^{(N_t)}/\bar{N}^{1/2})] \\ &= x_t^{1/2}\gamma_t\end{aligned} \tag{6.6}$$

where $x_t = \lim_{N\to\infty}[N/\bar{N}]$, γ_t is a continuous-valued cross section, and $\gamma_t = \lim_{N\to\infty}[\mathcal{E}_t^{(N_t)}/\bar{N}^{1/2}]$ is the fast varying speckle term that is a complex valued Ornstein-Uhlenbeck process with SDE

$$d\gamma_t = -\frac{1}{2}\mathcal{B}\gamma_t dt + \mathcal{B}^{1/2}d\xi_t \tag{6.7}$$

This compound scattering model can be used to describe a K-distributed intensity. By assuming a certain birth-death-immigration model on the scattering population, the resulting rescaled population variate $x \to \alpha x$ satisfies the SDE

$$dx_t = \mathcal{A}(\alpha - x_t)dt + (2\mathcal{A}x_t)^{1/2}dW_t^{(x)} \tag{6.8}$$

for independent Wiener process $W_t^{(x)}$, and α is a scaling parameter related to the aforementioned birth-death-immigration process. The asymptotic distribution of (6.8) can be shown to be the Γ-distribution. This Γ-distributed cross section is referred to as the texture, or RCS, in radar. With this notation, the K-distribution shape parameter is $\alpha - 1$. This shape parameter encodes the sea state, and characterizes extent of the nonlinearity [75]. The application of Itò's formula to ψ_t and the cross section SDE in (6.8), the dynamics of the scattered amplitude satisfy

$$\frac{d\psi_t}{\psi_t} = \left[\mathcal{A}\left(\frac{2(\alpha - x_t) - 1}{4x_t}\right) - \frac{1}{2}\mathcal{B}\right]dt + \left(\frac{\mathcal{A}}{2x_t}\right)^{1/2}dW_t^{(x)} + \frac{\mathcal{B}^{1/2}}{\gamma_t}d\xi_t \tag{6.9}$$

The parameters $\mathcal{A}$ and $\mathcal{B}$ have the physical dimensions of frequency. It is discussed in [75] that these parameters can be understood as the inverse of the decorrelation time for the RCS and speckle, respectively. The constant $\mathcal{B}$ is an electromagnetic property, while $\mathcal{A}$ is determined as an intrinsic property of the statistics of the scattering surface. In radar sea clutter situations of interest, these parameters are such that $\mathcal{A} \ll \mathcal{B}$. Another parameter of interest is the intensity defined as

$$z_t = |\psi_t|^2 \tag{6.10}$$

In summary, separating the speckle into its real and imaginary components leads to the three governing SDEs in (6.11). Further manipulations of these equations can give other possible observables, but knowledge of these main equations will suffice for the remainder of the chapter.

$$\begin{cases} dx_t = \mathcal{A}(\alpha - x_t)dt + (2\mathcal{A}x_t)^{1/2}dW_t^{(x)} \\ d\gamma_t^{(R)} = -\frac{1}{2}\mathcal{B}\gamma_t^{(R)}dt + \mathcal{B}^{1/2}d\xi_t^{(R)} \\ d\gamma_t^{(I)} = -\frac{1}{2}\mathcal{B}\gamma_t^{(I)}dt + \mathcal{B}^{1/2}d\xi_t^{(I)} \end{cases} \tag{6.11}$$

6.3 DMD Based Reconstruction of Sea Clutter

6.3.1 Radar Return

Assume that a high resolution maritime surface surveillance radar receives the backscattered returns from a range cell in the form of a complex time series of measurements. The decision problem for sea-surface small target detection can be expressed as a binary hypothesis testing problem for the cell under test (CUT). In each range cell, assume that the radar receives a time series of the backscattered signal $x(t)$. The hypothesis testing problem becomes to determine if the CUT contains a target in the presence of sea clutter, or just sea clutter. This can be expressed as

$$\begin{cases} H_0 : x(t) = \Psi_t \\ H_1 : x(t) = s(t) + \Psi_t \end{cases} \tag{6.12}$$

where $s(p)$ represents the returns of the target, and $\Psi(p)$ represents the combined contribution from sea clutter and noise. The difficulty in this problem lies in the clutter component, because as outlined previously, this is a dynamical system that can readily give rise to false alarms due to clutter spikes and doppler components in the return. The remaining purpose of this chapter is to apply these data driven techniques to discover a model of the dynamical system that is viable locally in the space and time of the CUT. From this model, targets would show up as anomalies in the returns. From (7.13), this effectively means to find a spatio-temporally local model of Ψ_t.

In addition to the stochastic framework described in Section II, a more realistic framework for a radar signal will include a contribution from thermal noise. Dropping the signal length dependence from the term, the observed scattered amplitude may be decomposed as

$$\Psi_t = \psi_t + n_t \tag{6.13}$$

This was studied in [73], a Wiener filtering approach was demonstrated to remove the thermal noise term from the expression. Again, underlying all methods directly utilizing the SDE model is the assumption of effectively estimating the parameters. This chapter instead continues with a data-driven approach not directly dependent on discovering anything about underlying equations. A Koopman based observer model was proposed in [68], that builds a linear observer model from the output DMD modes and eigenvalues of the dynamical system.

6.4 Simulation

6.4.1 Sea Clutter Stochastic Differential Equations

An example of the various components making up the signal can be seen in Fig. 6.1. From top to bottom, this plot shows γ_t, x_t, and $z_t = |\psi_t|^2$ from the SDE equations. The intensity plot shows the characteristic clutter spikes, and qualitatively resembles experimental data in the literature (see [7] for an example). The equation parameters chosen for Fig. 6.1 areg $\mathcal{B} = 100$, $\mathcal{A} = 1$, and $\alpha = 1$.

It is also beneficial to see how the behavior of the SDEs with different underlying parameters. Specifically, looking at changes in the RCS equation, with the parameters α and $\mathcal{A}$. These parameters describe properties of the scattering surface, so it is illuminating to see how these parameters change the qualitative behavior of the time series. First, the shaping parameter α has the effect of modifying the RCS as in Fig. 6.2. This parameter determines the gamma distribution associated with the RCS, which in turn influences the intensity K-distribution. Qualitatively, α provides a measure of the spikiness of the backscatter.

The effect of changing $\mathcal{A}$ can be seen in Fig. 6.3. This parameter can be seen to determing the fluctuation timescale of the texture process. It therefore also leaves the asymptotic statistics of the sea clutter invariant.

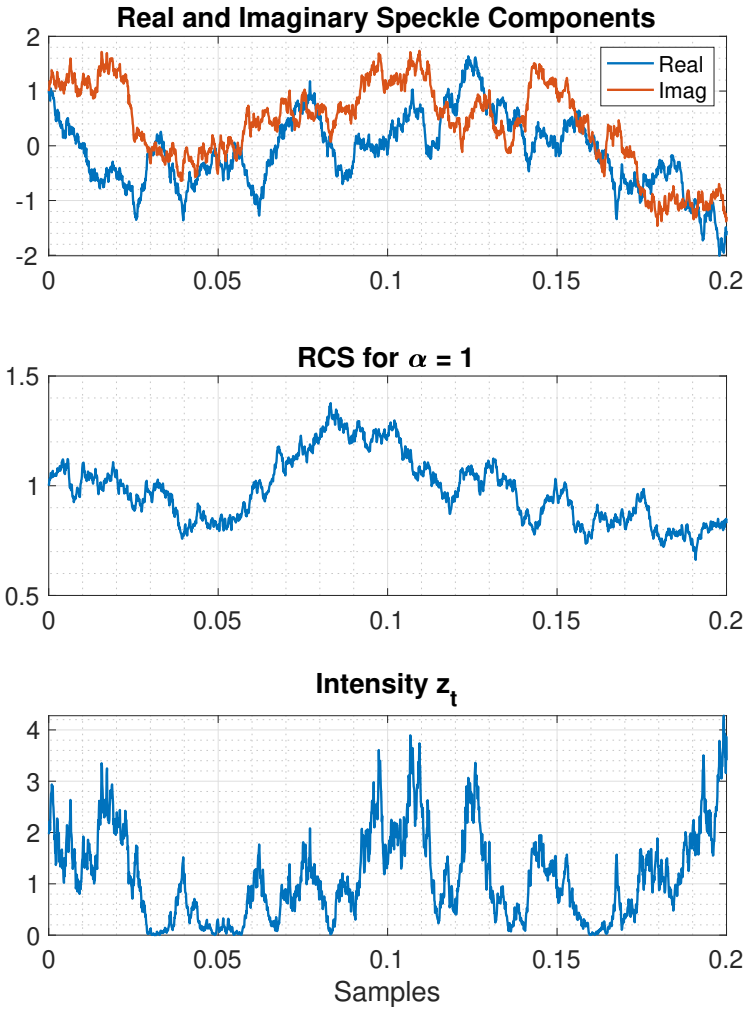


Figure 6.1: Example time series for (a) γ_t with parameter $\mathcal{B} = 100$, (b) cross section x_t with parameters $\mathcal{A} = 1$ and shaping parameter $\alpha = 1$, and (c) intensity $z_t = |\psi_t|^2$.

To test the efficacy of DMD and Koopman theory on data generated from this dynamical system, signals are generated using a similar procedure. Due to the nature of the SDEs, each run of the simulation will generate a different signal, though each will qualitatively resemble Fig. 6.1. To account for the randomness of the measurements in the simulations, each run generates a different initial

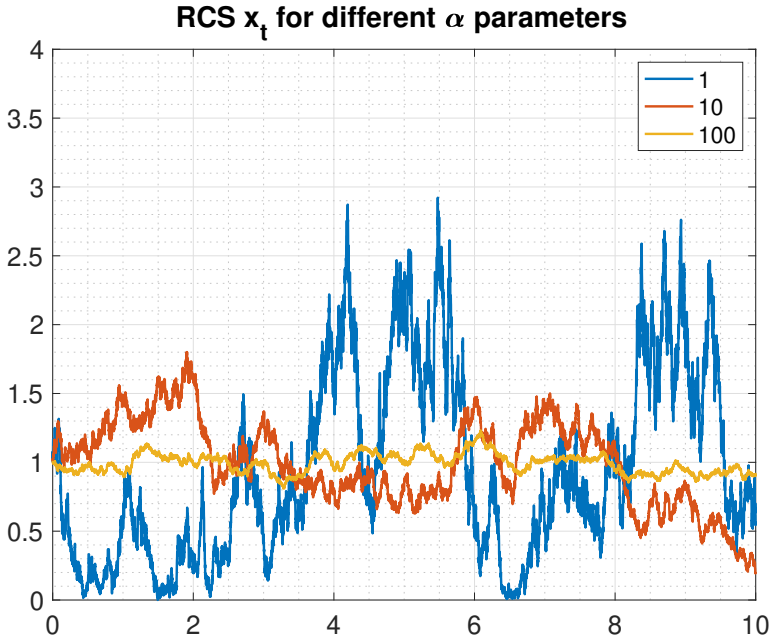


Figure 6.2: RCS time-series for various values of the shape parameter, simulated with $\alpha = 1, 10, 100$.

condition from a set uniformly distributed over some interval. In the next section, a preliminary study of applying the Koopman and DMD techniques to discover a model for the dynamics is presented. The initial goal will be to show how a discovered model can reconstruct the original signal, and then to see how prediction directly from the KMD performs. A modified Kalman-based estimation and prediction is reviewed and shown in the next chapter.

6.4.2 Koopman Mode Analysis of Sea Clutter SDEs

A data-driven analysis of this SDE based synthetic sea clutter data begins by effectively showing that the DMD and Koopman approach can effectively reconstruct the SDE time series from the calculated Koopman modes and eigenvalues in (6.14).

$$\mathbf{x}(t) = \sum_{j=1}^{r} \phi_j e^{\omega_j t} b_j = \mathbf{\Phi} \exp(\mathbf{\Omega} t)\mathbf{b} \tag{6.14}$$

It is assumed to have a collection of N sea clutter time series data for some

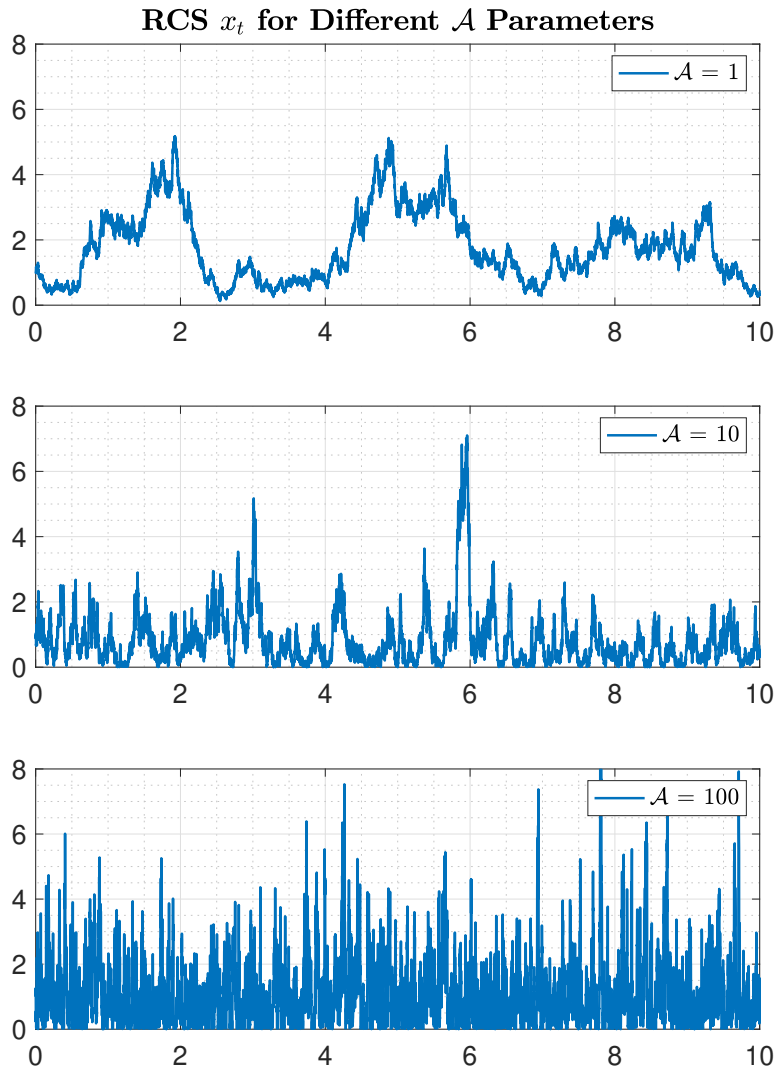


Figure 6.3: RCS time-series for various values of the $\mathcal{A}$ parameter, simulated with $\mathcal{A} = 1, 10, 100$.

CUT, where each time series is of length P. These snapshots will be stacked such that the first sample of each will be observation $g(\mathbf{x}_1)$. From these relatively longer time series, a shorter snapshot is taken such that a Hankel matrix of the data can be constructed as in (5.22). Exact DMD is performed on this data, where the goal

is to find the corresponding Koopman modes and eigenvalues.

For this simulation, ten instances of the clutter were generated. Each instance was made up of 10,000 samples with a time resolution of 10^{-4}. For the intensity, z_t, as well as the complex valued amplitude ψ_t, a 3000 sample length snapshot was taken from each. These were stacked into the aforementioned observation vector, and delay embedded for the Hankel matrix 2000 times. This results in a tall and skinny matrix of observations. To determine how many modes to use in modeling the signal, it is beneficial to look at the singular value spectrum. For the real part of Ψ_t, the first 200 singular values are plotted in Fig. 6.4. The nature of the underlying equation appears to make it so it is difficult to capture the dynamics through only a couple modes. The SVD was heuristically truncated to give 100 modes and singular values, though future work could include a disciplined approach to the cutoff. This has the effect of projecting the dynamics to some 100 dimensional space. These result in the KMD modes and eigenvalues used to reconstruct the signal using (6.14).

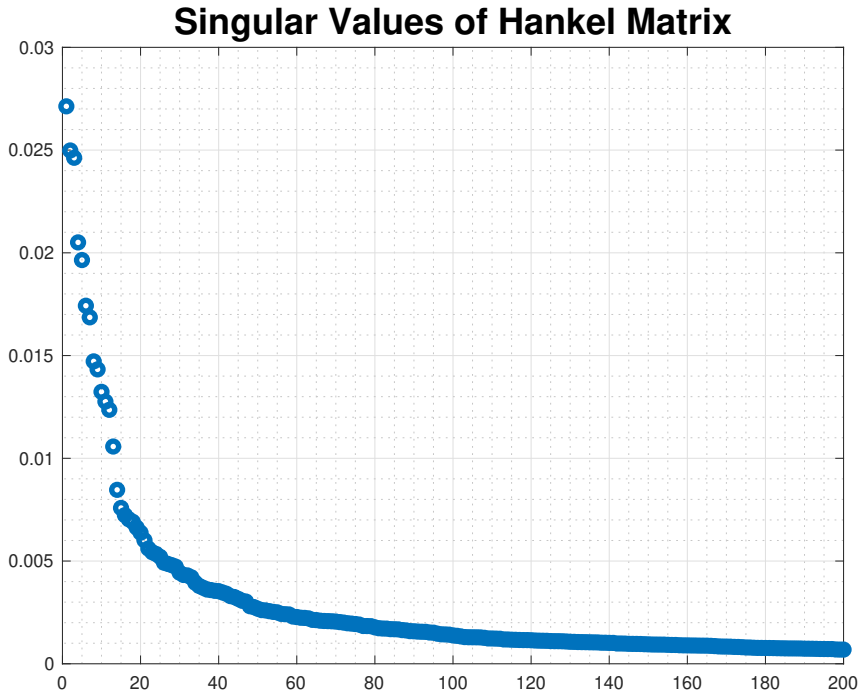


Figure 6.4: First 200 singular values of the Hankel matrix of Ψ_t.

For the intensity process, the result for reconstructing one intensity snapshot can be seen in Fig. 6.5, and similarly for the real part of the scattered amplitude Ψ_t in Fig. 6.6. In each case, exact DMD was performed on an additive white Gaussian noise version of the signal at 15 dB SNR. The KMD reconstruction of the signal was successful in capturing the major characteristics of the underlying SDE.

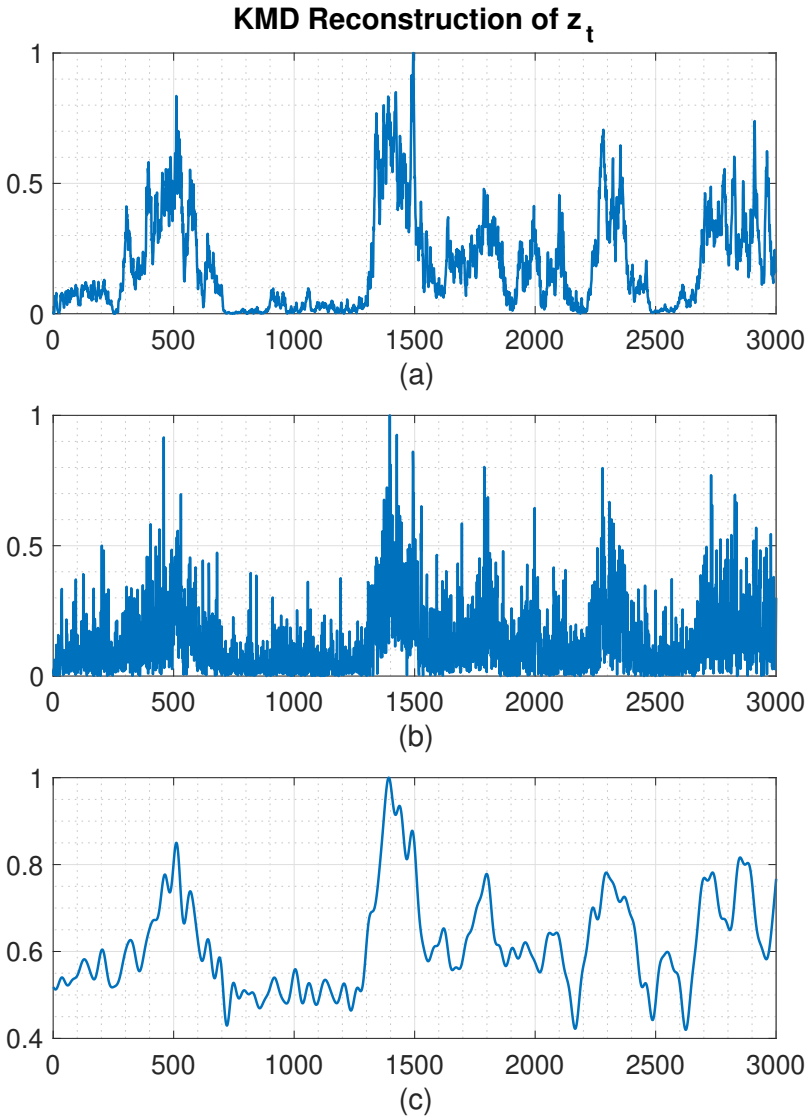


Figure 6.5: Example time series for (a) noise free intensity z_t (b) the same noise intensity at with additive white Gaussian noise at 15dB SNR, and (c) the Koopman mode decomposition reconstruction of the noise added intensity.

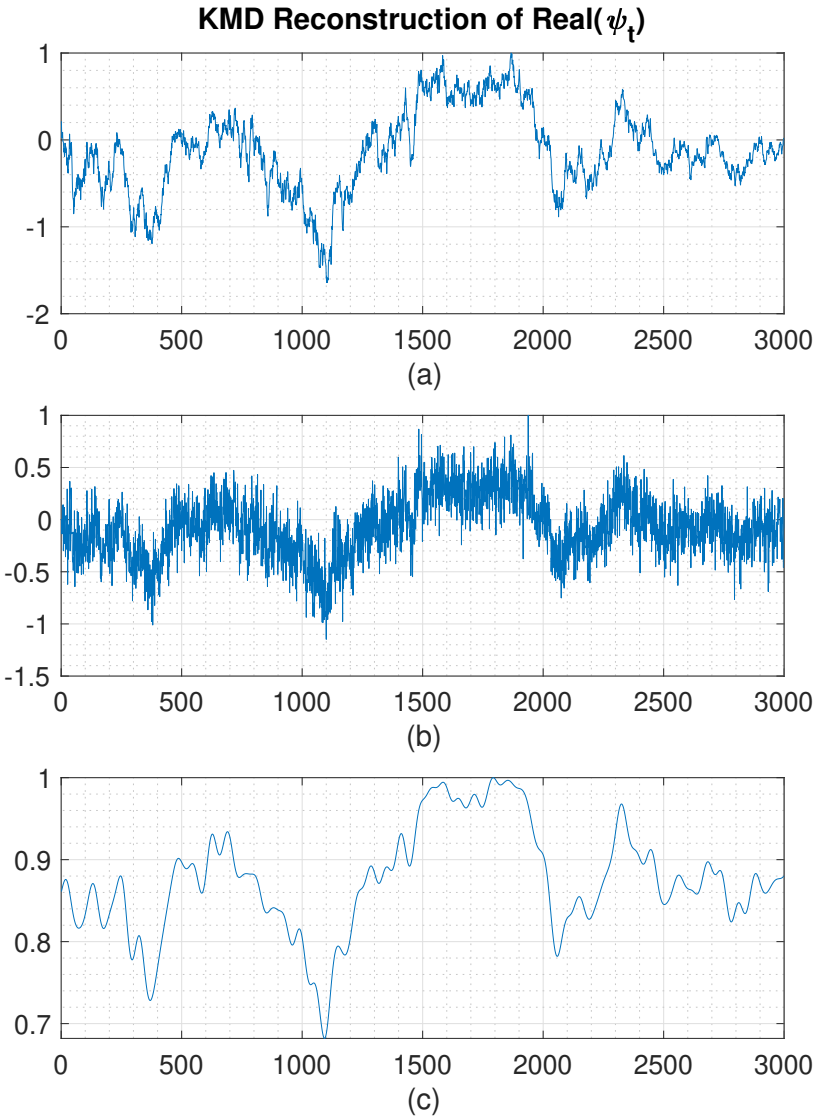


Figure 6.6: Example time series for (a) noise free intensity psi_t (b) the same noise intensity at with additive white Gaussian noise at 15dB SNR, and (c) the Koopman mode decomposition reconstruction of the noise addedc intensity.

The algorithm was able to reconstruct each of the separate time series snapshots used in the observation vector. This means that the algorithm was successful in finding a higher dimensional representation in which the dynamics propogate. However, exact DMD was not able to come up with an exact linear representation of

the signals. This can be seen by looking at the right singular vectors of the Hankel matrix SVD. It is known that sinusoids are eigenvectors of linear systems. In the case of a perfect linear representation of the dynamics, these singular vectors will look like sinusoids, because they have to do with the time dynamics of the discovered system [50]. Reconstructing the signals seems not to have been a problem, but longer term effective prediction of the signals could be an issue. Looking at the first five singular vectors in Fig. 6.7 gives an idea as to the linearity of the model.

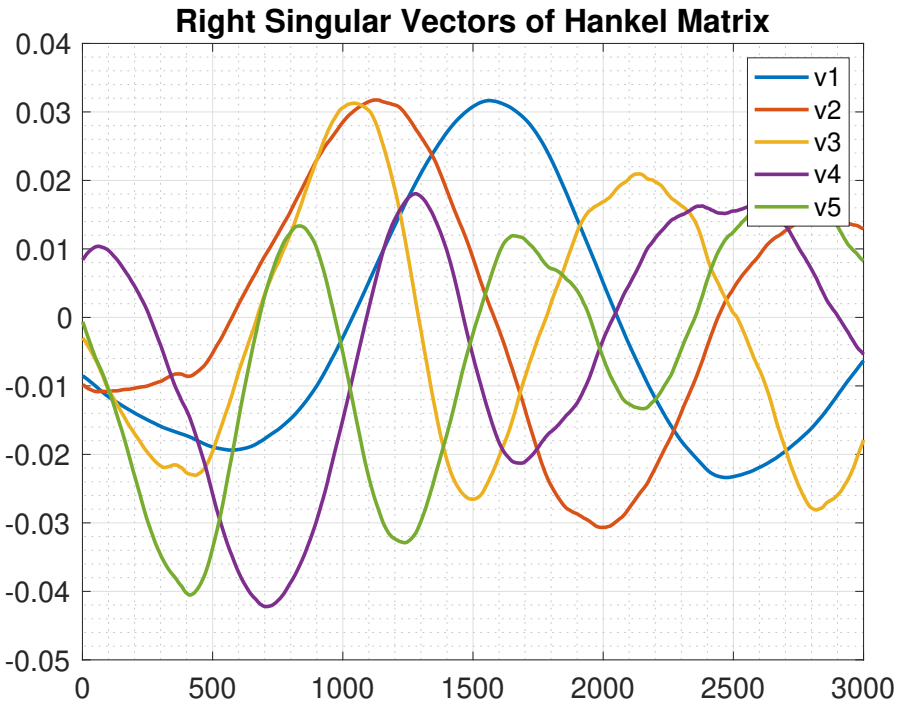


Figure 6.7: First five right singular vectors of the Hankel matrix of Ψ_t.

It is useful to look at the mean squared error of the model reconstruction, both over the length of the snapshot used in the Hankel matrix, as well as the prediction capabilities of the KMD by itself. An example of this is shown in Fig. 6.8. The vertical black bar is the length of the first snapshot used as a time-delay observable. The Hankel matrix will take information farther than this up to a 2000 sample embedding. It is apparent that the quality of the model to predict future behavior decays the further away the time series gets from the data used for learning. This

is expected, as it was shown that this did not build a perfect model. However, current results do show promise. In particular, in the next chapter the prediction model is put into a Kalman filtering framework. The discovered model does not need to be perfect, but is shown to be good enough for an initial anomaly detection in sea clutter data.

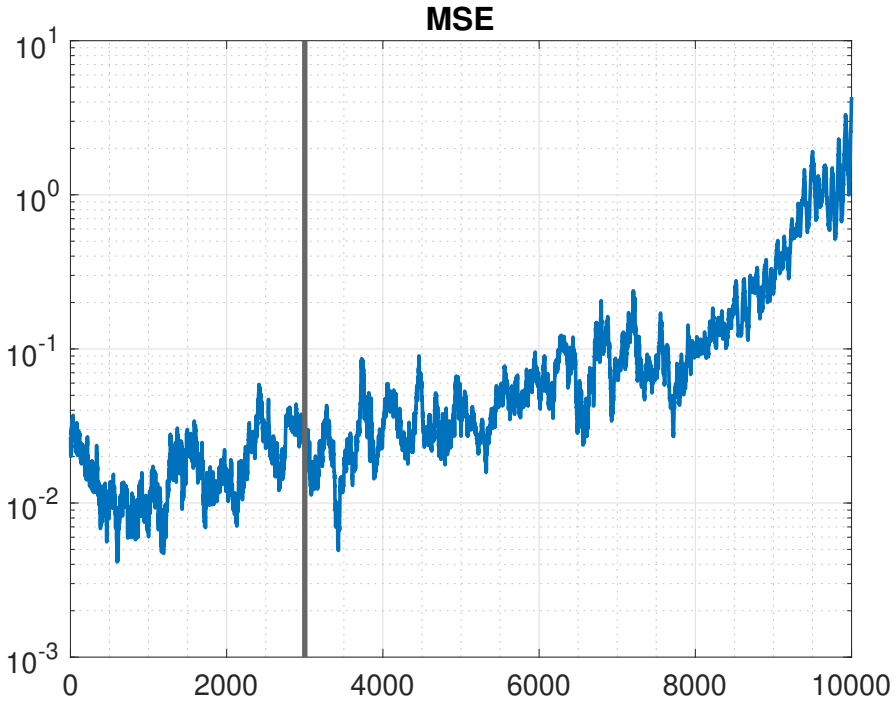


Figure 6.8: Mean squared error of KMD reconstruction and prediction over 100 runs.

CHAPTER 7

ANOMALY DETECTION ON DMD SEA CLUTTER MODEL

In this chapter an anomaly detection framework is reviewd and tested on the sea clutter model from the previous chapter. The overall goal is to demonstrate the ability to use the Koopman/DMD framework to detect anomalies as parameter changes in sea clutter, and set up future work in small target detection on these models. First, an anomaly detection framework presented in [42] is outlined, then this is applied to a test case of the Van der Pol oscillator. After demonstrating the ability to detect anomalies as parameter changes in the Van der Pol system, this framework is then similarly applied to the sea clutter equations from the previous chapter.

7.1 Anomaly Detection Framework

7.1.1 Koopman Observer Form

The basic anomaly detection approach presented here will rely on monitoring the innovation sequence of a Kalman filter. Before getting into this, it is necessary to show how an observer can be built from the Koopman framework. The most commonly used nonlinear observer and state estimation is the Extended Kalman

Filter (EKF), which despite its relative simplicity is based on linearization, which is not guaranteed to converge. There has been work showing the benefits of a Koopman based model compared to these techniques [68]. Work done by Surana et. al., [42], proposes nonlinear observer models using the Koopman operator framework. Successful application to the sea clutter problem would allow for the use of a Kalman-like linear observer to be used in nonlinear estimation.

Following notation, terminology, and results in [42], this section will review what is called the Koopman Observer Form (KOF). First, a distinction is made in the KMD between modes for full state observables, $\mathbf{g}(\mathbf{x}) = \mathbf{x}$, denoted as $\mathbf{v}^{\mathbf{x}}$ and called the Koopman Modes (KMs) and modes for other observables $\mathbf{h}$ as the Output Koopman Modes (OKMs), denoted as $\mathbf{v}^{\mathbf{h}}$. Here, time-delay observables will be used, and the superscript will be dropped.

Assume a model governed by the discrete time system, or discretized continuous time system, as follows

$$\begin{aligned} \mathbf{x}_t &= \mathbf{f}(\mathbf{x}_{t-1}) \\ \mathbf{y}_t &= \mathbf{h}(\mathbf{x}_t) \end{aligned} \tag{7.1}$$

where $\mathbf{x}_t \in \mathbf{X} \subset \mathbb{R}^d$ is the state vector, $\mathbf{f} : \mathbf{X} \to \mathbf{X}$ is a nonlinear state evolution, and $\mathbf{h} : \mathbf{X} \to \mathbb{R}^m$ is the measurement or observable of the state. In the sea clutter setting this is readily applied, as $\mathbf{x}_t$ is the sea state, and $\mathbf{y}_t$ is the radar return. Now, assume the observables can be decomposed into a finite sum of Koopman modes and eigenfunctions, i.e. that $\mathbf{h}(\mathbf{x}) \in \text{span}\{\phi_i\}_{i=1}^n$, such that

$$\mathbf{h}(\mathbf{x}) = \sum_{i=1}^{n} \phi_i(\mathbf{x})\mathbf{v}_i \tag{7.2}$$

For $\mathbf{\Phi} : \mathbb{R}^d \to \mathbb{R}^d$, define

$$\mathbf{z}_t = \mathbf{\Phi}(\mathbf{x}_t) \tag{7.3}$$

where this nonlinear change of coordinates is defined as

$$\mathbf{\Phi}(\mathbf{x}) = (\hat{\phi}_1(\mathbf{x}), \hat{\phi}_2(\mathbf{x}), \ldots, \hat{\phi}_n(\mathbf{x}))^*,$$

and constructed as follows

- $\hat{\phi}_i = \phi_i$ if ith KEF is real, and

- $\hat{\phi}_i = 2\text{Re}(\phi_i)$ and $\hat{\phi}_{i+1} = -2\text{Im}(\phi_i)$, if i and $(i+1)$'th KEFs are complex conjugate pairs

The authors refer to this as the Koopman Canonical Transform (KCT), and the coordinates $\mathbf{z}_t$ as the Koopman Canonical Coordinates (KCC). It can then be shown that

- $z_{i,t} = \lambda_i z_{i,t-1}$, if the ith KEF is real
- $(z_{i,t}, z_{i+1,t})^T = \mathbf{Q}_{\lambda_i}(z_{i,t-1}, z_{i+1,t-1})^T$, if the i and $(i+1)'$th KEFs are complex conjugate pairs, where

$$\mathbf{Q}_\lambda = |\lambda| \begin{bmatrix} \cos(\arg\lambda) & \sin(\arg\lambda) \\ -\sin(\arg\lambda) & \cos(\arg\lambda) \end{bmatrix} \tag{7.4}$$

The following state transition model is then

$$\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} \tag{7.5}$$

where the matrix $\mathbf{A}$ is an $n \times n$ real block diagonal and constructed as

- $\mathbf{A}$ has a diagonal entry $A_{i,i} = \lambda_i$, if the ith KEF is real
- $\mathbf{A}$ has a block diagonal entry $\begin{bmatrix} A_{i,i} & A_{i,i+1} \\ A_{i+1,i} & A_{i+1,i+1} \end{bmatrix} = \mathbf{Q}_{\lambda_i}$, if the i and $i+1$th KEFs are complex conjugate pairs

Additionally, the Koopman mode decomposition can be expressed in terms of KCC as

$$\mathbf{h}(\mathbf{x}_t) = \mathbf{C}\mathbf{z}_t. \tag{7.6}$$

The matrix $\mathbf{C} \in \mathbb{R}^{m \times n}$ is formed from the output Koopman modes, where the ith column of $\mathbf{C}$ is $\mathbf{v}_i$ if ith KEF is real, and $i, i+1$th columns are $\text{Re}(\mathbf{v}_i)$ and $\text{Im}(\mathbf{v}_i)$, respectively, if i and $i+1$th KEFs are complex conjugate pairs. This results in the following linear system

$$\begin{aligned} \mathbf{z}_t &= \mathbf{A}\mathbf{z}_{t-1} \\ \mathbf{y}_t &= \mathbf{C}\mathbf{z}_t \end{aligned} \tag{7.7}$$

For further explanation and details regarding this decomposition, see [42].

7.1.2 Anomaly Detection Using KOF

The purpose of the anomaly detection framework is to detect in the time series data a deviation in the normal behavior. In this chapter, the purpose is to show how this can be used to detect a change in the sea state, such that it could possibly be used for small target detection. The Koopman based observer can be used along with a Kalman filtering framework to find an estimate of $\hat{\mathbf{z}}$ from the radar history. Assume a noise added state and observables in the KOF ,

$$\begin{aligned} \mathbf{z}_t &= \mathbf{A}\mathbf{z}_{t-1} + \mathbf{w}_t, \\ \mathbf{y}_t &= \mathbf{C}\mathbf{z}_t + \mathbf{v}_t \end{aligned} \tag{7.8}$$

where $\mathbf{w}_t$ and $\mathbf{v}_t$ are zero mean Gaussian random variables with corresponding covariance matrices $\mathbf{Q}$ and $\mathbf{R}$. The innovations sequence is defined as the difference between the actual and predicted observation as

$$\mathbf{e}_t = \mathbf{y}_t - \mathbf{y}_t^p, \tag{7.9}$$

where $\mathbf{y}_t^p = \mathbf{C}\hat{\mathbf{z}}_t$ is the predicted observation. Anomaly detection can be performed as a hypothesis testing problem on this sequence. For example, a decision threshold η can be chosen and the decision rule could be summarized as

$$|e_t|^2 \underset{H_1}{\overset{H_0}{\lessgtr}} \eta \tag{7.10}$$

7.2 Van der Pol System

In this section a basic example of how the KOF Kalman filtering based model can be used to discover a data-driven model of a dynamical system, and use this model for anomaly detection. The Van der Pol oscillator is a second order differential equation with nonlinear damping, and is given by (7.11). The scalar parameter μ determines the strength of the damping and nonlinearity.

$$\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} - \mu(1 - x^2(t))\frac{\mathrm{d}x(t)}{\mathrm{d}t} + x(t) = 0 \tag{7.11}$$

In this example, the anomaly is modeled as a change in this μ parameter. This oscillator is often characterized in its two dimensional form by setting $x_1 = x$ and

$x_2 = \dot{x}_1$. This results in the set of equations

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \mu(1 - \dot{x}_1^2)x_2 - x_1 \end{aligned} \tag{7.12}$$

Plotting the limit cycle in these x_1 and x_2 coordinates results in the well known phase plane plot of the oscillator in Fig. 7.1. Also in this plot is a change in the signal happening when μ changes from 1.2 to 1.5, and is shown in red. The following example will take a sampled window of the x_2 signal, and input this into the Hankel DMD algorithm to approximate the Koopman mode decomposition of the signal.

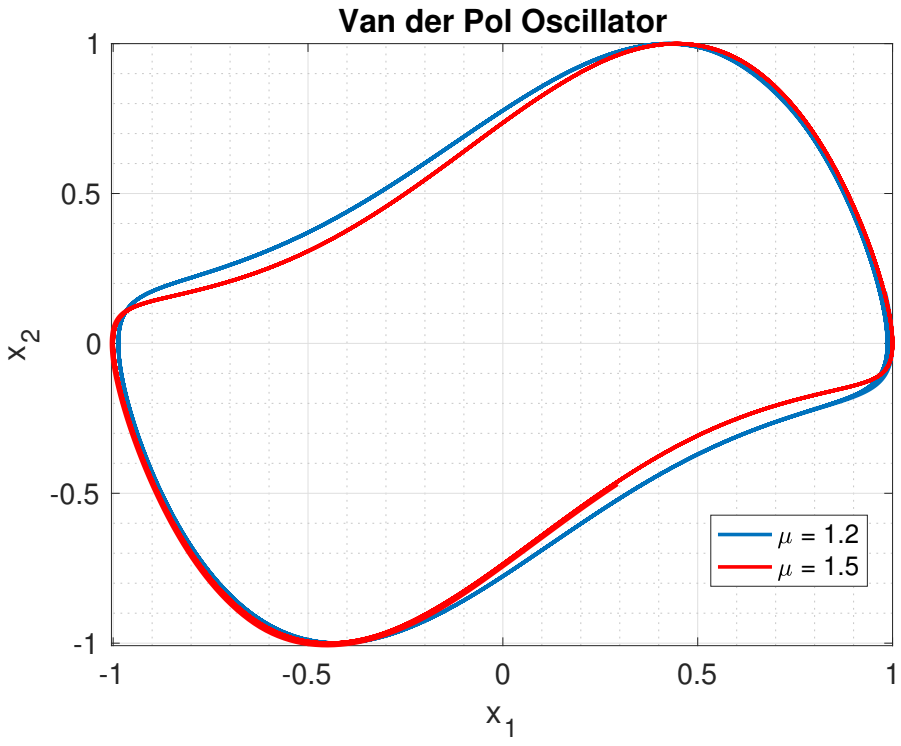


Figure 7.1: Van der Pol oscillator with parameter change.

The Hankel DMD algorithm is applied to the $\mu = 1.2$ signal, and the resulting Koopman modes and eigenvalues are used to construct the noise added KOF observer in (). This is then applied to a signal containing the parameter change in Fig. 7.1, where the vertical bar denotes the location of the parameter change.

The true and estimated signals are plotted in Fig. 7.2(a), and the squared error is plotted in Fig. 7.2(b). The peaks in the error plot demonstrate the ability of this approach to be used for anomaly detection.

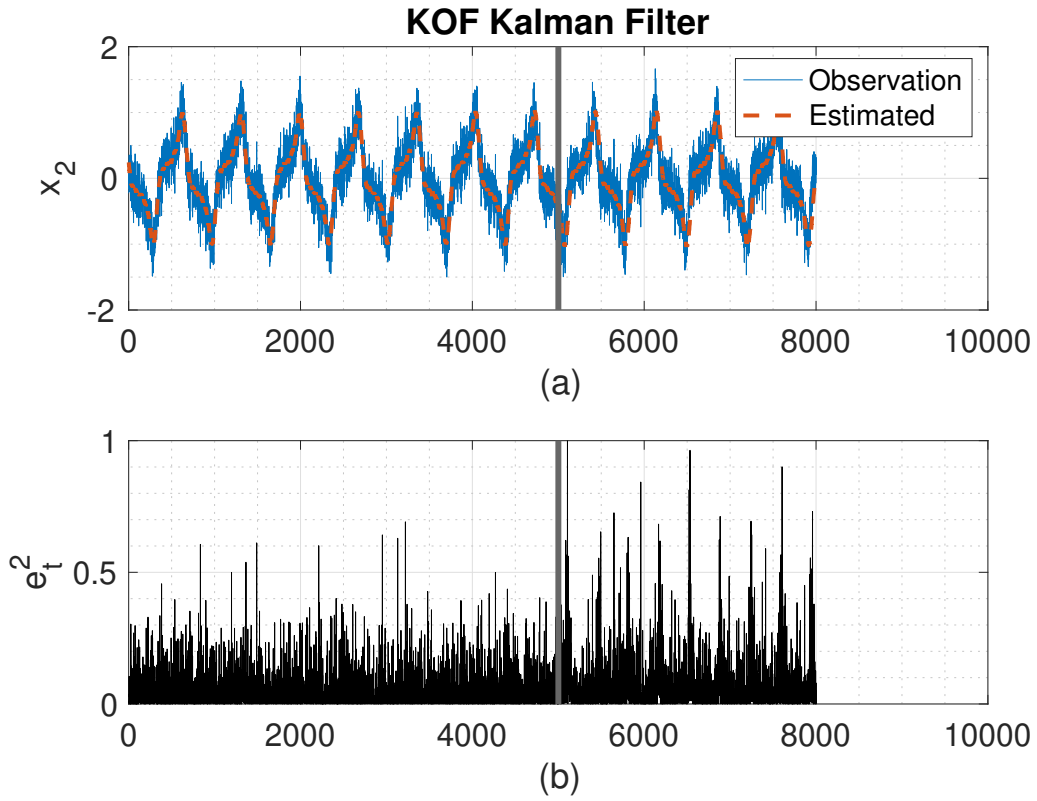


Figure 7.2: (a) True observation in blue, and KOF Kalman filter estimated state in dotted red line. (b) Innovations time series.

7.3 Sea Clutter SDEs

7.4 Radar Return

Assume that a high resolution maritime surface surveillance radar receives the backscattered returns from a range cell in the form of a complex time series of measurements. The decision problem for sea-surface small target detection can be expressed as a binary hypothesis testing problem for the cell under test (CUT). In

each range cell, assume that the radar receives a time series of the backscattered signal $x(t)$. The hypothesis testing problem becomes to determine if the CUT contains a target in the presence of sea clutter, or just sea clutter. This can be expressed as

$$\begin{cases} H_0 : x(t) = \Psi_t \\ H_1 : x(t) = s(t) + \Psi_t \end{cases} \tag{7.13}$$

where $s(p)$ represents the returns of the target, and $\Psi(p)$ represents the combined contribution from sea clutter and noise. The difficulty in this problem lies in the clutter component, because as outlined previously, this is a dynamical system that can readily give rise to false alarms due to clutter spikes and doppler components in the return. The remainder of this chapter applies these data driven techniques to the discover a model of the dynamical system that is viable locally in the space and time of the CUT. From this model, targets would show up as anomalies in the returns. From (7.13), this means effectively a temporally local model of $\Psi(p)$.

In addition to the stochastic framework described in Section II, a more realistic framework for a radar signal will include a contribution from thermal noise. Dropping the signal length dependence from the term, the observed scattered amplitude may be decomposed as

$$\Psi_t = \psi_t + n_t \tag{7.14}$$

This was studied in [73], a Wiener filtering approach was demonstrated to remove the thermal noise term from the expression. Again, underlying all methods directly utilizing the SDE model is the assumption of effectively estimating the parameters. We continue with a data-driven approach not directly dependent on discovering anything about underlying equations. A Koopman based observer model was proposed in [68], that builds a linear observer model from the output DMD modes and eigenvalues of the dynamical system.

7.5 Simulation

7.5.1 Effects of Parameter Changes on Asymptotic Distributions

The anomaly detection framework necessarily begins by defining an anomaly in the SDE generated sea clutter data. In this chapter, this is achieved by a change in the parameters $\mathcal{A}$, $\mathcal{B}$, and α defining the underlying SDE. A better understanding of what these represent will help clarify how and why this will be done. The continuum limit of the population model x_t is the gamma distribution, with parameter α. This parameter governs the variance and expected value of this distribution. The physical dimensions of $\mathcal{A}$ and $\mathcal{B}$ are frequency, with the reciprocals representing the correlation time scaled for the texture, or RCS modulation, and speckle components. The parameter $\mathcal{B}$ can be determined as a property of the transmitted wave. Constant $\mathcal{A}$ is determined as an intrinsic property of the scattering surface statistics, and is independent of the transmitted waveform [75]. Adjustment of $\mathcal{A}$ has the effect where smaller values yield longer duration between peaks in the RCS component. A change that could model the presence of a target could be an increase in either or both the $\mathcal{A}$ and α parameters. The reasoning being that presence of a target will result in regular and larger peaks in the RCS of the observed surface. A constant energy addition could instead be added, but the effect of waves and object motion could make this too simplifying of an assumption. In general, a desirable anomaly that could represent the presence of a small target would likely be expected to increase the RCS in that range cell, and/or give a higher density of peaks in that region of the backscattered signal.

To see how changes in the underlying parameters affect the model, and can thus be used to represent an anomaly, it is beneficial to look at the distribution behavior of the component parts. Specifically by looking at the stationary probabilities of the RCS and intensity for a change in α. Varying $\mathcal{A}$ was shown previously. Stationary probabilities of an SDE are found through the associated Fokker-Planch equation, and solving for the distribution after setting it equal to zero. Some results on this is shown in [8], and more thoroughly in [74]. The stationary, or asymptotic, distribution for the RCS is given by

$$p(\hat{x}_\infty = x) = \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}\mathbf{1}_{[0,\infty]}(x) \tag{7.15}$$

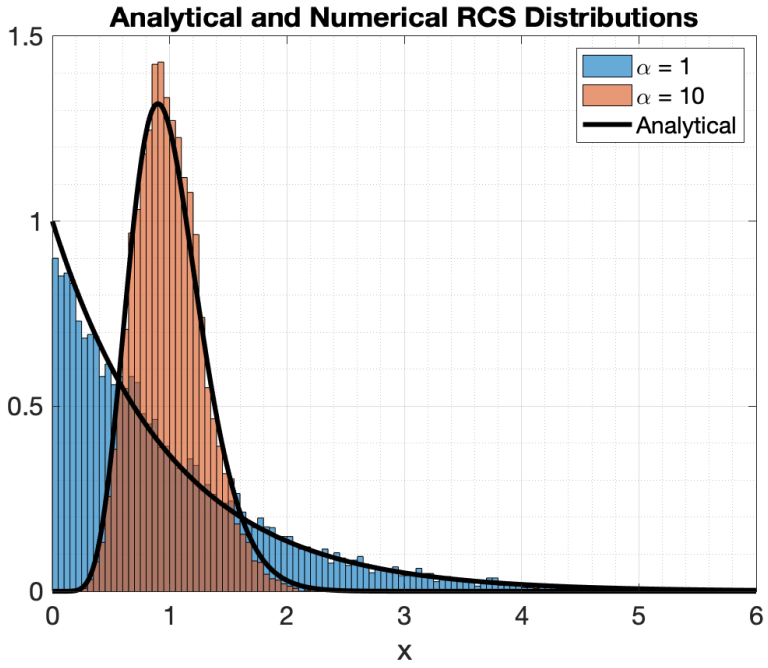
and for the unscaled version, i.e. $x_t \to \hat{x}_t/\alpha$, as

$$p(x_\infty = x) = \frac{\alpha(\alpha x)^{\alpha-1}e^{-\alpha x}}{\Gamma(\alpha)}\mathbf{1}_{[0,\infty]}(x). \tag{7.16}$$

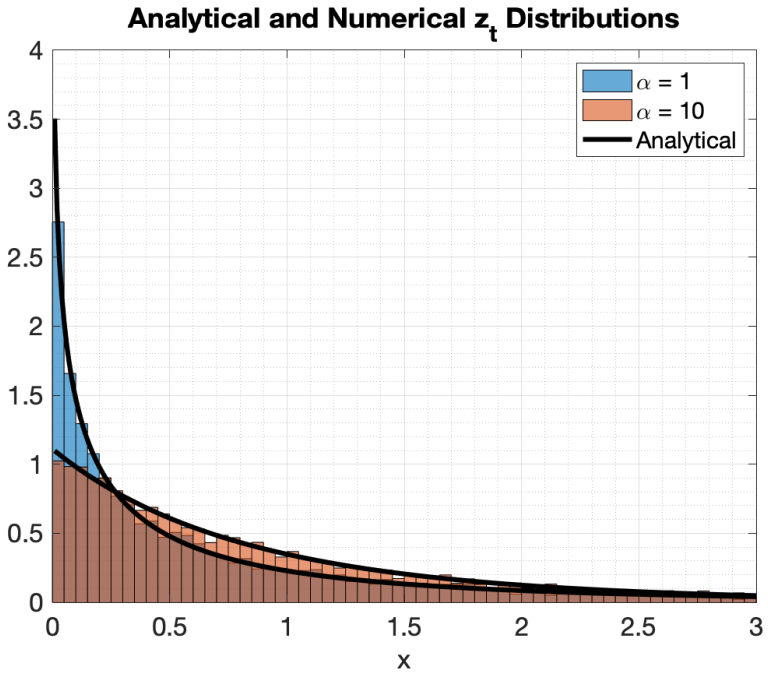
Similarly, the unscaled intensity, $z_t = |\Psi_t|^2$, distribution is found as

$$p(z_\infty = x) = \frac{2\alpha^{\frac{\alpha+1}{2}}x^{\frac{\alpha-1}{2}}}{\Gamma(\alpha)}K_{\alpha-1}(2\sqrt{\alpha x}), \tag{7.17}$$

where K is the modified Bessel function of the second kind. These equations only depend on α and not on the $\mathcal{A}$ parameter, as expected. These are compared to the numerical results in Fig. 7.3. In these plots, 10,000 runs of the SDEs were simulated for $\mathcal{B} = 100$, $\mathcal{A} = 1$, and varying α. The results were collected into the following histograms, and are shown to closely match the analytical solutions. These results shown the effect that changing α can have on the problem. It shifts the distribution of the RCS to a higher expected value, and spreads the distribution of the intensity to stronger values. This seems like a good candidate change, along with $\mathcal{A}$, to represent the presence of a small target in the data. The Koopman based framework will next be applied to time series containing an anomolous section with these parameter changes, with the goal to see an increase in the error sequence.



(a)



(b)

Figure 7.3: Histograms showing 10,000 run comparison of analytical and numerical (a) stationary RCS distribution for unscaled equation, and corresponding (b) stationary z_t distribution.

7.5.2 Complex Scattered Field

To test the KOF, first DMD was run using the same method as previously outlined on a simulated set of data to generate the $\mathbf{A}$ and $\mathbf{C}$ matrices in the KOF. This data was only used to generate the KOF, so other sea clutter time series data was generated to be used in testing. An anomalous section of the time series was generated through a parameter change in $\mathcal{A}$ in the SDE equations. The variance for the noise and measurement covariance matrices was set to 0.01. A Kalman filter was run on new time series, and the innovations were monitored. The goal is to see if there is a spike in the innovations sequence during the anomalous section of the time series, relative to the non-anomalous sections around it. An example of the intensity time series for anomalous data generated for the non-scaled equations with $\mathcal{A} = 50$, and non-anomalous sections around it at $\mathcal{A} = 1$ is shown in Fig. 7.4. The anomalous section is outlined by the black vertical bars. In this section, due to the parameter change there is a stronger returned intensity. The Kalman filter was run direction on the I and Q data of Ψ_t. The true time series, the estimated time series, and the innovations sequence is shown in Fig. 7.5. The vertical bars outline the same anomalous section, and it can be seen that within this section there is the largest spike in the innovations sequence. A similar example for an anomalous section generated with $\mathcal{A} = 10$ and $\alpha = 10$ for the scaled version of the equations is shown in Figs. 7.6 and 7.7. Both of these examples show the desired result of an increase in the error during the anomalous data. Future research can optimize the detection and hypothesis testing procedure for more robust anomaly detection and small target detection in sea clutter data.

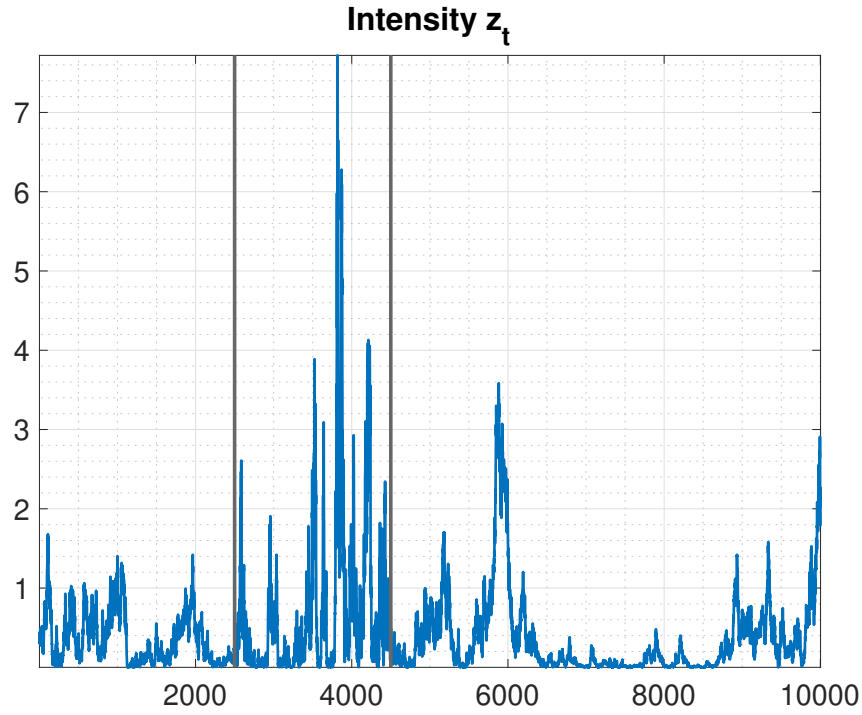


Figure 7.4: Intensity z_t of time series with anomalous segment. The segment between the black vertical bars represents a parameter shift from $\mathcal{A} = 1$ to $\mathcal{A} = 50$ in the unscaled equations. The I and Q samples of this are used in Kalman filtering algorithm.

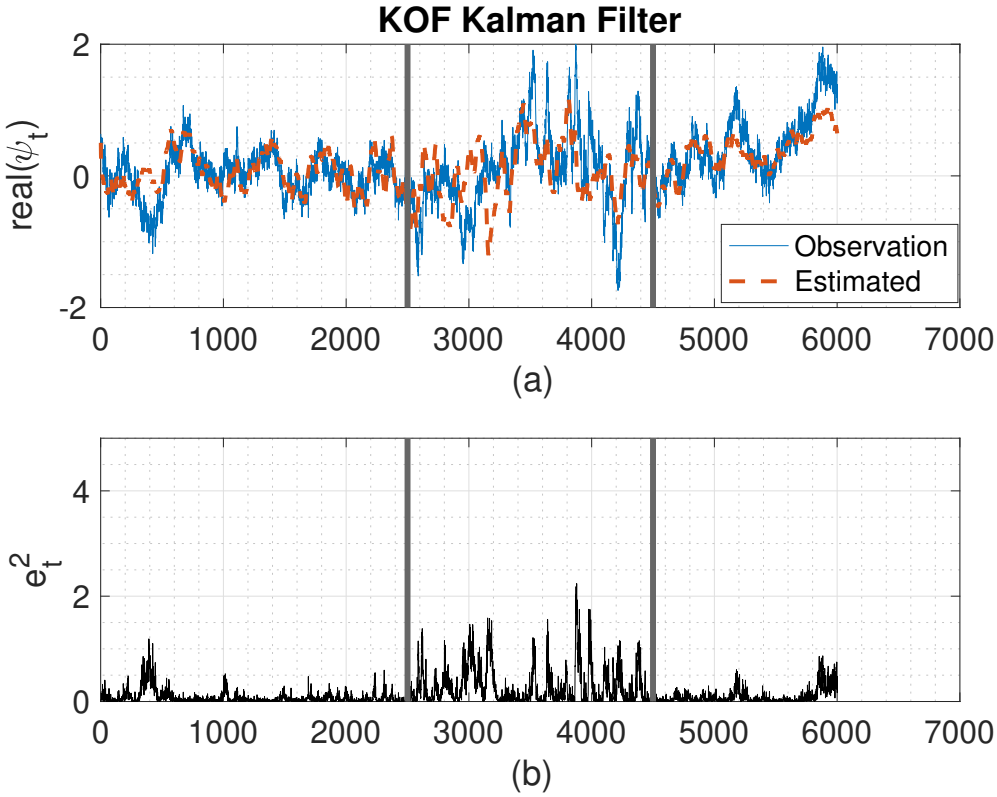


Figure 7.5: (a) True observation in blue, and KOF Kalman filter estimated state in dotted red line. (b) Innovations time series.

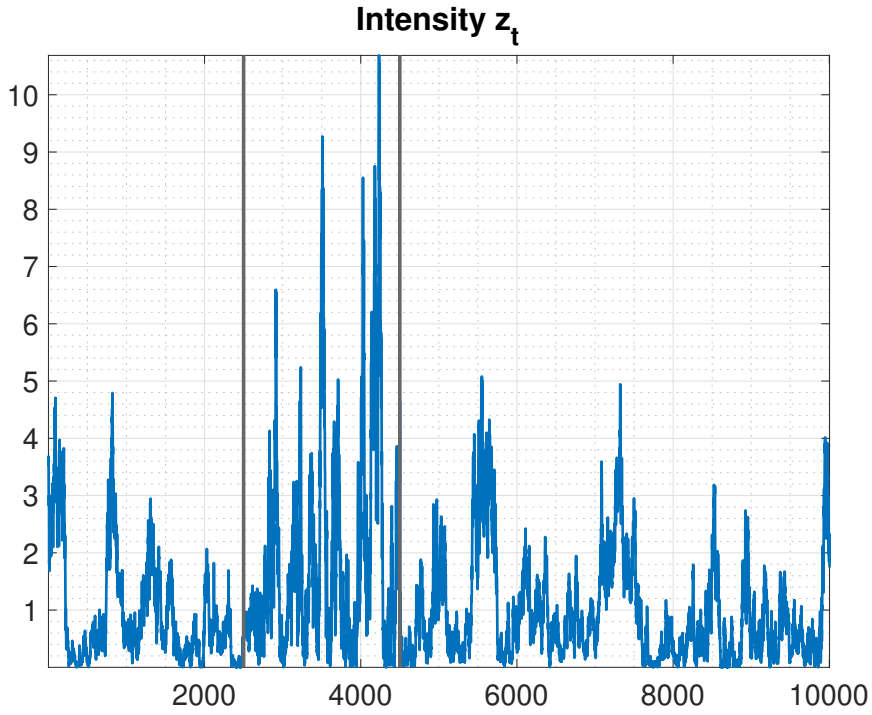


Figure 7.6: Intensity z_t of time series with anomalous segment. The segment between the black vertical bars represents a parameter shift from $\mathcal{A} = 1$ to $\mathcal{A} = 10$ and $\alpha = 1$ to $\alpha = 10$ for the scaled equations. The I and Q samples of this are used in Kalman filtering algorithm.

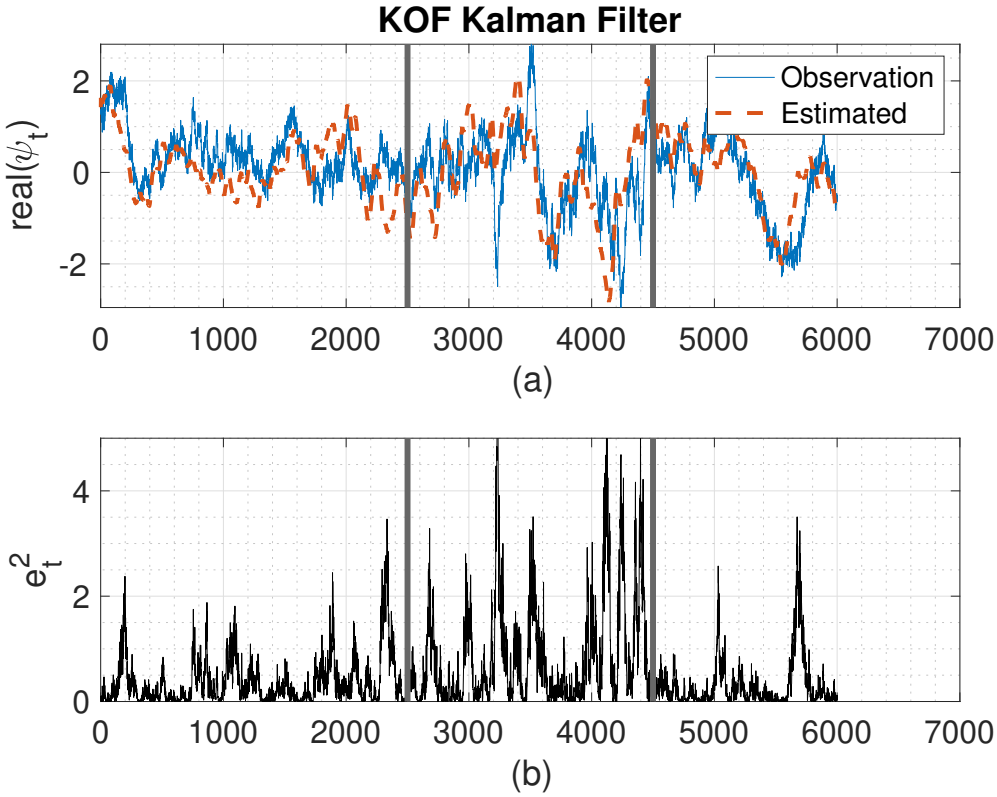


Figure 7.7: (a) True observation in blue, and KOF Kalman filter estimated state in dotted red line. (b) Innovations time series.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

This dissertation is primarily focused on various data-driven and operator theoretic approaches to radar modeling and signal processing. This began in Chapter 2 by reviewing and extending a linear operator theoretic model of the radar problem that can handle the presence of Doppler. The often used linear time-invariant signal model is shown to be but a special case of the more general class of operators discussed. Connections to the mathemaical formalism of quantum mechanics were shown, as well as connections to later chapters of the dissertation.

Two target classification receivers were developed in Chapter 3 that take advantage of a library of target response data across a large swath of target aspect angles. The first receiver focused on performing a sparse regression of a received signal onto this library of responses to classify not only the target, but the specific aspect angle that generated the return. The second receiver model is generated by deriving a target-tailored matched filter bank to the different target libraries. The approach classifies target returns to the target, and is invariant to the aspect angle generating the return. These receivers were tested in simulation against four different target CAD models, where the library of data consists of RCS data of each target across 180 degrees of aspect angles. Performance of the receivers was demonstrated on this library.

Waveform design approaches to maximize the signal-to-interference-plus noise ratio are often derived as eigenfunction solutions to maximizing the SINR. Finding an waveform to actually fit this solution is a comparably more difficult problem. In Chapter 4, a filtering approach for shaping a noise radar waveform to the SINR maximizing eigenfunction solution. Some preliminary simulations were shown to demonstrate its possible utility.

Small target detection in the presence of sea clutter is the next focus of study in this dissertation. Sea clutter is a difficult radar problem, as the clutter component of the recieved signal is a nonlinear dynamical system. This dissertation applies techniques from Koopman operator theory and dynamic mode decomposition to discover a model of sea clutter modeling stochastic differential equations directly from measurement data. The preliminary study demonstrates the potential for this approach to detect small targets as an anomaly in an anomaly detection framework. This work presents a preliminary study for future work on combining sea clutter dynamical models, and modern operator theoretic tools from dynamical systems theory.

8.2 Future Work

- For the operator theoretic model of radar, a general model and results were given that generalizes the usual LTI notion of radar processing. The reviewed work primarily rephrases existing theory into this alternative model. Future work could include not only further theoretical development, but exploring potential additional implications and applications that may be unique visible through this model.

- For the target classification recievers, future work could include joint optimization of the transmitted waveform and receiver, or further optimization of the receiver accounting for the presence of different target or clutter scenarios. Additionally, due to constraints in obtaining simulations of radar cross section data from CAD models, simulations were only performed on four different targets at one incident angle. With data over a small range of incident angles, the robustness of the different approaches to slight variations could be analyzed. Also, having numerous examples from the same class,

say having a few vehicles of the same general type (sedan, truck, etc.), there may be similarities in the response profiles due to having similar dominant scattering centers. One could then extract dominantly correlated structures representing the class more generally, and therefore classify vehicles not used in constructing the filters.

- For the noise waveform design, future work could include providing additional constraints to control the autocorrelation and range sidelobe behavior, optimizing the algorithm for spectrum notching, and experimentally demonstrating the gains that can be provided by pulse shaping. Additionally, providing the theory for the adaptive transmitter to take advantage of other radar system degrees of freedom would allow for greater versatility. This may include bandwidth, pulse duration, pulse repetition frequency, etc.

- For the sea clutter DMD modeling problem, future work will include further analysis and advancements to the anomaly detection framework for small target detection in sea clutter, such as improved metrics and robustness of the algorithms. Applying other variations of DMD can potentially improve the noise resistance of the algorithm. Beyond noise, methods such as compressed DMD and multiresolution DMD could prove to provide fruitful results. A new toolbox for radar algorithms in sea clutter could result.

- Further mathematical and theoretical analysis of Koopman operator theory and dynamic mode decomposition is still very available in this rapidly growing area of research. In the context of this dissertation, this could involve improved model discovery of stochastic differential equations. Further work could also include studying other related techniques to the Koopman operator.

Appendix A
Some Hilbert Space Background

Hilbert-Schmidt operators are a class of compact operators on Hilbert spaces. The aim of this appendix is to present some of the main definitions and tools necessary for completeness. See [87] for more details and proofs on the topic of Hilbert-Schmidt operators, and [48] for a more on Hilbert spaces and inner product spaces in general.

Definition A.0.1. (Hilbert Space) A Hilbert space, $\mathcal{H}$, is a real or complex inner product space that is complete with respect to the norm induced by the inner product.

Denote by $L^2(\mathbb{C})$ the Hilbert space of all square integrable functions. Then for $f, g \in L^2(\mathbb{C})$ the inner product is defined by

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t) g^*(t) dt,$$

where $g^*(t)$ denotes the complex conjugate of $g(t)$. The norm induced by the inner product is given by

$$\|f\| = \left(\int_{\mathbb{R}} |f(t)|^2 dt \right)^{1/2}$$

Definition A.0.2. (Riesz-Representation Theorem) For every bounded linear functional f on a Hilbert Space $\mathcal{H}$, there exists a unique element $y \in \mathcal{H}$ such that

$$f(x) = \langle x, y \rangle, \quad \text{for all } x \in \mathcal{H}$$

Definition A.0.3. (Compact Operator) Let E ande G be two normed spaces. A linear operator $T : E \to G$ is compact if and only if for every bounded sequence

$\{f_n\} \subset E$, the sequence $\{Tf_n\}$ has a convergent subsequence. If T is a compact operator, then it is also bounded.

Definition A.0.4. (Separable Hilbert Spaces) Let $\mathcal{H}$ be a Hilbert space. Then, $\mathcal{H}$ is a separable Hilbert space (i.e. $\mathcal{H}$ has a countable subset that is dense in $\mathcal{H}$) if and only if every orthonormal basis of $\mathcal{H}$ is countable.

A.1 Hilbert-Schmidt Operators

Definition A.1.1. Let $\mathcal{H}$ be a separable Hilbert space, and $\{e_n\}$ an orthonormal basis. Then, for any positive operator on $L(\mathcal{H})$ define $Tr(T) = \sum_{n=1}^{\infty} \langle Te_n, e_n \rangle$. The number $Tr(T)$ is independent of the orthonormal basis chosen.

This generalizes the linear algebraic notion of the trace as a sum of the diagonal entries of a matrix. However, because this involves infinite sums, not all operators will have a trace.

Definition A.1.2. (Hilbert-Schmidt Operator) Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. If in any orthonormal basis $\{e_n\}$ of $\mathcal{H}$, we have $Tr(T^*T) < \infty$, then the operator T is called Hilbert-Schmidt. If T is HS, then it is also a compact operator.

The Hilbert-Schmidt norm of T is defined as $\|T\|_{HS} = (\sum_{n=1}^{\infty} \|Te_n\|^2)^{1/2}$

Definition A.1.3 (Hilbert-Schmidt Integral Operator)**.** Let $D \subset \mathbb{R}^n$ be a bounded domain. A function $k : D \times D \to \mathbb{R}$ is called a Hilbert-Schmidt kernel if

$$\int_D \int_D |k(x,y)|^2 dxdy < \infty$$

That is, $k \in L^2(D \times D)$. Then define the integral operator T on $L^2(D)$, where $T : L^2(D) \to L^2(D)$, by

$$(Tf)(x) = \int_D k(x,y) f(y) dy$$

It can be shown that K is a bounded operator on $L^2(D)$. An integral operator K defined in this way is called a *Hilbert-Schmidt operator* with Hilbert-Schmidt norm

$$\|T\|_{\text{HS}} = \|k\|_{L^2}$$

Let D be a bounded domain in $\mathbb{R}^n$ and let $k \in L^2(D \times D)$ be a Hilbert-Schmidt kernel. Then, the integral operator $K : L^2(D) \to L^2(D)$ given by the above is a continuous and compact operator.

REFERENCES

[1] D. Cochran, S. D. Howard, and B. Moran, "Operator-theoretic modeling and waveform design for radar in the presence of doppler," in *2012 IEEE Radar Conference*. IEEE, 2012, pp. 0774–0777.

[2] J. E. Gray, "An interpretation of woodward's ambiguity function and its generalization," in *2010 IEEE Radar Conference*. IEEE, 2010, pp. 859–864.

[3] L. Cohen, *Time-Frequency Analysis*. Prentice hall New Jersey, 1995, vol. 778.

[4] L. G. Weiss, "Wavelets and wideband correlation processing," *IEEE signal processing magazine*, vol. 11, no. 1, pp. 13–32, 1994.

[5] J. R. Guerci, *Cognitive Radar: A Knowledge-Aided Fully Adaptive Approach*. Artech House, 2020.

[6] W. L. Melvin and J. Scheer, Eds., *Principles of Modern Radar: Advanced techniques*, ser. Radar, Sonar amp; Navigation. Institution of Engineering and Technology, 2012.

[7] S. Haykin, *Adaptive Radar Signal Processing*. Wiley Online Library, 2007.

[8] T. R. Field, *Electromagnetic Scattering from Random Media*. Oxford University Press, 2009.

[9] P. Tait, *Introduction to Radar Target Recognition*. Stevenage, Herts, UK: Institution of Engineering and Technology, 2005, vol. 18.

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

[11] J. Bae and N. A. Goodman, "Automatic target recognition with unknown orientation and adaptive waveforms," in *2011 IEEE RadarCon (RADAR)*, Kansas City, MO, May 2011, pp. 1000–1005.

[12] F. Gini, A. De Maio, and L. Patton, Eds., *Waveform Design and Diversity for Advanced Radar Systems*. Stevenage, Herts, UK: Institution of Engineering and Technology, 2012.

[13] 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.

[14] N. A. Goodman, P. R. Venkata, and M. A. Neifeld, "Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 105–113, 2007.

[15] D. A. Garren, M. K. Osborn, A. C. Odom, J. S. Goldstein, S. U. Pillai, and J. R. Guerci, "Optimal transmission pulse shape for detection and identification with uncertain target aspect," in *Proceedings of the 2001 IEEE Radar Conference*, Atlanta, GA, May 2001, pp. 123–128.

[16] S. Z. Alshirah, S. Gishkori, and B. Mulgrew, "Frequency-based optimal radar waveform design for classification performance maximization using multiclass fisher analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 3010–3021, 2020.

[17] Y. Wei, H. Meng, Y. Liu, and X. Wang, "Extended target recognition in cognitive radar networks," *Sensors*, vol. 10, no. 11, pp. 10 181–10 197, 2010.

[18] V. C. Chen, *The Micro-Doppler Effect in Radar*. Norwood, MA: Artech house, 2019.

[19] A. R. Persico, C. Clemente, D. Gaglione, C. V. Ilioudis, J. Cao, L. Pallotta, A. De Maio, I. Proudler, and J. J. Soraghan, "On model, algorithms, and experiment for micro-doppler-based recognition of ballistic targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1088–1108, 2017.

[20] C. Clemente, L. Pallotta, A. De Maio, J. J. Soraghan, and A. Farina, "A novel algorithm for radar classification based on Doppler characteristics exploiting orthogonal pseudo-Zernike polynomials," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 417–430, 2015.

[21] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar - Vol. I: Basic Principles*, Raleigh, NC: SciTech Publishing, 2010.

[22] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[23] Grabcad. [Online]. Available: https://grabcad.com/library

[24] R. M. Narayanan, A. Z. Liu, P. G. Singerman, and M. Rangaswamy, "Information elasticity in radar systems," *Electronics Letters*, vol. 54, no. 17, pp. 1049–1051, 2018.

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

[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, April 2011.

[27] B. Ravenscroft, J. W. Owen, J. Jakabosky, S. D. Blunt, A. F. Martone, and K. D. Sherbondy, "Experimental demonstration and analysis of cognitive spectrum sensing and notching for radar," *IET Radar, Sonar Navigation*, vol. 12, pp. 1466–1475, December 2018. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-rsn.2018.5379

[28] R. M. Narayanan, "Through-wall radar imaging using UWB noise waveforms," *Journal of the Franklin Institute*, vol. 345, no. 6, pp. 659 – 678, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0016003208000331

[29] C. J. Pici and R. M. Narayanan, "Multifunctional radar and communications waveform using chaos," in *NAECON 2018 - IEEE National Aerospace and Electronics Conference*, July 2018, pp. 568–572.

[30] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 1996.

[31] S. Kay, *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*.

[32] A. Lasota and M. MacKey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, ser. Applied Mathematical Sciences. Springer-Verlag, 1994.

[33] B. O. Koopman, "Hamiltonian systems and transformation in hilbert space," *Proceedings of the national academy of sciences of the united states of america*, vol. 17, no. 5, p. 315, 1931.

[34] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2019.

[35] H. Arbabi, "Introduction to koopman operator theory of dynamical systems," 2020.

[36] B. O. Koopman and J. v. Neumann, "Dynamical systems of continuous spectra," *Proceedings of the National Academy of Sciences*, vol. 18, no. 3, pp. 255–263, 1932.

[37] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics*, vol. 41, no. 1, pp. 309–325, 2005.

[38] I. Mezić and A. Banaszuk, "Comparison of systems with complex behavior," *Physica D: Nonlinear Phenomena*, vol. 197, no. 1-2, pp. 101–133, 2004.

[39] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of fluid mechanics*, vol. 641, pp. 115–127, 2009.

[40] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of fluid mechanics*, vol. 656, pp. 5–28, 2010.

[41] I. Mezić, "Analysis of fluid flows via spectral properties of the koopman operator," *Annual Review of Fluid Mechanics*, vol. 45, pp. 357–378, 2012.

[42] A. Surana, "Koopman operator framework for time series modeling and analysis," *Journal of Nonlinear Science*, vol. 30, no. 5, pp. 1973–2006, 2020.

[43] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz, "Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition," *Journal of neuroscience methods*, vol. 258, pp. 1–15, 2016.

[44] S. L. Brunton, M. Budišić, E. Kaiser, and J. N. Kutz, "Modern koopman theory for dynamical systems," *arXiv preprint arXiv:2102.12086*, 2021.

[45] S. Klus, P. Koltai, and C. Schütte, "On the numerical approximation of the perron-frobenius and koopman operator," *Journal of Computational Dynamics*, vol. 3, pp. 51 – 79, 09 2016.

[46] A. Mauroy, I. Mezić, and Y. Susuki, *The Koopman Operator in Systems and Control: Concepts, Methodologies, and Applications*. Springer, 2020, vol. 484.

[47] A. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, "A review on outlier/anomaly detection in time series data," *arXiv preprint arXiv:2002.04236*, 2020.

[48] A. W. Naylor and G. R. Sell, *Linear operator theory in engineering and science*. Springer, 1982.

[49] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, "Chaos as an intermittently forced linear system," *Nature communications*, vol. 8, no. 1, pp. 1–9, 2017.

[50] H. Arbabi and I. Mezić, "Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator," *SIAM Journal on Applied Dynamical Systems*, vol. 16, no. 4, pp. 2096–2126, 2017.

[51] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics*, vol. 1, no. 2, pp. 391–421, 2014.

[52] M. Kamb, E. Kaiser, S. L. Brunton, and J. N. Kutz, "Time-delay observables for koopman: Theory and applications," *SIAM Journal on Applied Dynamical Systems*, vol. 19, no. 2, pp. 886–917, 2020.

[53] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data–driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.

[54] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical systems and turbulence, Warwick 1980*. Springer, 1981, pp. 366–381.

[55] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of atmospheric sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[56] S. Luzzatto, I. Melbourne, and F. Paccaut, "The lorenz attractor is mixing," *Communications in Mathematical Physics*, vol. 260, no. 2, pp. 393–401, 2005.

[57] P. Walters, *Ergodic Theory – Introductory Lectures*, ser. Springer Lecture Notes in Mathematics (458). Springer-Verlag, 1975.

[58] N. Črnjarić-Žic, S. Maćešić, and I. Mezić, "Koopman operator spectrum for random dynamical systems," *Journal of Nonlinear Science*, pp. 1–50, 2019.

[59] L. Arnold, *Random Dynamical Systems*. Springer, 1998.

[60] M. T. Wanner and I. Mezić, *Robust Approximation of the Stochastic Koopman Operator*. University of California, Santa Barbara, 2020.

[61] B. Øksendal, *Stochastic differential equations: an introduction with applications*. Springer, 2003.

[62] K. Ward, "Compound representation of high resolution sea clutter," *Electronics letters*, vol. 16, no. 17, pp. 561–563, 1981.

[63] K. Ward, C. Baker, and S. Watts, "Maritime surveillance radar. part 1: Radar scattering from the ocean surface," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 137, no. 2. IET, 1990, pp. 51–62.

[64] T. R. Field and R. J. Tough, "Stochastic dynamics of the scattering amplitude generating k-distributed noise," *Journal of Mathematical Physics*, vol. 44, no. 11, pp. 5212–5223, 2003.

[65] ——, "Diffusion processes in electromagnetic scattering generating k-distributed noise," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 459, no. 2037, pp. 2169–2193, 2003.

[66] C. J. Roussel, A. Coatanhay, and A. Baussard, "Estimation of the parameters of stochastic differential equations for sea clutter," *IET Radar, Sonar & Navigation*, vol. 13, no. 4, pp. 497–504, 2019.

[67] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," pp. 391–421, 2014. [Online]. Available: /article/id/1dfebc20-876d-4da7-8034-7cd3c7ae1161

[68] A. Surana and A. Banaszuk, "Linear observer synthesis for nonlinear systems using koopman operator framework," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 716–723, 2016.

[69] Y. Zhang, L. Jiang, and H. T. Ewe, "A novel data-driven modeling method for the spatial-temporal correlated complex sea clutter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2022.

[70] E. Jakeman, "On the statistics of k-distributed noise," *Journal of Physics A: Mathematical and General*, vol. 13, no. 1, p. 31, 1980.

[71] T. R. Field, "Observability of the scattering cross-section through phase decoherence," *Journal of mathematical physics*, vol. 46, no. 6, p. 063305, 2005.

[72] T. R. Field and R. J. Tough, "Dynamical models of weak scattering," *Journal of mathematical physics*, vol. 46, no. 1, p. 013302, 2005.

[73] P. Fayard and T. R. Field, "Estimation of the k-scattered amplitude in additive noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 4, pp. 1668–1674, 2010.

[74] C. J. Roussel, A. Coatanhay, and A. Baussard, "Forward and backward probabilistic inference of the sea clutter," *Waves in Random and Complex Media*, vol. 29, no. 3, pp. 540–568, 2019.

[75] T. R. Field and S. Haykin, "Nonlinear dynamics of sea clutter," *International Journal of Navigation and Observation*, 2008.

[76] Z.-X. Guo and P.-L. Shui, "Anomaly based sea-surface small target detection using k-nearest neighbor classification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4947–4964, 2020.

[77] S.-N. Shi and P.-L. Shui, "Sea-surface floating small target detection by one-class classifier in time-frequency feature space," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6395–6411, 2018.

[78] X. Wu, H. Ding, N.-B. Liu, and J. Guan, "A method for detecting small targets in sea surface based on singular spectrum analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2021.

[79] S. Liu, Y. Ma, and Y. Huang, "Sea clutter cancellation for passive radar sensor exploiting multi-channel adaptive filters," *IEEE Sensors Journal*, vol. 19, no. 3, pp. 982–995, 2018.

[80] M. Lv and C. Zhou, "Study on sea clutter suppression methods based on a realistic radar dataset," *Remote Sensing*, vol. 11, no. 23, p. 2721, 2019.

[81] H. Xing and Y. Yan, "Detection of low-flying target under the sea clutter background based on volterra filter," *Complexity*, vol. 2018, 2018.

[82] S. P. Sira, D. Cochran, A. Papandreou-Suppappola, D. Morrell, W. Moran, S. D. Howard, and R. Calderbank, "Adaptive waveform design for improved detection of low-rcs targets in heavy sea clutter," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 56–66, 2007.

[83] L. Zhang, N. Wei, and X. Du, "Waveform design for improved detection of extended targets in sea clutter," *Sensors*, vol. 19, no. 18, p. 3957, 2019.

[84] N. F. Ghalyan and A. Ray, "Measure invariance of ergodic symbolic systems for low-delay detection of anomalous events," *Mechanical Systems and Signal Processing*, vol. 159, p. 107746, 2021.

[85] Y. Rajabzadeh, A. H. Rezaie, and H. Amindavar, "A dynamic modeling approach for anomaly detection using stochastic differential equations," *Digital Signal Processing*, vol. 54, pp. 1–11, 2016.

[86] S. Qian and C.-A. Chou, "A koopman-operator-theoretical approach for anomaly recognition and detection of multi-variate eeg system," *Biomedical Signal Processing and Control*, vol. 69, p. 102911, 2021.

[87] M. Reid and B. Simon, "Functional analysis (methods of modern mathematical physics)," *Perspectives in Mathematics, Academic Press, Boston*, 1980.

Vita

Caden J. Pici

Education

The Pennsylvania State University, University Park, PA, USA

Ph.D. Electrical Engineering,	August 2023
M.A. Mathematics,	August 2022
M.S. Electrical Engineering,	December 2020
B.S. Electrical Engineering with minor in Physics,	May 2017

Work Experience

U.S. Naval Research Laboratory, Washington, DC, USA

Research Scientist July 2022 - Present

U.S. Naval Research Laboratory, Washington, DC, USA

NREIP Program Student Intern June 2019 - August 2019

Academic Research Experience

U.S. Naval Research Laboratory, Washington, DC, USA

Karle Fellowship July 2020 - July 2022

The Pennsylvania State University, State College, PA, USA

Graduate Research Assistant August 2017 - May 2022

Publications

Journal Papers

- C. J. Pici, S. Kompella and R. M. Narayanan, "A Data-Driven Approach to Modeling Sea Clutter Stochastic Differential Equations" in preparation.
- C. J. Pici, S. Kompella and R. M. Narayanan, "Radar Target Classification Receiver Using Sparse Regression and Target Tailored Matched Filters," in IEEE Transactions on Aerospace and Electronic Systems, vol. 59, no. 1, pp. 184-195, Feb. 2023, doi: 10.1109/TAES.2022.3187387.