

The Pennsylvania State University
The Graduate School
College of Engineering

**COMPRESSIVELY SAMPLED RADAR USING RANDOM
WAVEFORMS**

A Dissertation in
Electrical Engineering
by
Mahesh C. Shastry

© 2013 Mahesh C. Shastry

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2013

The dissertation of Mahesh C. Shastry was reviewed and approved* by the following:

Ram M. Narayanan
Professor of Electrical Engineering
Dissertation Adviser, Chair of Committee

Vishal Monga
Assistant Professor of Electrical Engineering

John F. Doherty
Professor of Electrical Engineering

James Brannick
Assistant Professor of Mathematics

Kultegin Aydin
Department Head, Electrical Engineering

*Signatures are on file in the Graduate School.

Abstract

Compressive sensing (also referred to as compressed sensing) refers to the theory and practice of exploiting sparsity in physical measurements to acquire fewer samples than dictated by the conventional Shannon-Nyquist-Kotelnikov sampling theorem. In this thesis, we explore the utility of compressive sensing in radar imaging problems. In order to achieve good imaging resolution, we seek to build radar systems with hardware and algorithms intended for processing signals of high bandwidth. Traditionally, radar systems utilized analog processing systems. Over the last few years, with advances in computing, it has been possible to process ultra-wideband (UWB) signals digitally. The current state of analog-to-digital converter (ADC) technology limits our ability to effectively acquire UWB radio-frequency (RF) signals in the context of radar imaging. However, radar signals that are scattered from common target scenes are sparse when represented in appropriate basis functions. We can thus apply the theory of compressive sensing to circumvent the limitations of ADC technology and design radar systems capable of imaging at higher resolutions than conventionally thought possible.

In this thesis, we study the utility of random, noise-like transmit waveforms in compressive radar imaging. Our focus is on developing the theory and methods for the basic radar signal processing tasks of imaging, detection, and waveform design in the framework of compressive sensing. Compressive radar signal processing systems are more complex and less robust to noise and perturbations than conventional systems. We demonstrate using phase-transition diagrams that compressive noise radar is a feasible technology. Further, phase transition diagrams can be used for calibrating radar systems in real applications. The original contributions of this thesis are in developing the theory of compressive radar

imaging using random stochastic waveforms. We demonstrate that compressive sensing works in real scenarios by applying it to experimental ultra-wideband noise radar signals. Hypothesis testing for target detection in the context of compressive radar imaging is different than the conventional setting since signal recovery is iterative and non-linear. We propose a method based on extreme value statistics to characterize the detection performance and determine thresholds for target detection in compressive radar imaging. We validate the effectiveness of this approach by applying it to experimental noise-radar imaging data. Radar systems often require designing waveforms that are optimally suited for specific applications and target scenarios. However, optimizing the transmit waveform can make the system matrix unsuitable for compressive signal recovery. In our work, we propose a waveform design algorithm that optimizes the waveform while keeping the system matrix suitable for compressive sensing. We analyze its performance and demonstrate its effectiveness through numerical simulations.

Contents

List of Figures	viii
List of Tables	xii
List of Symbols	xiii
Acknowledgments	xv
1 Introduction	1
1.1 Radar Imaging	3
1.2 Introduction to Compressive Sensing	6
1.2.1 Restricted Isometry Property	7
1.2.2 Mutual Coherence	8
1.2.3 Phase Transitions	8
1.2.4 Recovery Algorithms	9
1.2.5 Applications	10
1.3 Structure of the Thesis and Summary of Contributions	11
2 Compressive Sensing for Radar Imaging	13
2.1 State of the Art in Compressive Radar Imaging	13
2.2 Basics of Compressive Stochastic Waveform Radar	16
2.2.1 Compressive Radar	16
2.2.2 Performance of Compressive Sensing	18
2.3 Sensitivity of the BPDN algorithm	20

2.4	Phase Transitions in Radar Imaging	22
2.4.1	Uniform Norm of Estimation Error in Radar Imaging	25
2.4.2	l_∞ Norm in Radar Imaging via Least Squares Estimation	27
2.5	l_∞ Norm in Compressive Radar Imaging	29
2.6	Correlations in the Circulant Matrix	30
2.7	Experiments	31
2.8	Analysis of Experimental Data	32
2.8.1	Characterization of Waveforms	32
2.8.2	Measure of Performance	36
2.8.3	Imaging Performance	38
3	Detection Strategies for Compressive Noise Radar	45
3.1	Compressive Sensing Detection	45
3.1.1	Compressive Sensing Detection: An Empirical Examination	47
3.2	Detection Algorithms	48
3.2.1	Thresholding Based on Extreme Value Theory	48
3.3	Statistics of the Error of Compressive Signal Recovery	49
3.4	Algorithm for Estimating the Thresholds for Compressive Detection	49
3.4.1	Threshold Estimation for Compressive Detection	52
3.4.2	GPD and Compressive Sensing	55
3.4.3	Performance of GPD-based Threshold Estimation	56
3.4.4	Performance of Algorithm on Experimental Data	58
3.5	Computational Complexity	59
4	Waveform Design and Compressively Sampled Radar	63
4.1	The Need for Designing Waveforms	63
4.2	Deterministic Nonidealities in Compressive Radar	68
4.2.1	Partial Correlated Random Matrices	68
4.2.2	Degradation in Restricted Isometry Constants	71
4.2.3	RIP for Circulant Matrices	71

4.3	Waveform Design Algorithm	75
4.3.1	Performance of Compressive Recovery with Optimized Transmit Waveform	77
4.4	Simulations	78
5	Conclusions and Future Work	89
5.1	Conclusions	89
5.1.1	Compressive Noise Radar Imaging and Detection	89
5.2	Open Problems	90
5.2.1	Waveform Design	90
5.2.2	Compressive Noise Radar Hardware	93
	Bibliography	93
	Appendix Mathematical Derivations	101
A.1	Compressive Signal Recovery: A Model-Selection Perspective	101

List of Figures

1.1	Figure showing trends in the progress of ADC technology since 1975 to 2010, borrowed from Jonsson [1]. ENOB stands for <i>effective number of bits</i> and refers to the achievable number of quantization levels.	4
2.1	Sensitivity of the l_2 norm of the estimation error to mismatch in σ for $\delta = 25\%$ and $\rho = 10\%$	22
2.2	Sensitivity of the l_∞ norm of the estimation error to mismatch in σ with for $\delta = 25\%$ and $\rho = 10\%$	23
2.3	Phase transition diagram for a Gaussian random circulant matrix with signal to noise ratio of 20 dB and random under-sampling.	25
2.4	Phase transition diagram for a Gaussian random circulant matrix with signal to noise ratio of 20 dB and uniform under-sampling.	25
2.5	Phase transitions at low SNRs. (a) SNR = 10 dB, and (b) SNR = 0 dB . .	26
2.6	Least squares uniform norm of estimation residue for different matrices of comparable conditioning number at 20 dB SNR.	28
2.7	Compressive sensing simulations showing the dependence of the uniform norm error on the row-wise sum of elements of the system matrix. 20 dB SNR.	29
2.8	20 dB SNR. Phase transition plot on a 32×32 grid on $[0, 1] \times [0, 1]$ for noise radar imaging, with \mathbf{A} as partial random Toeplitz.	30
2.9	The effect of filtering on the TPSF. On the y-axis is $\chi\mathbf{G}$. It is seen that the performance of compressive sensing is expected to deteriorate if the random transmit waveforms are highly correlated (narrow filters).	32
2.10	Photograph of experimental setup.	33

2.11	Close-up photograph of the corner reflector target.	34
2.12	Power spectrum estimates (using a covariance estimator) of (a) mm wave transmit waveform, (b) S-band transmit waveform.	35
2.13	(a) QQ-plot of the normalized transmit waveform for mm wave radar in comparison with standard normal, (b) QQ-plot of the normalized transmit waveform for an s-band radar in comparison with standard normal	36
2.14	The error metric $rPSN$ for different sampling rates. An indication of the phenomenon of phase transition is seen in this plot. Up to a sampling rate of around 10%, recovery is near impossible. However with more than 10% of the samples we can fully recovery the target information.	38
2.15	Corner reflector at a distance of 100 ft (millimeter wave radar)	40
2.16	Two cylindrical targets at a distance of about 40 feet from the radar, and separated from each other by 2 foot.	41
2.17	Two cylindrical targets at a distance of about 40 feet from the radar, and separated from each other by 1 foot.	42
2.18	Two corner reflectors separated by a distance of 7 feet, located at a distance of 100 ft from the radar.	43
2.19	Two corner reflectors separated by a distance of 3 feet, located at a distance of 100 ft from the radar.	44
3.1	The probability of error (red curve) and false alarm (blue curve) plotted as a function of the rate of fraction of samples acquired $SNR = 20dB$	48
3.2	The receiver operating curves for different values of signal to noise ratio, with $(\delta = 25\%, \rho = 10\%)$	48
3.3	Quantile-quantile plot of synthetic normalized data. Blue markers represent the statistics of the points that satisfy \hat{s}_i with i such that $s_i = 0$. Since the blue markers deviate significantly from the red line, the data cannot be modeled as Gaussian.	50

3.4	Quantile-quantile plots of real normalized data. Blue markers represent the statistics of the points that satisfy \hat{s}_i with i such that $s_i = 0$. If the data are distributed according to the standard normal distribution, then the blue markers coincide with the red line.	51
3.5	Plot showing application of Algorithm 1 for GPD-based extrapolation of the cumulative distribution of empirical residue data. Compressive recovery with $N = 1024$	57
3.6	The probability distribution of the estimates of the threshold value for various desired values of P_{FA} . Compressive recovery was performed with $S = 10$, $M = 256$, $N = 1024$ and SNR = 10 dB. The estimate of the P_{FA} was done using $\alpha = 0.01$. The pdf of the threshold estimation itself was generated using 50 instances of algorithm 1.	58
3.7	Tail estimation for real data. We observe that the empirical CDF corresponds closely with the GPD estimate for the tail. The target scenario involved one corner reflector target at 100 ft imaged multiple times using the millimeter wave radar described in Chapter 3.	59
3.8	Computational complexity of SPGL 1 convex optimization signal recovery in terms of running time of algorithm on an Intel i7 2.8 GHz with 8 GB of RAM. 62	
4.1	Examples of specific basis functions used in Section 4.4. These are the columns of the matrix Ψ . The index j refers to the index of the column of the basis within the matrix $\Psi^{(k)}$	79
4.2	CDF plots comparing the inter-column coherence of system matrices corresponding to naive random and optimized waveforms.	80
4.3	A typical transmit waveform, optimized in this simulation for a dictionary of rectangular basis functions and the identity matrix, with $D = 11$	81
4.4	Target recovery performance (left- $\Psi\hat{\mathbf{s}}$, right- $\hat{\mathbf{s}}$) with naively random waveforms and waveforms optimized for the dictionary model generated from rectangular window basis function.	82

4.5	Target recovery performance (left- $\Psi\hat{\mathbf{s}}$, right- $\hat{\mathbf{s}}$) with naively random waveforms and waveforms optimized for the dictionary model generated from a Gaussian window basis function.	83
4.6	Target recovery performance (left- $\Psi\hat{\mathbf{s}}$, right- $\hat{\mathbf{s}}$) with naively random waveforms and waveforms optimized for the dictionary model generated from Gaussian derivative-wavelet basis function.	83
4.7	Error metric ξ for 20 realizations of the rectangular bases function case, $\nu = 0.1$.	84
4.8	Ground truth for simulated target scene, targets representable in basis set of rectangular functions.	84
4.9	Compressive imaging using naive random waveforms, targets representable in basis set of rectangular functions.	85
4.10	Compressive imaging using optimized random waveforms, targets representable in basis set of rectangular functions.	85
4.11	Ground truth for simulated image, targets representable in basis set of Gaussian functions.	86
4.12	Compressive imaging using naive random waveforms, targets representable in basis set of Gaussian functions.	86
4.13	Compressive imaging using optimized random waveforms, targets representable in basis set of Gaussian functions.	87
4.14	Ground truth for simulated image, targets representable in basis set of Gaussian wavelets.	87
4.15	Compressive imaging using naive random waveforms, targets representable in basis set of Gaussian wavelets.	88
4.16	Compressive imaging using optimized random waveforms, targets representable in basis set of Gaussian wavelets.	88

List of Tables

3.1	Table summarizing the Kolmogorov-Smirnov test comparing a few instances of the normalized data with the standard normal distribution.	50
-----	---	----

List of Symbols

$x(t)$	Radar transmit waveform (continuous)
$s(t)$	Radar target scene model (continuous)
$y(t)$	Fully sampled receive waveform (continuous)
$\eta(t)$	Additive noise (continuous)
$\#\{\dots\}$	Cardinality of the set.
$\mathbf{x} \in \mathbb{R}^N$	Radar transmit waveform (discrete)
$\mathbf{s} \in \mathbb{R}^N$	Radar target scene model (discrete)
$\mathbf{y} \in \mathbb{R}^N$	Fully sampled receive waveform (discrete)
$\boldsymbol{\eta} \in \mathbb{R}^N$	Additive noise (discrete)
$\mathbf{z} \in \mathbb{R}^M$	Discrete undersampled receive signal.
$\mathbf{X} \in \mathbb{R}^{N \times N}$	Circulant matrix generated from \mathbf{x} .
$\mathbf{R}_\Omega \in \mathbb{R}^{M \times N}$	Restriction matrix representing undersampling operation.
S	Number of non-zero entries in the vector \mathbf{s}
$\delta = M/N$	Undersampling rate
$\rho = S/M$	Sparsity ratio
$[N]$	Set of all non-zero integers up to and including N .
P_{fa}	Probability of false alarm.
P_d	Probability of detection.
γ, ζ	Parameters of Generalized Pareto Distribution
μ	Mutual coherence
$F(\cdot)$	Cumulative distribution function

$\Psi \in \mathbb{R}^{N \times DN}$	Dictionary matrix
\mathbf{S}	Permutation matrix
\mathbf{F}	Discrete Fourier Transform Matrix.
$\mathbf{\Lambda}^{(i)}$	Eigenvalue matrices corresponding to $\Psi^{(i)}$.
$\hat{\cdot}$	Estimated quantity
T	Transpose of a matrix
H	Conjugate transpose of a matrix
$\ \cdot\ _{l_p}$	l_p -norm
$\ \cdot\ _p$	l_p -norm
$ \cdot $	Absolute value

Acknowledgments

I would like to thank my adviser, Professor Ram Naryanan for motivating and guiding me through my formative years as a researcher. His support, advice, and inspiration have all been invaluable. I gratefully acknowledge the insightful guidance provided by our collaborator, Dr. Muralidhar Rangaswamy from the Air Force Research Lab in WPAFB, Ohio. I would like to thank Professor James Brannick for discussions on multi-grid methods and the computational aspects of compressive sensing. I thank my committee members, Professors John Doherty and Vishal Monga for their valuable comments. Our research was supported by the Air Force Office of Scientific Research (AFOSR) contract number FA9550-09-1-0605. I would like to thank Dr. Jon Sjogren of AFOSR for his valuable suggestions related to our work. I am grateful for the great time I spent with fellow Raiders in Professor Narayanan's lab. I would like to acknowledge my productive collaborations with Yangsoo Kwon. I am deeply indebted to Sonny Smith and Kyle Gallagher for their extraordinary help in conducting experiments in freezing January weather to experimentally validate compressive noise radar imaging.

There are so many people to thank, friends and family, for helping me along in my quest to extend the frontiers of knowledge. Words are not sufficient to express my gratitude for my parents and my brother. They were primarily responsible for inspiring me to pursue research. I would like to thank my wife for her love and support through the trials and tribulations of grad school.

Chapter 1

Introduction

A desirable goal of research in imaging is to design systems that can *do more with less*. In more concrete terms, our objective is to maximize the information we can extract from physical signals while minimizing the cost of measurement, acquisition, and processing. Before we embark on a review of the state of the art, we address the question of whether measurement, acquisition, and processing represent different paradigms. Measurement of physical signals involves understanding the underlying physical phenomena and constructing hardware to measure physical quantities. Optical imaging involves measuring using optical sensors, the intensity of light at different points in space and time. Radio-frequency (RF) signal acquisition involves measuring surface currents induced in antennas by electromagnetic fields. Magnetic resonance imaging involves measuring radio-frequency signals emitted by excited water molecules returning to their ground states. These signals are then sampled and quantized for discrete operations. The final step in modern systems is to store and process these data to extract meaningful information and context about the real world. The conventional approach to imaging involves treating the problem of improving system performance as two distinct sub-problems: (1) designing efficient hardware for measurement and acquisition and (2) designing efficient algorithms for processing the discrete data. Over the last two decades, new results in electrical engineering have made this distinction less clear.

Compressive sensing [2] refers to the idea of making use of the mathematical structure

of a signal to reduce the burden on the hardware used for measurement and acquisition. Traditional sensing systems worked by first sampling a sensed physical signal, storing the entire record, and then utilizing the mathematical structure for efficient processing, storage, transmission, and visualization. The primary reason why we can do this efficiently is that signals are structured. Mathematically speaking, the data acquired by such sensors are highly redundant. We refer to this concept as *sparsity*. The useful information contained in a signal is much smaller in dimension compared to the apparent dimension of the signal itself. Compressive sensing provides the theoretical and algorithmic framework for simultaneously *compressing* and *sensing* the signal. It opens up the possibility of non-adaptively acquiring fewer data than before, while not compromising on the amount of useful information that we can extract. The hope is that by efficiently implementing compressive sensing, one can reduce the expensive burden of acquisition and measurement.

What are the specific challenges in modern systems that necessitate the deployment of compressive sensing? Commentators in news media and elsewhere call our times the era of ‘Big Data’. In the context of signal processing, this ‘data deluge’ [3] is largely due to the ubiquity of sensor systems that measure electromagnetic and physical waves. The burden on acquisition and processing systems is also proportionally larger. In some imaging problems, the information relevant to the application lies in a sub-space of smaller dimension than the acquired data [3]. In compressive sensing, the amount of data measured by sensors is decided by the information content, rather than the frequency content. In this context of *data deluge*, we can hope to do more with our signal processing systems without necessarily increasing the burden on sensors. The idea of shifting the burden from physical sensors to computational algorithms has exciting implications for modern systems. We can potentially design portable and energy efficient sensor systems and sampling schemes for applications in radar imaging, digital communications, infrared imaging, acoustic imaging, and optical sensing.

In our work, we present new results on compressive sensing for signal processing in the context of linear inverse problems that occur in radar imaging. Specifically we propose to use noise radar technology for compressive radar imaging. We outline the specific prop-

erties of noise radar systems that makes them suitable for building compressively sensed radar. While we can reduce the burden on sensors, deploying compressive sensing remains a computationally expensive prospect. The important problem of characterizing the effect of noisy signals in compressive sensing also remains open. In this thesis, we attempt to tackle the first problem by proposing an efficient approach based on extreme value statistics for formalizing the task of target detection (binary decision). We then explore the possibility of using compressive sensing principles to enhance the performance of radars in typical applications. Based on redundant dictionaries, we develop models to leverage supplementary information to improve the performance of radar imaging. The algorithms we propose enable significant improvements to the system performance at little extra cost. Our research opens up new possibilities for exploiting sparsity in practical sensor systems to ‘do more with less’.

1.1 Radar Imaging

In this thesis, we look at radar imaging as a linear inverse problem. The idea of using directed radio frequency electromagnetic radiation for imaging applications was first proposed during World War Two. In a typical radar system, an electromagnetic wave is transmitted into free space using a transmitting antenna and the reflections are measured by a receiving antenna. In this thesis, we use the terms ‘target scene’ and ‘target environment’ to refer to the region of space in which unknown targets are present. The terms ‘transmitted waveform’ and ‘received’ or ‘reflected’ waveform refer to the time series represented by the transmitted and reflected fields respectively. The received and transmitted waveforms are discretized for processing, and the discrete locations of targets in each scene are referred to as ‘cells’.

The most basic and widespread use of radar systems is in estimating range profiles of the target scene, velocity of targets, and in performing target detection tasks. Conventional systems have employed techniques such as matched filtering, synthetic aperture radar, inverse synthetic aperture radar, and threshold detection [4] for these tasks. These approaches to radar imaging have proved quite effective for basic tasks. Conventional radar systems used

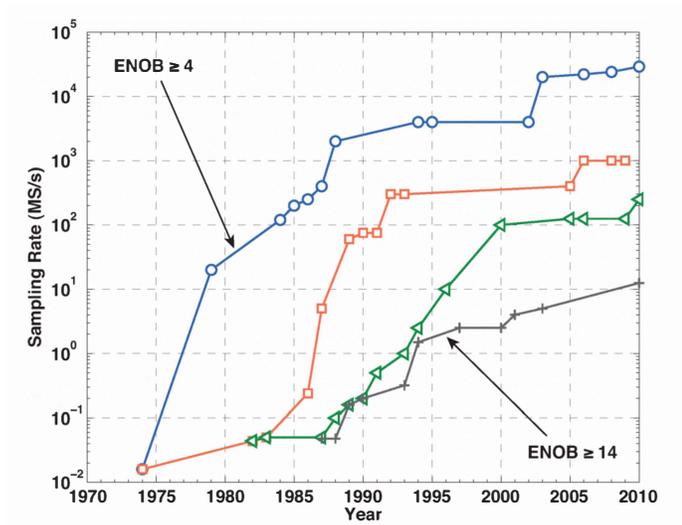


Figure 1.1: Figure showing trends in the progress of ADC technology since 1975 to 2010, borrowed from Jonsson [1]. ENOB stands for *effective number of bits* and refers to the achievable number of quantization levels.

analog processing systems for the most part. One of the main reasons for this is the desire to maximize the bandwidth of the transmit signals. In this thesis, we concern ourselves primarily with imaging problems that involve building range profiles. A high bandwidth signal is crucial to achieving a high range resolution. For most of the last six decades, the absence of efficient sampling hardware and digital processing limited the adoption of digital processing for radar systems. Since the 1990s [5], digital radar receivers are slowly becoming pervasive in radar imaging applications. This is largely thanks to the development of efficient high-rate analog-to-digital convertors (ADCs) and digital signal processing hardware. However, ADC technology has advanced at a much slower pace than the growth in computational capabilities. The move to digital systems has enabled us to incorporate ever more advanced signal processing techniques at the receivers to improve the functionality of radar systems. The trends in the evolution of analog-to-digital converter technology from 1975 to 2010 is shown in Figure 1.1 (the figure has been borrowed from a 2010 survey by Jonsson [1]). For example, the best 12-bit quantization ADCs available today can sample at a rate of around 2 gigasamples per second. The development of high resolution radar technology is hindered by the cost of sampling the analog signal.

Noise radar technology involves transmitting random-noise like continuous waveforms. The received signals are incoherently processed using a matched filter for range detection. Noise radar imaging was first proposed around fifty years ago in [6]. Over the years [7], the trend has been towards the adoption of *ultra-wideband* signals for noise radar. Ultra-wideband waveforms are defined as signals that have a bandwidth of either 500MHz or higher or 20% of the center frequency. In our applications, we use waveforms of 500MHz bandwidth. The ultrawideband nature of the waveform enables us to achieve a high range resolution. The basic ranging problem involves modeling the target scene as a linear filter with transfer function, $s(t)$, so that, for a given transmit signal $x(t)$ and additive noise $\eta(t)$, the reflected signal is simply given by

$$y(t) = x(t) * s(t) + \eta(t), \quad (1.1)$$

where $*$ represents the operation of linear convolution. In the past, noise radar signals have mainly been processed using a matched filter. Matched filtering based target recovery is premised on the fact that delayed versions of the transmitted signal are orthogonal to each other. If we discretize the signal model, we obtain,

$$\mathbf{y} = \mathbf{X}\mathbf{s} + \boldsymbol{\eta}, \quad (1.2)$$

where $\mathbf{y}, \mathbf{s}, \boldsymbol{\eta} \in \mathbb{R}^N$ and $\mathbf{X} \in \mathbb{R}^{N \times N}$. With the orthogonality assumption the recovered target is given by

$$\mathbf{s} = \mathbf{X}^{-1}\mathbf{y} = \mathbf{X}^T \mathbf{y}. \quad (1.3)$$

The advantage of using the matched filter is that its implementation is efficient in terms of cost and processing latency. In order to extract the target image, all we need to do is to compute the cross-correlation between the reflected waveform and the transmitted waveform.

In this thesis, we propose to use a least-squares based approach to noise radar imaging. We extend the signal model such that, $\mathbf{X} \in \mathbb{R}^{M \times N}$ and we cast the problem as one

of minimizing a least squares-based cost function such as $\min_{\mathbf{s}} \|\mathbf{z} - \mathbf{X}\mathbf{s}\|_{l_2}$. While least squares processing can potentially be more expensive computationally than correlation processing and more difficult to process in real time, such a model enables us to improve the functionality of the radar system. As we show in Chapter 2, with the least-squares formulation, we can utilize sparsity to more efficiently sample the signals. Least-squares and sparsity-based approach are harder to analyze and target recovery involves significant computational cost. In Chapter 3, we propose to overcome these problems with the help of an efficient data-driven detection algorithm. We strengthen our theoretical and empirical claims in Chapter 2 and 3 by testing them on experimental noise-radar data. In Chapter 4, we leverage the least-squares formulation for designing an algorithm that enables us to build a knowledge-based radar system.

1.2 Introduction to Compressive Sensing

Compressive sensing enables us to recover signals from asymptotically fewer samples than traditional approaches mandate. This assertion has interesting consequences to solving estimation problems. The compressive sensing problem that we consider in this thesis involves solving the linear system given by

$$\mathbf{z} = \mathbf{A}\mathbf{s} + \boldsymbol{\eta}, \tag{1.4}$$

where, we assume $\mathbf{z} \in \mathbb{R}^M$, $\mathbf{A} \in \mathbb{R}^{M \times N}$, and $\mathbf{s} \in \mathbb{R}^N$. In the absence of any information about the vector \mathbf{s} , for a unique solution to exist, the matrix \mathbf{A} should have at least as many rows as the dimension of \mathbf{s} . In other words, the system should be fully determined or overdetermined, with $M \geq N$. An important question in many applications is whether we can find a unique solution when $M \leq N$. Compressive sensing asserts that when \mathbf{s} is *sparse*, underdetermined systems have unique and stable solutions, under certain conditions. The basic approach to reconstruction involves looking for the sparsest solution to this linear system. We follow the convention used in signal processing to represent the sparsity of a vector as the so-called l_0 norm (which is not technically a norm). $\|\mathbf{u}\|_{l_0} = \#\{i : u_i \neq 0\}$.

We can then cast the inversion problem as one of minimizing the l_0 -norm

$$\hat{\mathbf{s}} = \arg \min \|\mathbf{s}\|_{l_0} \text{ s.t. } \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \leq \sigma. \quad (1.5)$$

While this formulation is theoretically correct, it is of little practical use. Minimizing the l_0 norm is an NP-hard combinatorial optimization problem. The main theoretical result that enables the application of compressive sensing is that we can replace the l_0 -norm cost function with the tractable l_1 -norm. This results in the following optimization problem

$$\hat{\mathbf{s}} = \arg \min \|\mathbf{s}\|_{l_1} \text{ s.t. } \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \leq \sigma. \quad (1.6)$$

The intuitive reason behind using the l_1 norm is that it better approximates the attribute of sparsity than other cost functions. Further, l_1 optimization problems are easy to solve using techniques from the mature fields of linear programming and convex optimization. There are three important factors that decide whether this l_1 -optimization problem results in the sparsest solution. The first is the *sparsity*, S of the solution as defined by the number of non-zero elements in the vector \mathbf{s} . Secondly, the stability and accuracy of the solution is dependent on the *noise*, $\boldsymbol{\eta}$ and our estimate of the noise content, as represented by the parameter, σ . The third and arguably, the most important criterion from a systems-design perspective, are certain properties of the system matrix \mathbf{A} . These properties are in some sense surrogates for the much desirable property of orthogonality which applies for square matrices. A discussion of two such properties follows.

1.2.1 Restricted Isometry Property

For each S , the *isometry constant* δ_S of a matrix \mathbf{A} , is defined as the smallest number that satisfies the property that for all vectors \mathbf{u} with sparsity S [2],

$$(1 - \delta_S)\|\mathbf{u}\|_{l_2} \leq \|\mathbf{A}\mathbf{u}\|_{l_2} \leq (1 + \delta_S)\|\mathbf{u}\|_{l_2}. \quad (1.7)$$

This condition is referred to as the restricted isometry property (RIP). For different values of the constant δ_S , it has been shown that the solution to the optimization problem coincides with the actual solution of the problem. If $\delta_S \leq \sqrt{2} - 1$, then, Candes, et al. [2] derived the bound for the error to be, $\|\hat{\mathbf{s}} - \mathbf{s}\|_{l_2} \leq C_0 \|\hat{\mathbf{s}} - \mathbf{s}_S\|_{l_1} / \sqrt{S} + C_1$ for some constants C_0, C_1 . We present an exposition of this result in Appendix A.1.

1.2.2 Mutual Coherence

If \mathbf{a}_i represent the columns of the matrix \mathbf{A} , then the mutual coherence of the matrix \mathbf{A} is defined as,

$$\mu(\mathbf{A}) \triangleq \max_{i \neq j} \frac{\langle \mathbf{a}_i, \mathbf{a}_j \rangle}{\|\mathbf{a}_i\|_{l_2} \|\mathbf{a}_j\|_{l_2}}. \quad (1.8)$$

This property is more intuitive and easier to visualize and compute than restricted isometry. μ is a measure of the correlations between the different basis functions that comprise the matrix. For compressive signal recovery, we desire the mutual coherence to be as low as possible. In general [8], the solution to the optimization problem is exact with a high probability if for some positive constant, the dimension M satisfies, $M \geq C\mu^2(\mathbf{A})S \log N$. The use of mutual coherence predates the restricted isometry constant and compressive sensing. The mutual coherence is also indicative of whether a linear inverse problem is amenable to greedy pursuit methods [9].

Transform Point Spread Functions (TPSF) is a computationally feasible surrogate to the concept of mutual coherence [10]. We define the matrix $\mathring{\mathbf{A}}$ by normalizing each column with its l_2 norm. Then, the TPSF of a linear system is defined simply as the Gram matrix $\mathbf{G} = \mathring{\mathbf{A}}^T \mathring{\mathbf{A}}$.

1.2.3 Phase Transitions

The concept of *phase transitions* was first introduced by Donoho and Tanner [11] in 2009. In the context of compressive sensing, the term *phase* refers to the phase space mapped out by δ and ρ , which are defined as $\delta = M/N$ and $\rho = S/M$. The *transitions* refer to the

fact that for certain pairs of (δ, ρ) , the system is solvable while for others it is not. This transition indicates whether the dimension of a system is sufficient for recovering vectors of given sparsity. The phase transition approach, however provides us an empirical *post hoc* approach to validating the suitability of compressive sensing for our linear system. The one advantage of using phase transitions comes from the fact that they are universal [11] across different system matrices, particularly random matrices drawn from sub-Gaussian distributions. This universality provides us with a standard to compare against. We know that the a system matrix that is a Gaussian random matrix permits compressive signal recovery. Thus, we can use the phase transition map for random matrices as a benchmark to characterize the performance of our signal model and recovery algorithms.

1.2.4 Recovery Algorithms

Compressive signal recovery algorithms can be classified to be of five types: greedy pursuit methods, convex optimization algorithms [12], message passing algorithms [13], Bayesian methods [14], and non-convex optimization algorithms [15]. The first and second types of algorithms are currently the most widely used. In general, convex optimization methods result in more exact solutions and are more robust to noise than other methods. Message passing approaches, specifically, the approximate message passing algorithm has been theoretically proved to converge for systems with random matrices. They have been shown to be equivalent in performance with convex optimization approaches. Bayesian and non-convex optimization techniques do not currently come with theoretical guarantees of recovery and convergence. Non-convex approaches, which are characterized by non-convex cost functions, are harder to analyze and slower than convex optimization and greedy algorithms. In their survey paper from 2011, Tropp and Wright [12] present a thorough review of the different algorithms in literature.

Greedy pursuit methods include matching pursuit [16], orthogonal matching pursuit [17], compressive sampling orthogonal matching pursuit [18]. These methods work by searching, in an algorithmically *greedy* manner, through a set of basis functions looking for that set of basis functions that best models the vector \mathbf{z} . The presence of perturbations in linear

systems adversely affects the performance of greedy pursuit methods.

Convex optimization approaches are an exact formulation of the compressive signal recovery problem. The basic approach is to solve variants of the l_1 -norm minimization problem given by $\min_{\mathbf{s}} \|\mathbf{s}\|_{l_1}$ subject to $\mathbf{z} = \mathbf{A}\mathbf{s}$. The following formulations of the convex optimization problem are equivalent for some constants σ, λ, τ [19]:

$$\min_{\mathbf{s}} \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} + \lambda \|\mathbf{s}\|_{l_1} \quad (1.9)$$

$$\min_{\mathbf{s}} \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \text{ s.t. } \|\mathbf{s}\|_{l_1} \leq \tau \quad (1.10)$$

$$\min_{\mathbf{s}} \|\mathbf{s}\|_{l_1} \text{ s.t. } \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \leq \sigma. \quad (1.11)$$

There are several approaches for solving the above problems. In convex optimization, these formulations fall under the well-studied classes of linearly constrained quadratic program (LCPQ) and quadratically constrained linear program (QCLP) problems [20]. There are several numerical algorithms for solving such problems. In our thesis we use the spectral projected gradient l_1 (SPGL1) [19] algorithm for solving the optimization problem. There are two advantages to using the SPGL1 algorithm. Firstly, the SPGL1 algorithm is computationally efficient since its iterations do not involve matrix-matrix products, but only matrix-vector products. As a result, SPGL1 scales well. Secondly, the formulation of the problem is natural for electrical engineering applications since the parameter σ simply represents the signal to noise ratio. In practical systems, the signal-to-noise ratio and hence the parameter σ be chosen using any of the several blind SNR estimation techniques in literature [21].

1.2.5 Applications

This dissertation looks at one of the applications of compressive sensing, in radar imaging. Successful recovery is defined as the ability of the algorithm to solve the underdetermined linear system to arrive at a stable and accurate sparse solution. We were successful in testing compressive sensing in a real scenario using experimental data. Since around 2006,

there have been about a couple of thousand papers on compressive sensing. The earliest application to be developed was the single-pixel camera [22]. The single pixel camera utilized spatial sparsity to reduce the number of sensors required for capturing images. The incoherence and randomness that are crucial for compressive sensing were implemented by randomly modulating acquired light intensity. This has potentially important applications in infrared imaging, where sensors are expensive. The problem of magnetic resonance imaging (MRI) involves acquiring images which are sparse in the frequency domain. Time domain undersampling is easy to implement, and hence compressive sensing can prove inexpensive for deployment [10].

Compressive sensing is particularly applicable in applications involving linear inverse problems. In seismic imaging, compressive sensing has been applied to reduce problem dimensions in large-scale inverse problems [23]. Random probes were employed for compressive channel separation in seismic imaging applications by Romberg and Neelamani [24]. Inverse problems also occur in digital communications in the context of channel estimation. The problem of estimating a communications channel is equivalent mathematically to the problem of radar imaging. Paredes, et al. [25] proposed a compressive sensing-based approach for using UWB pulses in channel estimation. Haupt, et al. [26] derived results for the restricted isometry constant of random Toeplitz matrices in the context of channel sensing. Berger, et al. [27] examined the problem of sparse channel estimation in the context of underwater acoustic channels and proposed sparsifying representations and acquisition schemes based on compressive sensing. Since the publication of these three important papers, there have been several new papers proposing novel techniques for improving channel estimation [28, 29, 30, 31].

1.3 Structure of the Thesis and Summary of Contributions

The principal contributions of this thesis are as follows:

1. We formulate compressive sensing as a problem involving the inversion of a linear system with circulant random matrices. In Chapter 2, we use this linear system model to develop the theory of compressive noise radar imaging. We justify the

circulant-matrix model by applying it to experimental noise radar data. We analyze the performance of compressive sensing through theoretical and empirical arguments. For the first time in literature, we experimentally verify the possibility of recovering targets from compressively sampled UWB noise radar. We conducted experiments using a millimeter-wave radar to validate the practicality of compressive noise radar.

2. In compressive noise radar systems, target recovery is achieved by using non-linear convex optimization solvers. This complicates the task of target detection. Under nonlinear recovery, it is difficult to derive theoretical closed form expressions for probabilities of detection and false alarm. The probability of false alarm and the statistics of the recovered vector are necessary to determine detection thresholds. In Chapter 4, we propose a data-driven tail estimation algorithm based on the theory of extreme value statistics. We fit a generalized Pareto distribution to the tail distribution of the detection variables to efficiently derive empirical expressions relating the probability of false alarm and the detection threshold. We test our algorithms on data acquired from experiments with real noise radar systems.
3. In Chapter 4, we consider the problem of incorporating *a priori*, general models of radar target scenes. We propose a knowledge-based waveform design algorithm for adaptive systems. Information about radar target scenes are modeled using redundant dictionaries. Signal models involving redundant dictionaries can be inverted using principles from compressive sensing. However, the redundancy in the signal models may, in general, degrade the performance of compressive signal recovery. We compensate for this degradation by proposing a radar waveform design algorithm that makes the system matrix more suitable for compressive signal recovery.

Chapter 2

Compressive Sensing for Radar Imaging

2.1 State of the Art in Compressive Radar Imaging

Radar systems can be classified to be of two types [32]: (i) those employing an analog system for detecting the targets using an analog matched filter, and (ii) systems that sample and quantize the reflected radar signal and process it digitally. The latter type of radars offer operational flexibility and can be used to acquire a variety of information about target scenes. The resolution of digital radar systems is limited by the ADC technology.

Random noise radar [6] involves transmitting waveforms that are generated as stochastic processes. The technology is headed towards utilizing ultra-wideband transmit waveforms [7] with which, high range resolutions can be obtained. The use of randomly generated waveforms in noise radar makes signals immune to interception and jamming to an extent. Early stochastic waveform radars used analog processing to detect targets [7]. Increasingly noise radar systems are using digital processing [33] for imaging in real time. Digital noise radar systems use high rate analog to digital converters (ADC). The sampling rates of the ADC limits the maximum achievable range resolution of the system. In order to circumvent this bottleneck, we employ compressive sensing principles. Practical compressive noise radar systems have been implemented for ground penetrating radar applications using stepped

frequency waveforms [34, 35]. Bar-Ilan, et al. [36] applied the Xampling [37] framework to develop a pulse radar prototype for joint range-velocity radar imaging using Doppler focusing. Ender [38] presented results on compressive radar imaging performed pulsed chirp waveforms. Our experimental work differs from these in that we use incoherent ultrawideband continuous wave noise radar waveforms for imaging. Continuous wave noise-like waveforms have the advantage of being instantaneous wideband. Further, the waveforms are robust to additive noise, jamming, and interception [7].

The use of random transmit waveforms makes noise radar particularly suitable for compressive sensing. Random waveforms in compressive radar imaging were first suggested by Baraniuk and Steeghs [32] in the context of random demodulators. The recovery performance of compressive sensing estimators depends on the system matrix satisfying certain properties. The two most commonly studied properties are the restricted isometry property (RIP) and mutual coherence [2, 39]. The compressive noise radar imaging problem, as described in this chapter, involves a circulant random system matrix. The suitability of circulant random matrices for compressive sensing is less well studied than the standard case of the random matrix. In the context of *random demodulators*, Romberg [40] showed that with specially designed circulant random system matrices recovery with probability $1 - O(n^{-1})$ is possible with $O(S \log^3 N)$ measurements. Bernoulli random circulant matrices were considered by Haupt, et al [26] in the context of compressive channel estimation. They showed that based on the restricted isometry property of the system matrix, $O(S^2 \log N)$ measurements of the Toeplitz random matrix are sufficient for stable recovery using the Dantzig selector [2] recovery algorithm. Rauhut, et al [41] derived the restricted isometry property for circulant matrices, suggesting that $O(\max((S \log N)^{1.5}, S \log^2 S \log^2 N))$ measurements guarantee stable recovery. Herman and Strohmer derived mutual coherence results for high-resolution radar imaging in [42], and proposed the use of Alltop sequences as transmit waveforms, under the narrowband approximation. They also alluded to the effectiveness of random waveforms in high-resolution radar imaging. Our work extends the state of the art in considering ultrawideband random waveform radar systems which are primarily used for range imaging. An advantage of practical noise radar systems is their simplicity. Our analyses of

practical issues relating to compressive radar imaging are intended to push the field towards real world applications.

Simulations indicate that the number of measurements follows the optimal compressive sensing asymptotics of $O(S \log N)$. In this chapter, we use the phase transitions formalism developed by Donoho, et al. [13] to analyze the behavior of compressive sensing. This chapter is structured to reflect the approach we propose for analyzing general ultrawideband compressive radar systems. In Section 2.2, we formulate the compressive radar imaging problem, and state assumptions and approximations we make to mimic real systems. The convex optimization approaches to compressive signal recovery involve certain assumptions about the sparsity and the SNR to decide on the number of acquisitions required for stable recovery. In this chapter, we use experimental data to validate our results. The main contributions of this chapter as follows:

1. In Section 2.3, we approach basis pursuit denoising as a problem in model selection and study its sensitivity to the accuracy of prior information about the signal-to-noise ratio (SNR) of the system.
2. In Section 2.4, we follow the phase transition formalism to show that in various signal to noise ratio regimes, the behavior of compressive noise radar imaging follows the standard case of compressive sensing with random matrices. We propose using phase transition diagrams as calibration charts for guiding the operation of practical compressive radar systems.
3. In Section 2.7, we outline the experimental set up that we used for testing the theory and methods we developed.
4. In Section 2.8, we analyze the performance of a typical compressively sampled noise radar. We work on an extensive set of experimental data collected using UWB millimeter wave and S band radar systems.

Some results in this chapter have been published in the following papers, (1) Shastry, M. C., Narayanan, R. M., Rangaswamy, M. *Compressive radar imaging using white stochastic waveforms*, In Proc. 5th Int'l Waveform Diversity Design Conf (pp. 90-94), (2) Shastry,

M. C., Kwon, Y., Narayanan, R. M., Rangaswamy, M, *Analysis and design of algorithms for compressive sensing based noise radar systems*, In The 7th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM) 2012, pp. 333-336.

2.2 Basics of Compressive Stochastic Waveform Radar

2.2.1 Compressive Radar

The range estimation problem in radar imaging involves the inversion of a linear system. Let $x(t)$ be the transmit waveform, $s(t)$ denote a target scene, and $\eta(t)$ the additive noise in the system. The reflected waveform $y(t)$ can be modeled as

$$y(t) = \int_{-\infty}^{\infty} x(\tau - t)s(\tau)d\tau + \eta(t). \quad (2.1)$$

The discrete form of this linear convolution problem results in the expression

$$y_n = \sum_k x_{n-k}s_k + \eta_n, \quad (2.2)$$

where, $\mathbf{y}, \mathbf{s} \in \mathbb{R}^N$ and $\mathbf{X} \in \mathbb{R}^{N \times N}$. In real situations, the vector, \mathbf{s} is typically highly sparse. For signals of finite duration, this equation is represented exactly by a linear system with a Toeplitz system matrix $\hat{\mathbf{X}}$, so that $\hat{\mathbf{y}} = \hat{\mathbf{X}}\mathbf{s} + \boldsymbol{\eta}$. We invoke the sparsity of \mathbf{s} to approximate this as a linear system with a circulant system matrix \mathbf{X} , in order to simplify computation and analysis. We use x_k to denote the k^{th} row of \mathbf{X} and \hat{x}_k to denote the k^{th} row of $\hat{\mathbf{X}}$. We define $\boldsymbol{\epsilon}^{circ} = \hat{\mathbf{y}} - \mathbf{y}$ to be the vector representing the error of approximating the linear convolution problem by the circulant matrix system. Let k_j denote the indices where \mathbf{s} is

non-zero. Then, we have

$$\boldsymbol{\epsilon}^{circ} = \sum_{k=1}^N s_k \hat{\boldsymbol{x}}_k - s_k \boldsymbol{x}_k \quad (2.3)$$

$$\epsilon_n^{circ} = \sum_{k=1}^N \hat{x}_{n-k} s_k - x_{n-k} s_k \quad (2.4)$$

$$= \sum_{j=1}^S (\hat{x}_{n-k_j} - x_{n-k_j}) s_{k_j}. \quad (2.5)$$

We note that $S \ll N$. In addition, for some values of n , we have that $\hat{x}_{n-k_j} = x_{n-k_j}$, so that $\boldsymbol{\epsilon} \approx 0$. Alternatively, one can derive the bound on the approximation by considering the matrix-norm of the matrix $\boldsymbol{E} = \hat{\boldsymbol{X}} - \boldsymbol{X}$, so that:

$$\boldsymbol{\epsilon}^{circ} = (\hat{\boldsymbol{X}} - \boldsymbol{X}) \boldsymbol{s} \quad (2.6)$$

$$= \boldsymbol{E} \boldsymbol{s} \quad (2.7)$$

$$\|\boldsymbol{\epsilon}^{circ}\|_2 \leq \|\boldsymbol{E}\|_2 \|\boldsymbol{s}\|_2. \quad (2.8)$$

In sparse problems, $\|\boldsymbol{s}\|_2 \approx 0$. Further, the error is larger for non-zero elements that are situated towards the end of the row. In our experiments the approximation remains accurate because the non-zero elements occur close to the beginning. Physically, this is due to the fact that we use ultra-wideband waveforms, so that even a few microseconds of acquisition results in information related to a much larger range than required. The error in recovering accurately at the ends of the signal is intrinsic to processing signals of finite time duration [43]. However, this type of error can only be overcome only by increasing the dimensionality of our problem. Experimental justification for this approximation is illustrated by the accuracy of the experimental results described in Chapter 2.

In sparse target scenarios, compressive sensing theory allows us the luxury to under-sample $y(t)$. We model this operation of under-sampling as a premultiplication by the matrix $\boldsymbol{R}_\Omega \in \mathbb{R}^{M \times N}$, where \boldsymbol{R}_Ω consists of the rows of the identity matrix indexed by the

set $\Omega \subset [N]$. Thus, we have

$$\mathbf{z} = \mathbf{R}_\Omega(\mathbf{y} + \boldsymbol{\eta}) = \mathbf{R}_\Omega \mathbf{X} \mathbf{s} + \mathbf{R}_\Omega \boldsymbol{\eta}, \quad (2.9)$$

with $\mathbf{z} \in \mathbb{R}^M$. The performance of the compressive radar ranging problem depends on the properties of the matrix $\mathbf{R}_\Omega \mathbf{X} \triangleq \mathbf{A} \in \mathbb{R}^{M \times N}$. We assume that the continuous target scene is discretized into N grid points, with only a small percentage of the *cells* occupied by scattering targets. We justify this assumption with experimental results. The sparsity of the anticipated solution is characterized by the assumption that $\|\mathbf{s}\|_{l_0} \leq S$ with $S \ll N$. We define the quantities $\rho \triangleq S/M$ and $\delta \triangleq M/N$. If the matrix \mathbf{A} satisfies certain properties for vectors of given sparsity, then, the problem above can be inverted by solving the following convex optimization problem, called basis pursuit de-noising (BPDN):

$$\text{BPDN}(\rho, \delta; \sigma) : \min_{\mathbf{s} \in \mathbb{R}^N} \|\mathbf{s}\|_{l_1} \text{ subject to } \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \leq \sigma \quad (2.10)$$

This specific formulation of the problem is chosen over other formulations [44] of compressive sensing because it is useful for practical implementations. Specifically, the value σ has a natural association with the signal to noise ratio of the system. Practical radar systems can measure this either by aiming the radar signal at an area where targets are absent and noting the energy. When there is no opportunity to measure reflected signals in the absence of targets, this can be done by employing an SNR estimation approach [21]. In either case, the radar operator has access to an approximate estimate of the value σ .

2.2.2 Performance of Compressive Sensing

The RIP and mutual-coherence approaches to analyzing compressive sensing offer no exact results about the behavior of the residue of compressive estimation. The properties of the residue of compressive estimation are inferred via inequalities [2] that estimate the upper bound of residual error. Practical applications of compressive sensing require more detailed analyses of the residual error. Analyzing the sensitivity of the estimation error to various non-idealities typical of practical compressive radar imaging is the main motivation of this

chapter.

The state evolution approach [13, 45, 46] offers an empirical and theoretical framework for analyzing the behavior of compressive sensing algorithms in an exact manner. Donoho and Tanner [11] empirically studied the universality of *phase transitions* in compressive sensing. The basic idea of *phase transitions* is to derive the exact asymptotics of various compressive sensing models in the phase space of the parameters (ρ, δ) and the signal-to-noise ratio [47]. Bayati and Montanari [45] provide a way to compute norms of the estimation error when using the approximate message passing algorithm [13] for random measurement matrices. The equivalence between solutions to the approximate message passing algorithm and convex optimization approaches was proved for random Gaussian matrices in [46]. This equivalence provides the theoretical justification for studying the performance of the basis pursuit denoising problem using phase transitions. We show that the phase transitions for circulant matrices are similar to the phase transitions for basis pursuit denoising. The empirical results suggest that the results of [45, 46] also hold when the system matrix is random circulant. The sensitivity of phase transitions to additive noise was studied empirically in [47]. The contributions of our thesis are in presenting phase transition simulations for noisy compressive sensing using circulant matrices for various error metrics. Our use of phase transitions is motivated by their potential application to guide radar engineers in operating real systems. The AMP approach has been applied in a recent paper [48] for determining detection thresholds for constant false alarm rate applications. However, no theoretical guarantees exist for the optimality of the AMP algorithm for circulant and Toeplitz system matrices. We propose a new approach to relating the false alarm probability to the threshold.

The uniform norm for a vector \mathbf{u} is defined as $\|\mathbf{u}\|_{l_\infty} = \max_k |u_k|$. The uniform norm of the compressive estimation error is particularly useful to understanding system performance. There are two reasons for studying uniform norm: (i) the uniform norm is given by extreme value statistics which converge to known distributions in the large system limit, (ii) estimates for the distribution of the uniform norm via extreme value theory can be used as a benchmark for setting thresholds used in determining the presence of

targets in multiple-target scenarios. The literature on the uniform norm of compressive sensing residues is limited. In [49], the authors propose a computable fitness measure for compressive sensing that enables the analysis of the l_∞ norm of the residue. The limit theorems of [45] provide a way to compute the asymptotic behavior of the l_∞ norm in the phase transition/state evolution formalism. We adopt an empirical approach for analyzing the uniform norm in compressive radar imaging.

2.3 Sensitivity of the BPDN algorithm

In practical radar systems, the value of σ is known with some uncertainty. This can be viewed as the uncertainty in prior information possessed by the radar operator concerning the model describing the system. In this section, we examine the sensitivity and dependence of the solution to of BPDN(.) to the mismatch in the estimated and real values of σ . Let $\hat{\sigma}$ represent the estimated value, and σ^* , the actual value. Our goal is to characterize the errors

$$p(\rho, \delta; \sigma) = \|\mathbf{s}_{(\rho, \delta; \sigma)}^* - \mathbf{s}\|_{l_2} \text{ and} \quad (2.11)$$

$$q(\rho, \delta; \sigma) = \|\mathbf{s}_{(\rho, \delta; \sigma)}^* - \mathbf{s}\|_{l_\infty}. \quad (2.12)$$

In order to study the sensitivity of $q(., \sigma)$, we run the SPGL1 algorithm at all points on a discretized $(\sigma, \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2})$. On each point on the discretized grid, we compute the error metrics $p(.)$ and $q(.)$. This is a characterization of the mismatch in the assumed model for the signal and the . The line $\sigma = \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2}$ divides the phase space of $(\sigma, \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2})$ into two parts to reflect the practical problems of over- and under- estimating the values for σ . Clearly, the optimal solution occurs when $\sigma = \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2}$. This is reflected in the occurrence of a minima along the line $\sigma = \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2}$ in Figure 2.1. The interesting results in the simulations of Figure 2.2 concern the sensitivity of $q(., \sigma)$ to σ . It is clear that the uniform norm of the residue is more sensitive to the accuracy of the estimate of σ than the l_2 -norm. Particularly when the value σ is an overestimate of the actual noise variance, $\frac{\partial q}{\partial \sigma}$ is large.

Theorem 1. Let $\tilde{\mathbf{s}}$ and $\hat{\mathbf{s}}$ be the solutions to the problems, $\text{BPDN}(\rho, \delta; \tilde{\sigma})$ and $\text{BPDN}(\rho, \delta; \hat{\sigma})$. Further let us assume that the algorithm converges to vectors $\tilde{\mathbf{s}}$ and $\hat{\mathbf{s}}$ such that $\|\mathbf{z} - \mathbf{A}\tilde{\mathbf{s}}\|_{l_2} = \tilde{\sigma}$ and $\|\mathbf{z} - \mathbf{A}\hat{\mathbf{s}}\|_{l_2} = \hat{\sigma}$, and that the matrix \mathbf{A} satisfies RIP. Then, for RIP constants $\delta_{\tilde{\mathcal{S}}+\hat{\mathcal{S}}}$ the estimation error satisfies,

$$\frac{|\tilde{\sigma} - \hat{\sigma}|}{1 + \delta_{\tilde{\mathcal{S}}+\hat{\mathcal{S}}}} \leq \|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|_{l_2} \leq \sqrt{N}\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|_{l_\infty}. \quad (2.13)$$

Proof. Let the support of $\tilde{\mathbf{s}}$ and $\hat{\mathbf{s}}$ be $\tilde{\mathcal{S}}$ and $\hat{\mathcal{S}}$ respectively. The vector $\tilde{\mathbf{s}} - \hat{\mathbf{s}}$ then has the property $\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|_{l_0} \leq \tilde{\mathcal{S}} + \hat{\mathcal{S}}$. Using the restricted isometry criterion for the matrix \mathbf{A} we have,

$$\begin{aligned} (1 + \delta_{\tilde{\mathcal{S}}+\hat{\mathcal{S}}})\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|_{l_2} &\geq \|\mathbf{A}(\tilde{\mathbf{s}} - \hat{\mathbf{s}})\|_{l_2} \\ &= \|(\mathbf{z} - \mathbf{A}\hat{\mathbf{s}}) - (\mathbf{z} - \mathbf{A}\tilde{\mathbf{s}})\|_{l_2} \\ &= \left| \|\mathbf{z} - \mathbf{A}\hat{\mathbf{s}}\|_{l_2} - \|\mathbf{z} - \mathbf{A}\tilde{\mathbf{s}}\|_{l_2} \right| \\ &= |\hat{\sigma} - \tilde{\sigma}| \leq (1 + \delta_{\tilde{\mathcal{S}}+\hat{\mathcal{S}}})\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|_{l_2}. \end{aligned}$$

The above result follows from the reverse triangle inequality. Further we have for any vector, $\mathbf{u} \in \mathbb{R}^N$, $\|\mathbf{u}\|_{l_2} \leq \sqrt{N}\|\mathbf{u}\|_{l_\infty}$. \square

As a consequence of Theorem 1, we have that a lower bound of the estimation error is determined by the optimization constraint σ in the BPDN problem. Since exact information about the noise estimate σ provides us with the best results for estimation error, any deviation from this value increases the lower bound of the estimation error. This result suggests that it is important to use accurate information about the noise estimate.

The standard approach to solving $\text{BPDN}(\rho, \delta; \sigma)$ is to cast it as the unconstrained quadratic optimization problem, $\min_{\mathbf{s} \in \mathbb{R}^N} \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} + \lambda\|\mathbf{s}\|_{l_1}$. The correspondence between the parameters σ and λ has been argued [19] to be $\lambda = \sigma\sqrt{2\log N}$ in the case where the system matrix is orthogonal. However, in our case, \mathbf{A} is a circulant random matrix, is rectangular, and does not satisfy orthogonality. We use the SPGL1 algorithm, which is designed to solve convex optimization problem of the form given by $\text{BPDN}(\cdot)$. The SPGL1 algorithm

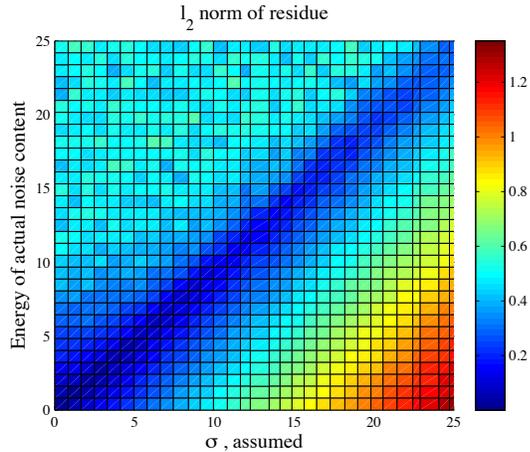


Figure 2.1: Sensitivity of the l_2 norm of the estimation error to mismatch in σ for $\delta = 25\%$ and $\rho = 10\%$.

is a spectral projected gradient algorithm that was proposed to solve the BPDN problem in an efficient manner. SPGL1 involves finding the solution $\text{BPDN}(\rho, \delta; \sigma)$ by repeatedly solving the related optimization problem given by $\text{Lasso}(\rho, \delta; \tau) : \min_{\mathbf{s} \in \mathbb{R}^N} \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2}$ subject to $\|\mathbf{s}\|_{l_1} < \tau$. The goal is to find $\tau^*(\sigma)$, such that the solution of this problem is identical to $\text{BPDN}(\cdot; \sigma)$. The motivation for recasting the problem is that the dual of $\text{Lasso}(\cdot)$ has useful properties. The solution is arrived at by inverting using Newton's method, the scalar equation given by, $\phi(\tau) = \sigma$ where $\phi(\tau) = \|\mathbf{z} - \mathbf{A}\mathbf{s}_\tau\|_{l_2}$. The evaluation of the function $\phi(\tau)$ and its derivative $\phi'(\tau)$ is done using a spectral projected gradient algorithm.

2.4 Phase Transitions in Radar Imaging

We first demonstrate simple simulations concerning the probability of detection for different rates of under-sampling, and the receiver operating characteristics. In our simulations, we used discrete sequences of arbitrary lengths to simulate the performance of the compressive detector, and compressive estimator. Numerically, the difference between the compressive radar imaging system proposed above and other applications of compressive sensing arises due to the fact that the matrix that maps the target scene to the under-sampled reflected signal is a random circulant matrix. The receiver operating characteristics (ROC) resulting

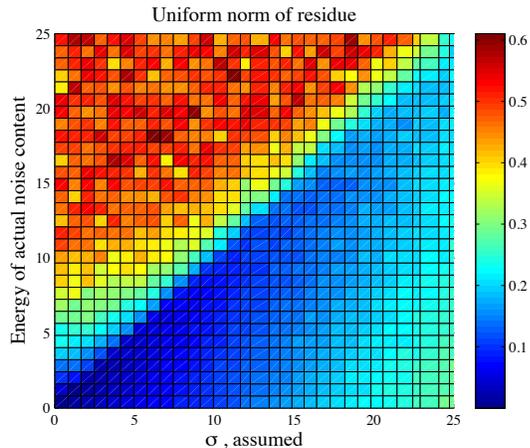


Figure 2.2: Sensitivity of the l_∞ norm of the estimation error to mismatch in σ with for $\delta = 25\%$ and $\rho = 10\%$.

from using circulant random matrices in compressive sensing is analyzed in Chapter 3.

In this section, we examine the behavior of $p(\rho, \delta; \sigma^*)$ and $q(\rho, \delta; \sigma^*)$, where σ^* is the exact estimate of the noise contained in the signal \mathbf{z} . The approach of using phase transitions for studying compressive sensing is well understood for random measurement matrices. In the past, it was postulated that a phase transition structure similar to random measurements also exists for other matrices. In this section, we study phase transition diagrams [11] in the context of compressive sensing to characterize the performance of random circulant matrices. The two sampling schemes we consider are: (1) random undersampling, i.e., choosing the sampling distance in a random manner, (2) uniform undersampling. The uniform undersampling case is more useful for practical systems because of the ease of implementation. Phase transition diagrams present a numerical study of the $l_1 - l_0$ equivalence that forms the basis of compressive sensing. In our simulations, we discretized the (δ, ρ) subspace, and at each point, we computed the error metric averaged over several realizations. The recovery was performed using SPGL1. The signal-to-noise ratio is defined prior to undersampling so that

$$\text{SNR} = 10 \log \frac{\|\mathbf{X}\mathbf{s}\|_{l_2}}{\|\mathbf{z} - \mathbf{X}\mathbf{s}\|_{l_2}}. \quad (2.14)$$

The phase transition diagram of Figure 2.3 was generated by acquiring at random intervals, samples of the reflected signal. The suitability of stochastic waveforms for compressive radar systems is indicated by the fact that the phase transition diagrams are close to the phase transition diagrams that arise from the general compressive sensing setting of acquisition using random measurements [11]. The boundary line between the red and blue areas indicates that as with random measurement matrices, $O(M \log N)$ measurements guarantee good recovery. Similar, desirable phase transitions are seen in Figure 2.4 when the reflected signal is uniformly undersampled. In the context of compressive sensing, phase transition diagrams are a representation of the success of compressive recovery in terms of the fraction of samples correctly recovered as plotted in the phase space of the pair (δ, ρ) where, $\delta = M/N$ corresponds to the number of samples acquired and $\rho = S/M$, is a ratio of the sparsity of the signal and the number of samples acquired.

From a radar systems perspective, phase transition diagrams provide a way of calibrating compressive radar systems. Compressive radar imaging is only non-blind in the sense that the operator of the system needs to have an approximate idea of the sparsity of the target scene. The uncertainty in the knowledge of the approximate sparsity of the target scene can be quantified in terms of the phase transition diagrams. For instance, if a radar system is intended to operate in an environment with 20% of the range cells populated by point scatterers and an ambient SNR is 20 dB, then the radar system can operate with 90% accuracy (as defined by the number of samples of \mathbf{s} recovered accurately) by acquiring 10% to 30% of the total samples, as can be inferred from Figure 2.3. The phase transitions at high SNR are identical to the sharp boundaries reported and shown theoretically for random measurement matrices. The performance of compressive estimation is seen to deteriorate for low SNRs in Figure 2.5 for SNR = 10 dB and SNR = 0 dB. As the noise content in the signal increases, the transition boundary becomes fuzzy. At 0 dB SNR, recovery is only possible when a large number of samples are acquired and that too only for very sparse target scenes.

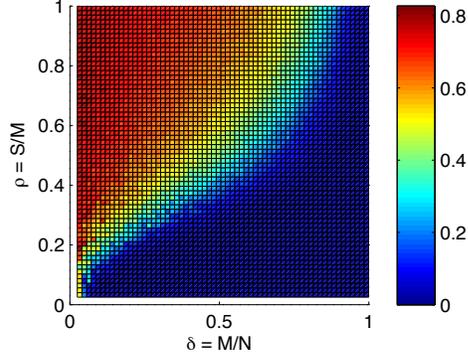


Figure 2.3: Phase transition diagram for a Gaussian random circulant matrix with signal to noise ratio of 20 dB and random under-sampling.

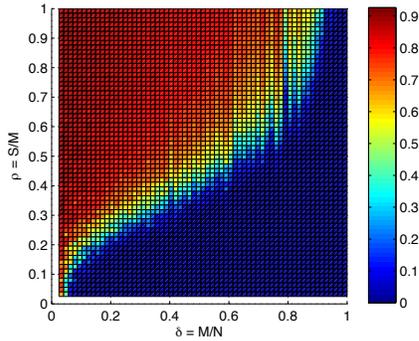


Figure 2.4: Phase transition diagram for a Gaussian random circulant matrix with signal to noise ratio of 20 dB and uniform under-sampling.

2.4.1 Uniform Norm of Estimation Error in Radar Imaging

The l_∞ norm on function spaces is defined as the supremum, i.e., for a function, $f : X \mapsto Y$, $\|f\|_{l_\infty} = \sup_{x \in X} |f(x)|$, for a discrete vector which can be treated as a function $\mathbf{s} : \mathbb{N} \mapsto \mathbb{R}$, this becomes, $\|\mathbf{s}\|_{l_\infty} = \max_{n \in \mathbb{N}} |\mathbf{s}(n)|$. To estimate \mathbf{s} , we need to invert the linear system, $\mathbf{z} = \mathbf{A}\mathbf{s} + \boldsymbol{\eta}_\Omega$ in the relevant paradigm of least squares (LS) or compressive sensing (CS) as

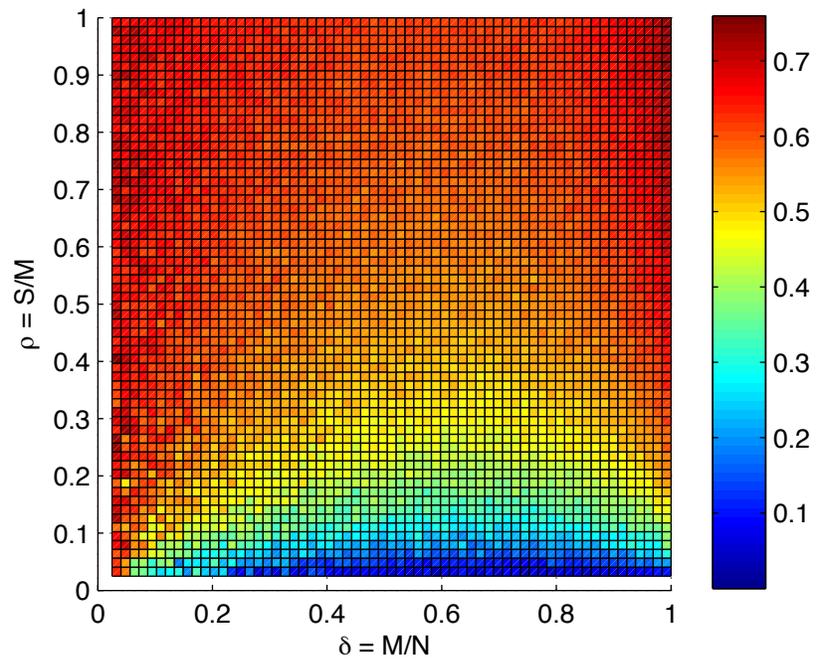
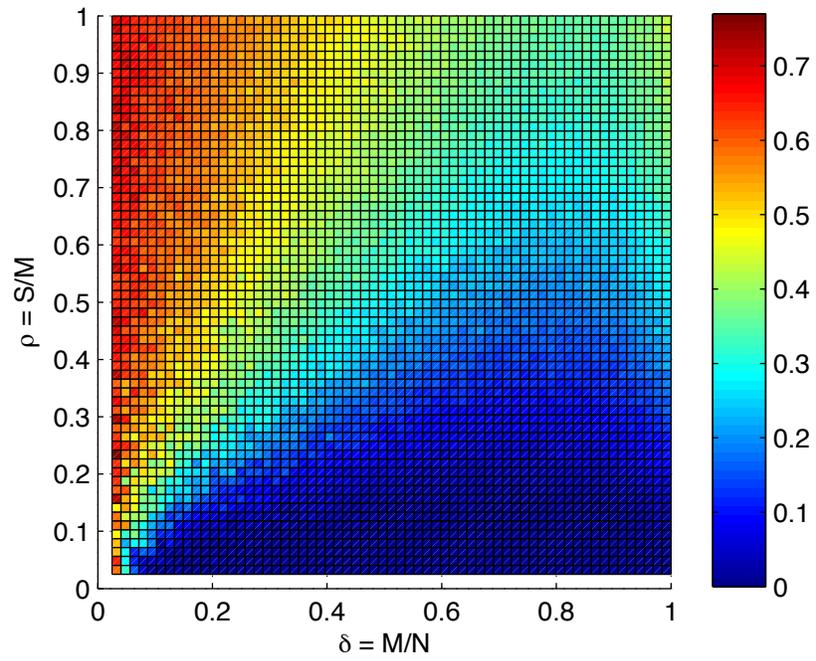


Figure 2.5: Phase transitions at low SNRs. (a) SNR = 10 dB, and (b) SNR = 0 dB

follows:

$$\text{LS: } \mathbf{s}_{ls}^* = \arg \min_{\mathbf{s}} \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \quad (2.15)$$

$$\text{CS: } \mathbf{s}_{cs}^* = \arg \min \|\mathbf{s}\|_{l_1} \text{ s.t. } \|\mathbf{z} - \mathbf{A}\mathbf{s}\|_{l_2} \leq \tau \quad (2.16)$$

We define the error in each case as, $\mathbf{r}_{ls} = \mathbf{s} - \mathbf{s}_{ls}^*$ and $\mathbf{r}_{cs} = \mathbf{s} - \mathbf{s}_{cs}^*$.

Claim: When \mathbf{s} is sparse, the theoretical maximum value of the estimation error is important to radar imaging.

Justification: We express in probabilistic terms, the fact that sparse target scenes contain very few targets. If \mathbf{s} is sparse, then the probability of a target being absent at any target cell is high. We can formalize this with the statement: for some n , $\exists \omega \in (0, 1) : \mathbb{P}[s[n] = 0] > \omega$. On recovery using an unbiased estimator, the cells where the target is absent will contain only random noise. Thus $\exists \omega^* \in (0, 1) : \mathbb{P}[s^*(n) < |\mathbf{r}|_{l_\infty} | s[n] = 0] > \omega^*$. Since the recovery does not assume any specific knowledge about the location of the target, the two events given by $[s(n) = 0]$ and $[s^*(n) < |\mathbf{r}|_{l_\infty}]$ are independent and hence, $\exists 0 \leq \omega^* \leq 1 : \mathbb{P}[s^*(n) < |\mathbf{r}|_{l_\infty}] > \omega^*$. Since, by the hypothesis of sparsity, ω^* is close to 1, it is nearly impossible to recover any target that has reflectivity lower than the theoretical computed l_∞ norm of the estimation residue.

2.4.2 l_∞ Norm in Radar Imaging via Least Squares Estimation

Theorem 2 (Uniform norm of \mathbf{r}_{ls}). *The uniform norm of the least squares residue is a function of the sum of rows of the system matrix \mathbf{A} .*

Proof. We first consider the case where $\mathbf{z} = \mathbf{A}\mathbf{s} + \eta$ is a fully determined (but ill-posed) system of linear equations. The optimal solution to this problem is given by the standard least squares solution, $\mathbf{s}_{ls}^* = \mathbf{A}^\dagger^{-1}\mathbf{z}$, where $\mathbf{A}^\dagger = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is the Moore-Penrose pseudoinverse. The estimation error is given by, $\mathbf{r}_{ls} = \mathbf{s}^* - \mathbf{s} = \mathbf{A}^\dagger^{-1}\eta$. For convenience, we define $\mathbf{B} \triangleq \mathbf{A}^\dagger^{-1}$. Then, $\mathbf{r}_{ls} \sim \mathcal{N}(0, \sigma^2\mathbf{B}'\mathbf{B})$. The uniform norm of the error is given by the

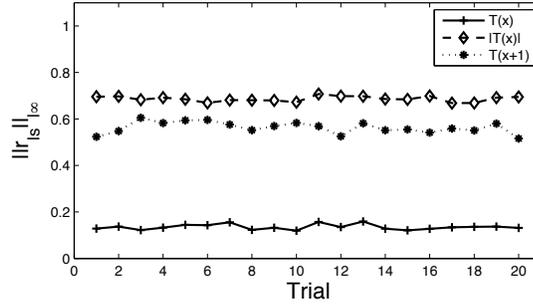


Figure 2.6: Least squares uniform norm of estimation residue for different matrices of comparable conditioning number at 20 dB SNR.

random variable

$$Q \triangleq \|\mathbf{r}_{ls}\|_{l_\infty} = \max_k |r_{ls}(k)| \quad (2.17)$$

$$\mathbb{P}[Q < q] = \mathbb{P}[|r_{ls}(1)| < q, \dots, |r_{ls}(N)| < q] \quad (2.18)$$

$$= \mathbb{P}\left[\bigcap_k \{ |(\mathbf{B}\boldsymbol{\eta})(k)| < q \}\right] \quad (2.19)$$

$$= \mathbb{P}\left[\bigcap_k \{ |\eta(k)| < (\mathbf{A}\mathbf{1})(k)q \}\right], \quad (2.20)$$

where $\mathbf{1} = (1, \dots, 1)'$ is the column vector of 1s. Let us define $c_k = (\mathbf{A}\mathbf{1})(k)$. Since $\{|\eta(k)| < c_k q\}$ are mutually exclusive,

$$F_Q(q) = \prod_k F_{|\eta|}(c_k q) \quad (2.21)$$

$$p_Q(q) = \sum_k \left(\prod_{j \in [N] \setminus k} F_{|\eta|}(c_j q) \right) c_k p_{|n|}(c_k q). \quad (2.22)$$

□

Note that Theorem 2 disregards the numerical properties of the algorithm used to solve the least squares problem. Thus, in a sense, the theorem should be verified numerically for matrices \mathbf{A} with similar conditioning numbers. In Figure 2.6 we demonstrate this effect for noise radar systems, where the system matrix is a random Toeplitz matrix. We denote by

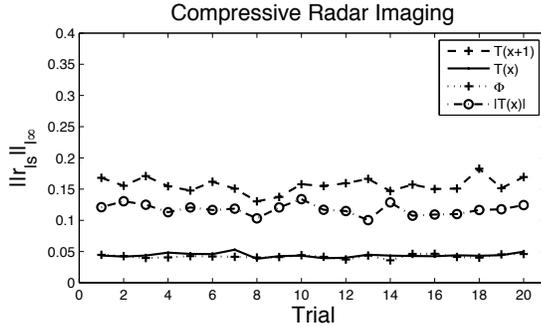


Figure 2.7: Compressive sensing simulations showing the dependence of the uniform norm error on the row-wise sum of elements of the system matrix. 20 dB SNR.

$\mathbf{T}(\mathbf{x})$, the Toeplitz matrix generated from the random standard normal vector \mathbf{x} . The effect of Theorem 2 is illustrated in the fact that the Toeplitz random matrix shows the least $\|\mathbf{r}_{ls}\|_{l_\infty}$ since $c_k \rightarrow 0$ unlike in the other two cases. If we consider the hypothetical *oracle*-estimator, where we subsample based on the assumption that we possess full knowledge of the right samples to acquire, then the same results from Theorem 2 hold since the number of columns remains unchanged. Among all under-determined estimator problems, the *oracle*-estimator presents a the best-case scenario and is used as a benchmark.

2.5 l_∞ Norm in Compressive Radar Imaging

Compressive sensing algorithms are, in general, iterative algorithms or convex solvers. The analysis of the estimation residue is much harder compared to linear least squares estimation. Current literature on uniform norm analysis [49] provides upper bounds that are determined up to multiplicative scalar constants. Such error estimates cannot be directly applied to the task of building and operating radar systems. The phase transition-state evolution approach [45] is a more useful approach to derive readily usable approximations of the residue. The approach of phase transitions and state evolution reduce the analysis of compressive sensing to the task of understanding the behavior of a single scalar recursion. However, the state evolution formalism has been proven theoretically only for problems where \mathbf{A} is a random matrix. The theoretical results (Proposition 1, [45]) do not directly generalize to the uniform norm.

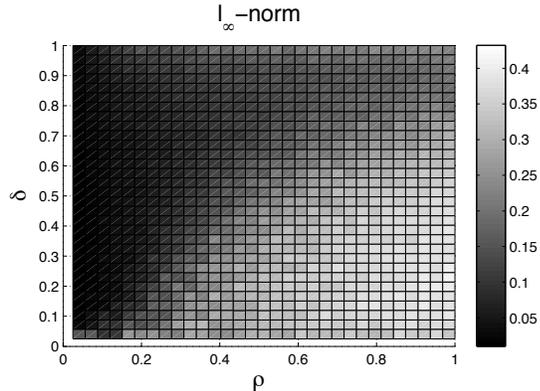


Figure 2.8: 20 dB SNR. Phase transition plot on a 32×32 grid on $[0, 1] \times [0, 1]$ for noise radar imaging, with \mathbf{A} as partial random Toeplitz.

The state evolution recursion is intimately related to the phenomenon of phase transitions [13]. Phase transition plots show the transition from regions of recovery to non-recovery on the plane of (ρ, δ) , where $\rho = K/N$ and $\delta = M/N$. It is seen in Figures 2.8, that the phase transition plots are similar to that obtained for l_2 residues and by using random \mathbf{A} , for which convergence has been shown theoretically [45]. This comparison demonstrates empirically that the state evolution formalism is valid for studying the uniform norm in the radar imaging scenario. The phase transition provides general values for the l_∞ norm for different combinations of ρ and δ , thus providing a direct way for radar operators to decide thresholds for distinguishing cells containing targets.

Similar to the least squares analysis, we consider four matrices with similar restricted isometries, but different values of c_k . The matrices considered were: Toeplitz random $\mathbf{T}(\mathbf{x})$, $\mathbf{T}(\mathbf{x} + 1)$, $|\mathbf{T}(\mathbf{x})|$, and the random standard normal matrix Φ . Figure 2.7 demonstrates the dependence on c_k . This suggests that compressive sensing residue behaves in a manner similar to Theorem 2.

2.6 Correlations in the Circulant Matrix

In practical compressive sensing systems, as we shall see in 2.8, transmit waveforms are not ideal. The assumption that transmit waveforms can be discretized into i.i.d random process is convenient for theoretical analysis of compressive recovery. In real systems how-

ever, without employing the kind of adaptive processing outlined in Chapter 4, we need to characterize the effect of correlations that exist due to hardware-related nonidealities. We model the transmit waveform as the correlation of an i.i.d random process $\tilde{x}(t)$ and a transfer function $h(t)$ that represents the bandlimiting non-idealities so that

$$x(t) = \tilde{x}(t) * h(t) \tag{2.23}$$

$$\mathbf{X} = \tilde{\mathbf{X}}\mathbf{H}. \tag{2.24}$$

In order to quantify the effect of correlations, we look at the transform point spread function (TPSF) of the system matrix. Ideally, for effective compressive signal recovery, we desire the non-diagonal elements of the normalized Gram matrix \mathbf{G} to be as low as possible. We look at the error metric given by

$$\chi(\mathbf{G}) = \sum_{i \neq j} |G_{i,j}|^2. \tag{2.25}$$

In Figure 2.9, we plot the values of $\chi(\mathbf{G})$ for various values of $\mathbf{G} = \mathbf{X}\mathbf{H}_l$, where l denotes the support of the power spectrum of the filters. We characterize the matrix \mathbf{H}_l with the parameter $l = \|P_h(f)\|_{l_0}$, where $P_h(f)$ refers to the power spectrum. Large values of l indicate that the waveform is highly uncorrelated. Lower values of l correspond to narrowly filtered random processes. The narrow power spectrum corresponds to highly correlated waveforms. The interesting result from this simulation is actually that compressive sensing is fairly robust to correlations in the transmit waveform. This conform with the experimental observations that we outline in 2.8. Even with a low-pass filter that only allows about 70% of the spectrum of the transmit waveform, we see in Figure 2.9 that χ is almost as good as the uncorrelated case.

2.7 Experiments

We used a millimeter wave radar system to test the possibility of using compressive sensing for noise radar. The bandwidth of the signal used in this system is 500 MHz. We operated

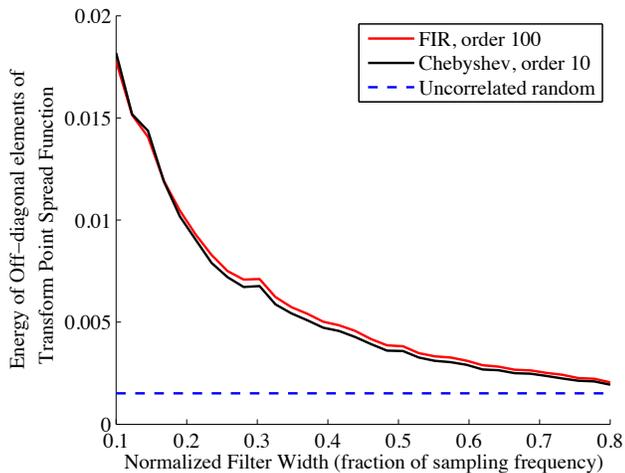


Figure 2.9: The effect of filtering on the TPSF. On the y-axis is $\chi\mathbf{G}$. It is seen that the performance of compressive sensing is expected to deteriorate if the random transmit waveforms are highly correlated (narrow filters).

the ADC at a rate of 1 gigasamples per second. The antenna has a half power beam-width of about 1° . The system consists of two conical antennas that are used for transmitting and receiving signals. The antennas have a half-power beam width of 1° . The conical antennas are connected to a high-power amplifier. The experiments were conducted in an outdoor setting. A photograph of the experimental set-up is shown in Figure 2.10. A typical target scenario is shown in Figure 2.11. We tested the imaging capability of the system at distances ranging from 40 ft to around 100 ft. We used tetrahedral *corner reflectors* and cylindrical scatterers as targets.

2.8 Analysis of Experimental Data

2.8.1 Characterization of Waveforms

Compressive sensing recovery with circulant matrices is possible when the waveform that generates the circulant system matrix is i.i.d random variables. Further, theoretically, it is unclear whether the distribution of the waveform affects the performance of compressive signal recovery. Our experiments indicate that, (a) waveforms in real systems are correlated as seen in Figure 2.12, (b) compressive signal recovery is tolerant to some extent to the

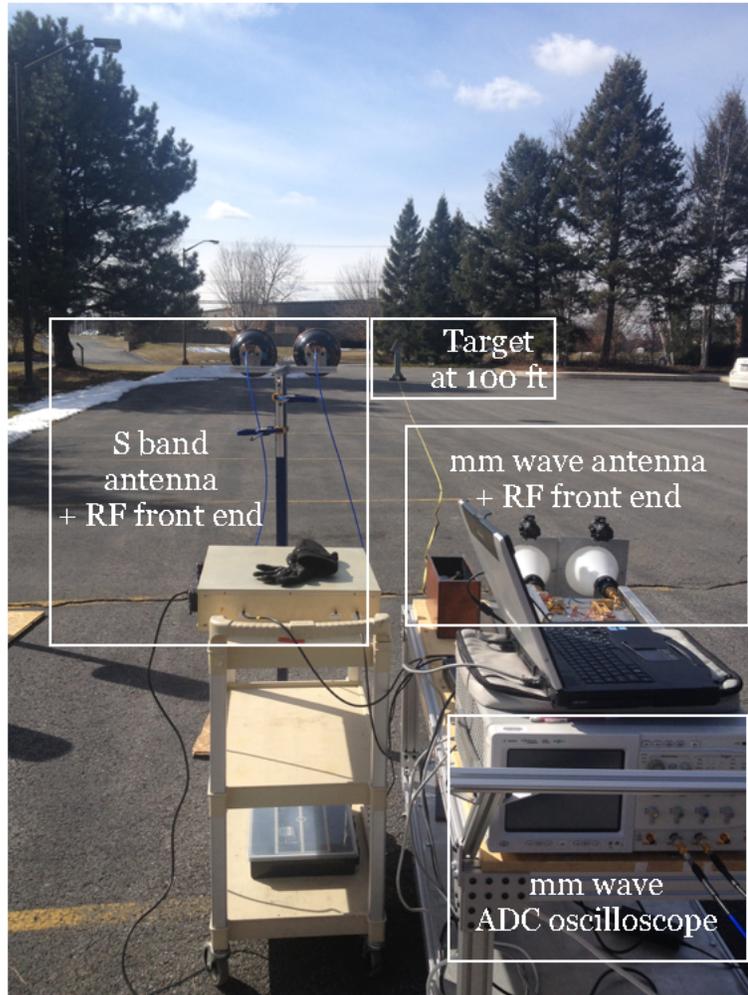


Figure 2.10: Photograph of experimental setup.



Figure 2.11: Close-up photograph of the corner reflector target.

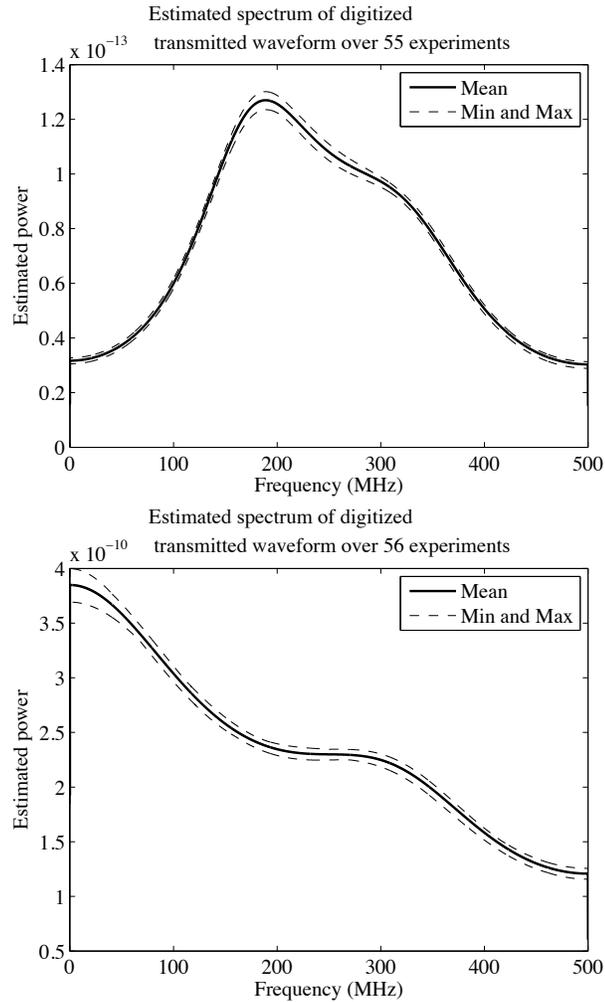


Figure 2.12: Power spectrum estimates (using a covariance estimator) of (a) mm wave transmit waveform, (b) S-band transmit waveform.

correlations in the system matrix. The correlations in the hardware are expected to worsen the performance of compressive signal recovery. However, as we saw in Figure 2.9 the degradation may not be significant as long as the autocorrelation is reasonably close to ideal. This is affirmed by the accuracy of recovery seen in Figures 2.15, 2.18, 2.19, 2.16, and 2.17.

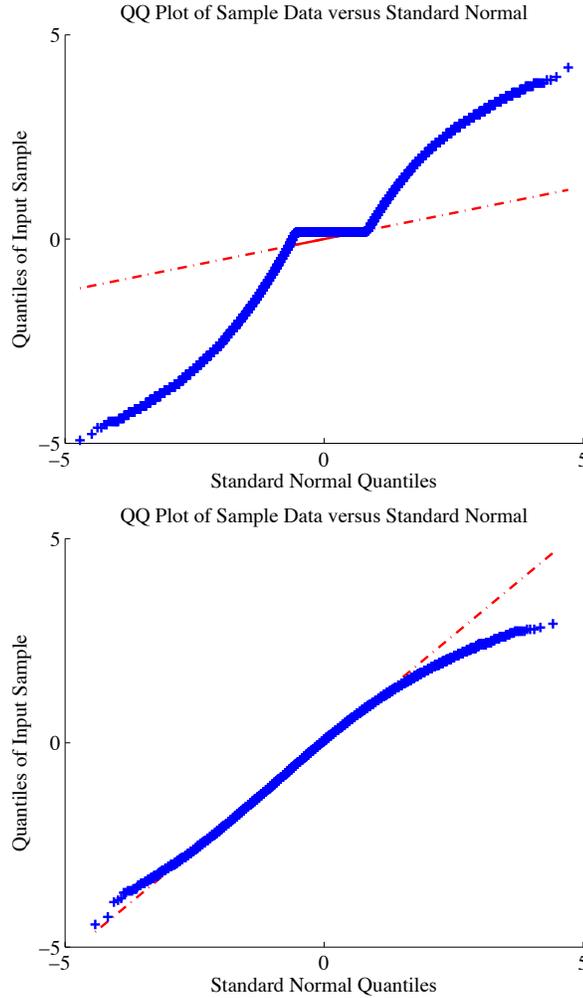


Figure 2.13: (a) QQ-plot of the normalized transmit waveform for mm wave radar in comparison with standard normal, (b) QQ-plot of the normalized transmit waveform for an s-band radar in comparison with standard normal

2.8.2 Measure of Performance

In order to measure the performance of the system, we adapted the standard peak-to-sidelobe ratio for compressive sensing. Peak-to-sidelobe ratio in the context of conventional radar imaging is defined as the ratio of the peak of the ambiguity function to the sidelobes. Since the ambiguity function does not have the same interpretation in the context of compressive radar imaging, we use a modified error metric. Rather than measuring the energies of the peak and sidelobe, we compared the uniform norm of the recovered vector within an

interval defined to contain the target and the rest of the signal. The uniform norm is directly related to the task of choosing thresholds to determining the significance of each pixel. We call this error metric the relative peak-to-sidelobe norm ratio (rPSN). Working within the universality of the phase transition framework, this error metric is specific to the matrix \mathbf{A} , the sparsity ρ , and the undersampling rate, δ . The reason for modifying the definition is that the corner reflector is not exactly a point target, but is actually represented by a few non-zero target cells. The traditional peak-to-sidelobe ratio is determined by the shape and spectrum of the transmit waveform. However, in this section that is not our objective. We wish to characterize the performance of real systems for different rates of under-sampling. Let us define the interval that contains the target as the set Γ . By considering the uniform norm, we can characterize the effect of thresholding and significance testing on each pixel. This metric is also appropriate because ensemble measures such as the l_2 norm are biased by the sparsity of the vector. The sparsity of a recovery vector doesn't necessarily indicate accurate recovery from a radar imaging perspective. Yet, for sparse vectors, the l_2 norm would show up as a low value. We write the relative complement as $\Gamma^c \triangleq [N] \setminus \Gamma$, where \setminus represents set-differences. Formally, we can then define the performance measure as

$$rPSN(\mathbf{A}, \rho, \delta) = \frac{\|\hat{\mathbf{s}}_\Gamma\|_{l_\infty}}{\|\hat{\mathbf{s}}_{\Gamma^c}\|_{l_\infty}}. \quad (2.26)$$

We demonstrate the stability of compressive noise radar imaging with respect to this error metric. We see that as we acquire fewer samples, $rPSN$ goes down. However, if the number of samples is too few, in this case $< 10\%$, the signal recovery fails. For the analysis shown in Figure 2.14, we considered the case of a single target. We plot the result of a number of experiments by indicating the lowest and highest errors that were seen across the different sets of experimental data. The error metric becomes worse as we perform recovery with around 90% or more of the samples. We surmise that this is because of two reasons, (i) approximating the continuous time problem with the discrete finite time convolution, (ii) as we increase the number of samples acquired, we also sample the noise more finely, resulting in more perturbations. Theoretical analysis of this phenomenon remains an open problem.

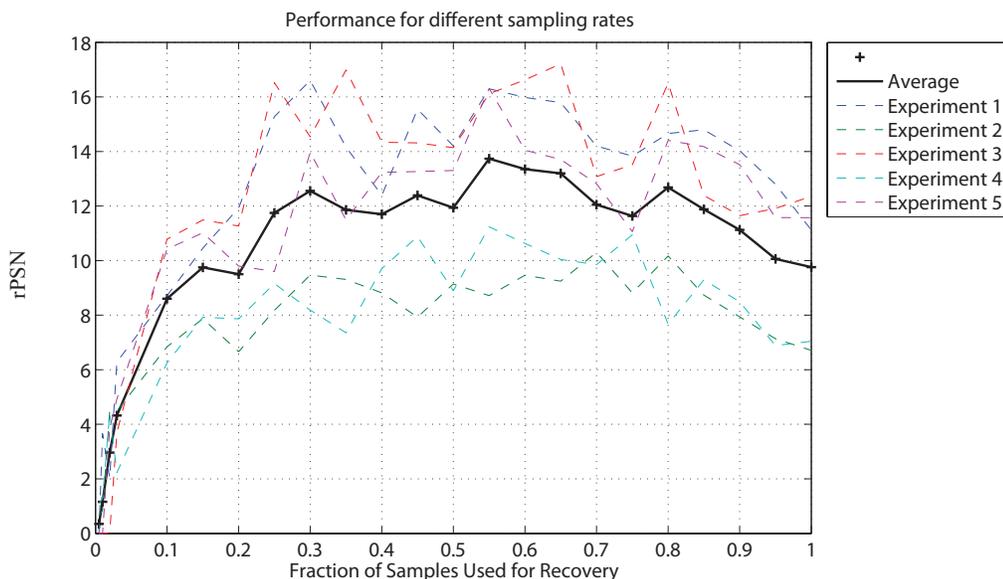


Figure 2.14: The error metric $rPSN$ for different sampling rates. An indication of the phenomenon of phase transition is seen in this plot. Up to a sampling rate of around 10%, recovery is near impossible. However with more than 10% of the samples we can fully recover the target information.

2.8.3 Imaging Performance

In this section, we compare the performance of compressive sensing recovery with traditional correlation processing and least squares. We see evidence for the fact that the performance of compressive recovery even with just 25% of the samples compares favorably with least squares and correlation processing. We repeated the experiment for different scenarios as described in the Figures 2.15, 2.18, 2.19, 2.16, and 2.17. The scenarios and the related experimental analyses are described below. In the descriptions, *target scene* refers to the sum total of all electromagnetic effects within the antenna beam.

1. The recovery of a target scene with one corner reflector placed at a distance of 100 ft from the antennas is illustrated in Figure 2.15. We chose a corner reflector because most of the energy backscatter is guaranteed to be along the antenna line of sight

of the antenna. This scenario was intended to demonstrate the feasibility of modeling compressive noise radar imaging as the problem of inverting a circulant random matrix.

2. Two cylindrical reflectors with the first placed at around 40 ft in the scenarios presented in Figures 2.16 and 2.17. The other is placed at different distances from the the first one, at locations from 40 - 45 ft. We see that the resolving capabilities of compressively sampled noise radar are comparable with that of conventional cross-correlation based processing. This favorable result is also seen in the case of a target scene consisting of two corner reflectors placed at 100 ft and separated by distances of 7 feet and 3 feet in Figures 2.18 and 2.19, respectively.

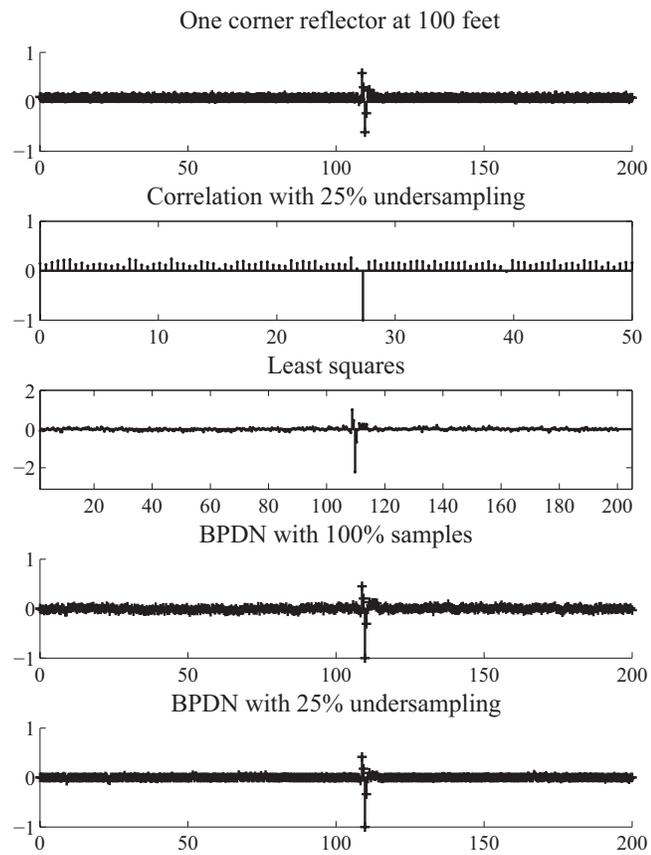


Figure 2.15: Corner reflector at a distance of 100 ft (millimeter wave radar)

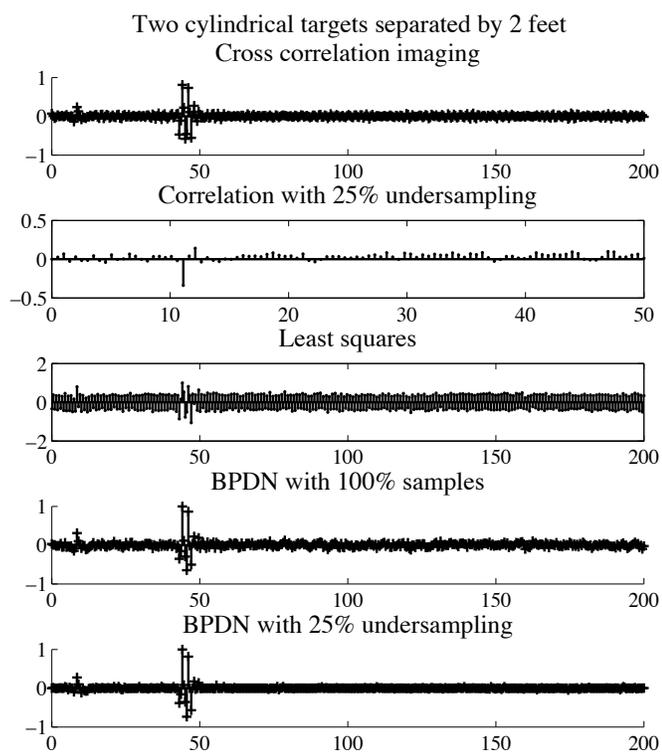


Figure 2.16: Two cylindrical targets at a distance of about 40 feet from the radar, and separated from each other by 2 feet.

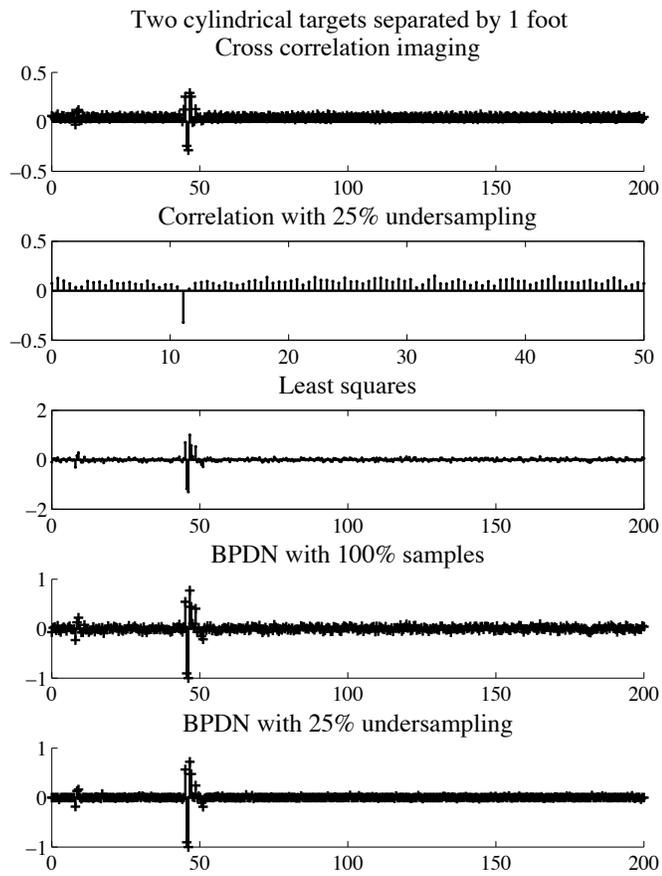


Figure 2.17: Two cylindrical targets at a distance of about 40 feet from the radar, and separated from each other by 1 foot.

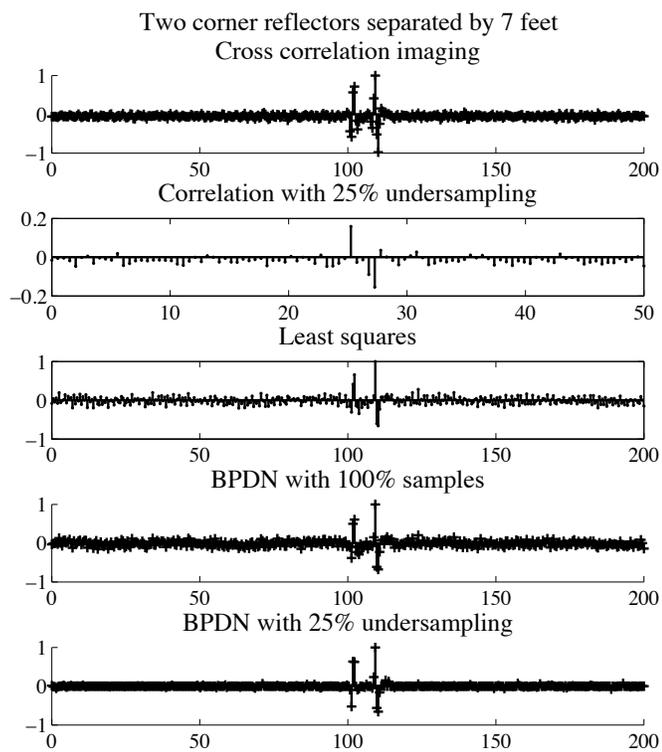


Figure 2.18: Two corner reflectors separated by a distance of 7 feet, located at a distance of 100 ft from the radar.

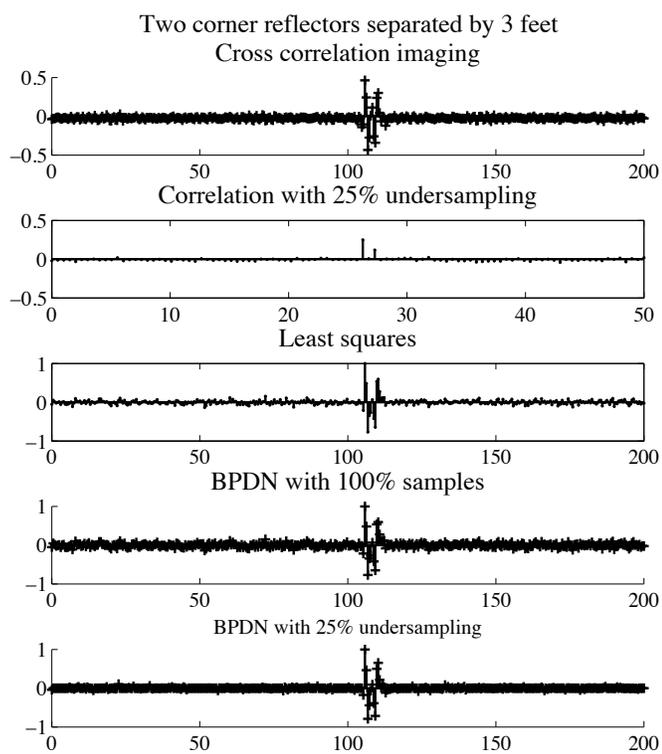


Figure 2.19: Two corner reflectors separated by a distance of 3 feet, located at a distance of 100 ft from the radar.

Chapter 3

Detection Strategies for Compressive Noise Radar

3.1 Compressive Sensing Detection

Target detection is a fundamental task in radar imaging systems. There are well established approaches to target detection in conventional radar imaging [4]. In conventional radar systems, with the exception of space-time adaptive processing (STAP), radar detection theory has largely been amenable to theoretical analysis. Threshold detection has proved particularly powerful in radar imaging. Improvements in the detection have primarily been achieved by optimizing objective functions constructed from expressions for the signal-to-noise ratio. We consider here, the problem of detecting each element of the recovered vector in compressive radar imaging. We continue with the signal model discussed in Chapter 1 and 2. We model the reflected signal as a convolution of the transmit waveform and the target (channel) impulse response. The fully determined system is given by

$$\mathbf{y} = \mathbf{X}\mathbf{s} + \boldsymbol{\eta}. \quad (3.1)$$

In order to simplify the acquisition hardware, we propose to just acquire a subset of the waveform \mathbf{y} . Thus, the linear system representing the target imaging process for compressive

noise radar imaging is given by

$$\mathbf{z} = \mathbf{R}_\Omega \mathbf{X} \mathbf{s} + \mathbf{R}_\Omega \boldsymbol{\eta}. \quad (3.2)$$

We define $\mathbf{A} = \mathbf{R}_\Omega$ and $\boldsymbol{\eta}_\Omega = \mathbf{R}_\Omega \boldsymbol{\eta}$ so that we have the system equation,

$$\mathbf{z} = \mathbf{A} \mathbf{s} + \boldsymbol{\eta}. \quad (3.3)$$

We leverage the sparsity of \mathbf{s} to invert this linear system using l_1 optimization-based algorithms. The recovered signal is given by

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathbb{R}^N} \|\mathbf{s}\|_{l_1} \text{ s.t. } \|\mathbf{z} - \mathbf{A} \mathbf{s}\|_{l_2} \leq \sigma. \quad (3.4)$$

Our detection problem seeks the significance of each recovered pixel of the vector \mathbf{s} . The null hypothesis is the absence of a target at that pixel.

$$\mathcal{H}_0^{(k)} : s(k) = 0 \quad (3.5)$$

$$\mathcal{H}_1^{(k)} : s(k) \neq 0. \quad (3.6)$$

Relating this ground truth hypothesis with the recovered target $\hat{\mathbf{s}}$, we get the probabilities,

$$P_D^{(k)} = \mathbb{P}[\hat{s}(k) > \xi | H_1^{(k)}] \quad (3.7)$$

$$P_{FA}^{(k)} = \mathbb{P}[\hat{s}(k) > \xi | H_0^{(k)}]. \quad (3.8)$$

$P_D^{(k)}$ and $P_{FA}^{(k)}$ henceforth refer to the probabilities of detection and false alarm, respectively, for each pixel indexed by k .

Detection of radar targets is accomplished by comparing different attributes of signals and images with thresholds. In our case, we look at the problem of determining the significance of each recovered pixel of the target scene. The important task of characterizing the detector requires us to derive relationships between thresholds and their corresponding probabilities of detection and false alarm. Thresholds are meaningless in the absence of

their association with detection performance metrics.

As we have discussed earlier in this thesis, the theory and algorithms of compressive sensing present interesting new possibilities for radar systems. However, there has been precious little work in casting compressive radar imaging in the context of conventional radar systems. The first paper to explore this problem from the perspective of signal processing was the 2010 paper on compressive signal processing by Davenport, et al. [50]. Their proposal was to use a threshold detection approach based on the sufficient statistic given by $\mathbf{z}^T(\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\hat{\mathbf{s}}$. This scalar detector is an ensemble quantity that is not useful for pixel-wise thresholding. The asymptotic results concerning the detectors of Davenport, et al. were derived with the assumption that the system matrix is Gaussian random. The most recent work related to our thesis is the 2012 paper by Anitori et al. [48] on the design and analysis of detectors for compressive sensing. Anitori et al. propose to apply the complex-approximate message passing algorithm to arrive at a closed form for the distribution of the recovery error. However, the approximate message passing algorithm has not been shown to converge for problems with circulant random matrices. Thus, we develop a data-driven approach to threshold detection in compressive noise radar.

3.1.1 Compressive Sensing Detection: An Empirical Examination

In order to demonstrate the effectiveness of compressive detection, we use a matching pursuit-based algorithm for detection [51]. The algorithm is essentially the traditional matching pursuit algorithm, with the modification that the matching pursuit iteration stops when the l_∞ norm, defined as the maximum element of the estimated vector is found to be above a fixed threshold. We look at some simulations to verify the performance of such an algorithm. Thus, a convergence to the solution would be seen in the presence of a target, with the subsequently recovered vector lying above the detection threshold. Figure 3.1 shows the curves for the probability of error and false alarm for an arbitrarily fixed threshold. In the absence of theoretical results, we study the application of this algorithm using Monte-Carlo simulations for computing the probabilities of false alarm and detection. The performance of the above compressive detector in the context of radar systems is seen

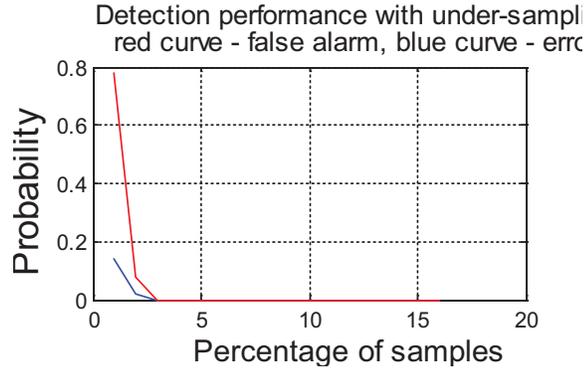


Figure 3.1: The probability of error (red curve) and false alarm (blue curve) plotted as a function of the rate of fraction of samples acquired $SNR = 20dB$.

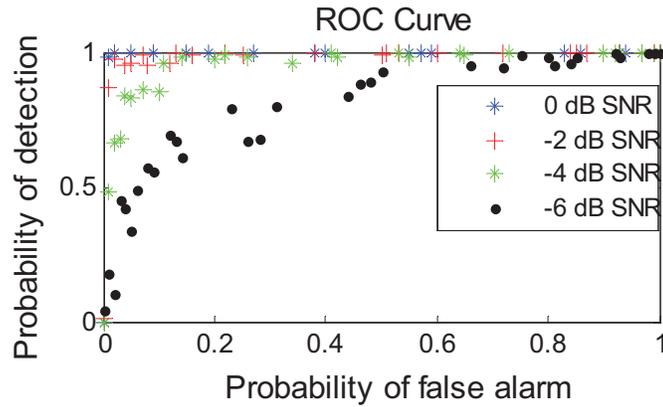


Figure 3.2: The receiver operating curves for different values of signal to noise ratio, with $(\delta = 25\%, \rho = 10\%)$.

in Figure 3.2 to be reasonably good at even low signal to noise ratios. The simulations demonstrate that compressive sensing performs reasonably well in the context of threshold detection.

3.2 Detection Algorithms

3.2.1 Thresholding Based on Extreme Value Theory

In the context of sparse radar imaging, fixing the threshold for detecting and defining significant targets is an important task. With a large percentage of the target cells being zero, the cost of false alarms is proportionally higher. The choice of the threshold is decided by

desired probabilities of detection and false alarm. In particular, radar imaging applications often require setting variable thresholds to maintain false alarm rates.

3.3 Statistics of the Error of Compressive Signal Recovery

There are no theoretical guarantees for the distribution of the residue of compressive sensing recovery when the system matrix is circulant random. In this section, we perform some empirical tests to study whether it is reasonable to approximate compressive signal recovery using normal distribution. We looked at those points of \hat{s} that are characterized by the indices $i : s_i = 0$. This is equivalent to the instances of falsely detecting targets. In order to test the 'Gaussianity' we use the Kolmogorov-Smirnov (KS) test and quantile-quantile (QQ) plots. The Kolmogorov-Smirnov test involves testing the null hypothesis that the distribution underlying the data is standard normal. We normalized the data by subtracting the mean and dividing by the standard deviation. Following this normalization, the empirical cdf of data is compared to the standard normal distribution. The results of the KS test are given in Table 3.1. The quantile-quantile plots (qq-plot) in Figures 3.3 and 3.4 compare the quantiles of empirical data with that of theoretical estimates. Deviations of the empirical data from the straight line indicate non-Gaussianity. These two tests provide evidence that it is unreasonable to model the recovery residue as being drawn from a Gaussian distribution. This precludes using sample mean and sample variance as estimates in computing detection thresholds and probabilities of false alarm and detection. In order to completely characterize threshold detection for compressive sensing, we need to have a fairly accurate idea of the underlying distributions.

3.4 Algorithm for Estimating the Thresholds for Compressive Detection

When convex optimization recovery schemes such as SPGL1 are used, the nonlinear and iterative nature of the estimator algorithm, and the arbitrary evolution of the iterates rule out the possibility of deriving theoretical expressions for the distribution of the residue. The

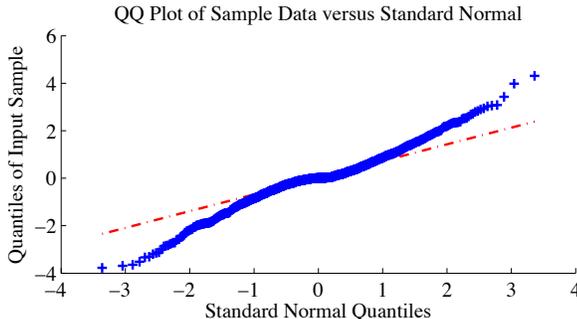


Figure 3.3: Quantile-quantile plot of **synthetic** normalized data. Blue markers represent the statistics of the points that satisfy \hat{s}_i with i such that $s_i = 0$. Since the blue markers deviate significantly from the red line, the data cannot be modeled as Gaussian.

Data Type	Null Hypothesis: Standard Normal	p-value
Synthetic	Reject	0
Experimental 1	Reject	0
Experimental 2	Reject	0
Experimental 3	Reject	0
Experimental 4	Reject	0

Table 3.1: Table summarizing the Kolmogorov-Smirnov test comparing a few instances of the normalized data with the standard normal distribution.

computational cost of convex optimization solvers presents a significant barrier to employing a brute force Monte Carlo approach to determining the threshold. For a desired probability of false alarm, P_{fa} , the number of instances of BPDN(.) should be at least of the order of $100/P_{fa}$ to compute the threshold. This becomes a problem when the desired P_{fa} is lower than values achieved in conventional detection schemes which is usually 10^{-4} or lesser. This is because each instance of running a convex optimization BPDN(.)-solver such as the SPGL1 algorithm, takes several seconds (run without parallelization on an Intel i7, 8 GB RAM PC) to converge, even for vectors as small as $N = 1000$.

There have been extensions to Monte Carlo simulations in the past to accommodate the occurrence of rare events [52]. One such approach involves estimating probabilities using extreme value theory [53, 52]. We propose an approach based on limit theorems from extreme value theory to overcome this problem. Extreme value theory refers to the study of probabilities of rare events in stochastic systems. The basic results concern the statistics

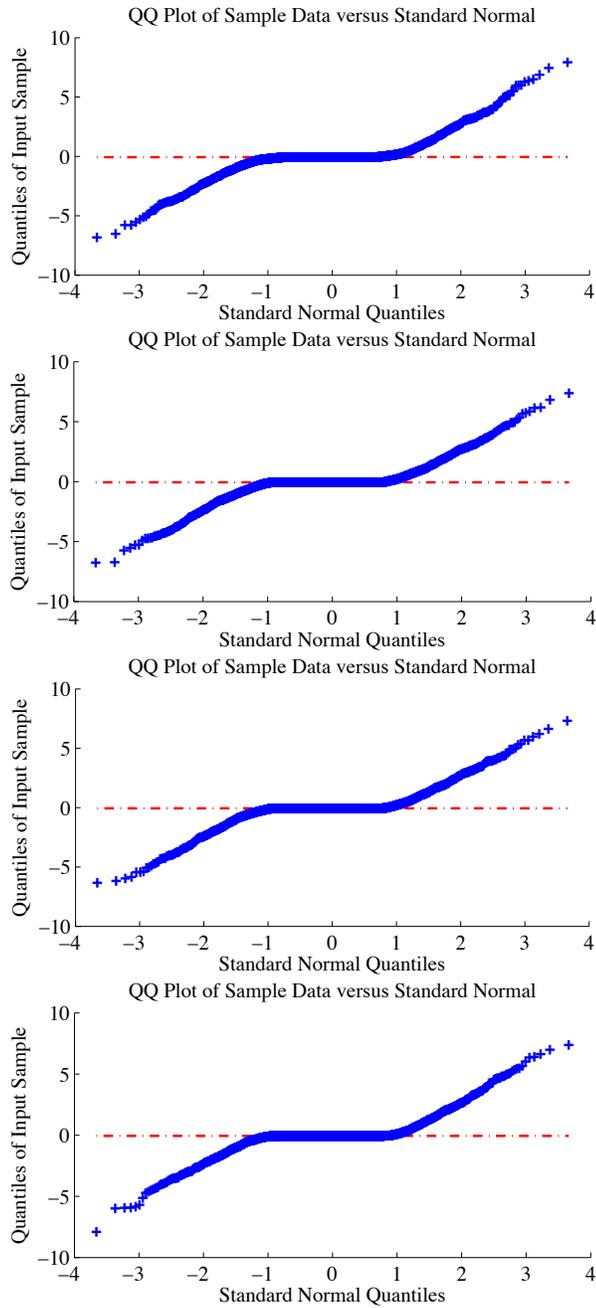


Figure 3.4: Quantile-quantile plots of **real** normalized data. Blue markers represent the statistics of the points that satisfy \hat{s}_i with i such that $s_i = 0$. If the data are distributed according to the standard normal distribution, then the blue markers coincide with the red line.

of the extremes of ordered random variables. In the past, it has been applied to problems in finance, climate sciences, and geophysical modeling. In electrical engineering, the utility of extreme value theory was first proposed for problems in detection theory by Ozturk et al. [53]. In the context of compressive sensing, our basic proposal is to extrapolate the probabilities of rare events from a few instances of solving the convex optimization problem. A manageable number of iterations of convex optimization problem are used to generate the statistics of compressive sensing for various values of (ρ, δ) and these are used to compute thresholds for small values of P_{fa} ($< 10^{-4}$).

3.4.1 Threshold Estimation for Compressive Detection

We adopt a data-driven approach to estimate thresholds for compressive detection. From the sparsity assumption, it follows that a large proportion of the recovered vector will be zeros. Thus, we start with the assumption that we have access to the oracle knowledge about the location of a few zeros in the recovered vector. In a real radar system, this may, for example, be knowledge acquired by visual examination of the target scene, and the realization that some locations consist of zeros. Let $Z \triangleq \{k : s(k) = 0\}$ denote the subset of $[N]$ about which we have knowledge that targets are absent. If for each individual pixel, we consider a simple threshold test with the two hypotheses given by,

$$\mathcal{H}_0^{(k)} : s(k) = 0 \tag{3.9}$$

$$\mathcal{H}_1^{(k)} : s(k) \neq 0 \tag{3.10}$$

$$\tag{3.11}$$

To determine the presence and absence of targets at each target cell, we adopt an approach of comparing each pixel with a threshold, ξ . Then, the probability of detection and false alarm for each pixel will be

$$P_D = \mathbb{P}[s^*(k) > \xi | H_1^{(k)}] \tag{3.12}$$

$$P_{FA} = \mathbb{P}[s^*(k) > \xi | H_0^{(k)}] \tag{3.13}$$

By observing the statistics of members of the set Z , across different realizations of the convex optimization solver for different ξ , we can determine the false alarm rate. Let $k_z^{(i)}$ denote elements of sub-sequence of $[N]$, such that, for each $k_z \in Z$,

$$P_{FA} = \mathbb{P}[s^*(k_z^{(i)}) > \xi]. \quad (3.14)$$

Let us assume that the random variable $s^*(k_z^{(i)})$ has the pdf $p_z(x)$; then we can compute P_{FA} as

$$P_{FA}(\xi) = \int_{\xi}^{\infty} p_z(x) dx. \quad (3.15)$$

If we employ Monte Carlo simulations to estimate \hat{p}_z^{MC} , the result will be

$$P_{FA}^{MC}(\xi) = \frac{1}{Q} \sum_1^Q \eta(s^*(k_z^{(i)}); \xi) \quad (3.16)$$

For a given P_{FA} , we thus require $Q \gg 1/P_{FA}$, which is computationally intractable when each $s^*(k_z^{(i)})$ is being generated from a convex optimization solver. Following past research [52, 53] in this area, we use the Generalized Pareto Distribution (GPD) to estimate the tail distribution and thus the P_{FA} . The cumulative distribution function of GPD is given by,

$$G(x) \triangleq 1 - \left(1 + \frac{\gamma x}{\zeta}\right)^{-\frac{1}{\gamma}}, \quad (3.17)$$

with

$$-\infty < \gamma < \infty, \zeta > 0, \gamma x > -\zeta. \quad (3.18)$$

GPD parametrizes numerous other distributions such as the exponential distribution when $\gamma = 0$ and the uniform distribution when $\gamma = -1$. As proved by Pickands [54], the following

result relates the GPD to the tails of general, unspecified distributions;

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sup_{0 \leq x < \infty} |\mathbb{P}[Y > y + u | Y \geq u] - 1 + G(y)| > \epsilon \right] = 0, \quad (3.19)$$

$$\forall \epsilon > 0. \quad (3.20)$$

Now, we note that the conditional expectation given by,

$$F_u(y) = \mathbb{P}[Y \leq y + u | Y > u] = \frac{F(u + y) - F(u)}{1 - F(u)}. \quad (3.21)$$

Setting $z = u + y$, we get

$$F(z) = F_u(z - u)(1 - F(u)) + F(u). \quad (3.22)$$

Further, we can define $\alpha \triangleq 1 - F(u)$, so that

$$F(z) = \alpha F_u(z - u) + (1 - \alpha). \quad (3.23)$$

We proceed by estimating the limit of $F_u(z - u)$ based on Equation (3.19), so that,

$$F_u(z - u) = G(z) = 1 - \left(1 + \frac{\gamma}{\zeta}(z - u) \right)^{-\frac{1}{\gamma}} \quad (3.24)$$

and

$$F(z) = \alpha \left(1 - \left(1 + \frac{\gamma}{\zeta}(z - u) \right)^{-\frac{1}{\gamma}} \right) + (1 - \alpha) \quad (3.25)$$

$$= 1 - \alpha \left\{ 1 + \frac{\gamma}{\zeta}(z - u) \right\}^{-\frac{1}{\gamma}}. \quad (3.26)$$

Thus, the general strategy for applying GPD to estimate rare-event probabilities is to first set a particular α and then estimate the tail of the unknown distribution using 3.25. A typically used value is $\alpha = 0.1$, for which the value of u can be computed using Monte Carlo simulations to be the value of the threshold representing the $100 \times (1 - \alpha)^{th}$ percentile of

the data. The parametric function is then fitted to the given data to arrive at values for γ and ζ . We follow the proposal of [53] and employ the Nelder-Mead algorithm [53] for solving the maximum-likelihood formulation given by

$$\left(\hat{\zeta}, \hat{\gamma}\right) = \arg \min_{\gamma, \zeta} \left(\alpha Q \log \zeta + \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^{\alpha Q} \log \left(1 + \frac{\gamma z_i}{\zeta}\right) \right). \quad (3.27)$$

Subsequently, the relationship between the probability of false alarm and the threshold can be derived based on the GPD estimate as follows:

$$P_{FA} = \alpha \left\{ 1 + \frac{\hat{\gamma}}{\hat{\zeta}} (\tau - u) \right\}^{-\frac{1}{\hat{\gamma}}}, \quad (3.28)$$

$$\tau = u + \frac{\hat{\zeta}}{\hat{\gamma}} \left(\left(\frac{P_{FA}}{\alpha} \right)^{-\hat{\gamma}} - 1 \right). \quad (3.29)$$

3.4.2 GPD and Compressive Sensing

With the above results about GPD established, we now proceed to apply it to compressive sensing. We treat the convex optimization solver as an experiment whose reconstruction error has an unknown distribution. We wish to estimate accurate thresholds for low P_{FA} . The advantage of using the approach based on GPD is that the results so derived are to a large extent, independent of the type of distributions of the target scene, \mathbf{s} , the noise η , and residue $\mathbf{s} - \hat{\mathbf{s}}$. Our methodology for deriving the thresholds for compressive sensing is as follows:

Algorithm 1 GPD-based tail, P_{FA} , and threshold estimation.

Estimation of P_{FA} as a function of threshold for compressive sensing using GPD.

Input: \mathbf{X} , \mathbf{s} .

Output: GPD parameters, $\hat{\gamma}^{(CS)}$, $\hat{\zeta}$.

for $j \in Z$ **do**

for $i = 1 \rightarrow Q$ **do**

 Solve the convex optimization problem given by BPDN(ρ, δ, σ);

 Use the entire recovered vector $\hat{\mathbf{s}}$ as a training set, i.e., $T = \{1, 2, \dots, N\}$ OR if available, choose a set of points for which the truth of the hypothesis is known, i.e. $T = \{i : s_i = 0\}$;

end for

 Using the points $r(j) : j \in T$, construct the distribution $\hat{p}_{r(j)}(x) = \sum_{t=1}^Q \mathbb{1}_{r(k_z^{(t)})}(x)$;

end for

Set $\alpha = 0.1$, and select $u = r(b)$ such that $\#\{t : r(t) > r(b)\} = \lfloor \alpha Q \rfloor$, and let $T \triangleq \{t : r(t) > r(b)\}$;

Let $z^{(u)}$ be a sequence such that $\forall j, z_j^{(u)} \in T$;

Estimate GPD parameters μ and ζ by applying the Nelder-Mead solver to solve the *maximum likelihood* function optimization problem given by $\min_{\gamma, \zeta} \alpha Q \log \zeta + \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^{\alpha Q} \log \left(1 + \frac{\gamma}{\zeta} (z_i^{(u)} - u)\right)$. Call this solution $\hat{\gamma}^{(CS)}$ and $\hat{\zeta}^{(CS)}$;

The threshold is computed from the probability of false alarm as

$$\tau^{(CS)}(P_{FA}) = u + \frac{\hat{\zeta}}{\hat{\gamma}} \left(\left(\frac{P_{FA}}{\alpha} \right)^{-\hat{\gamma}} - 1 \right) \quad (3.30)$$

With the relationship between P_{FA} and τ established, we can proceed to derive the probability of detection from the statistics of the non-zero values and establish the receiver operating characteristics.

3.4.3 Performance of GPD-based Threshold Estimation

We demonstrate the effectiveness of Algorithm 1 to data generated from applying the SPGL1 convex optimization solver to synthesized compressive noise radar problems. In Figure 3.5, we plot the complete real CDF of the absolute value of the residual data with the relatively small value of $Q = 5 \times 10^4$ for the training data. The parameters for the GPD-based estimate of the P_{FA} are derived from the data. The value of α has to be chosen carefully

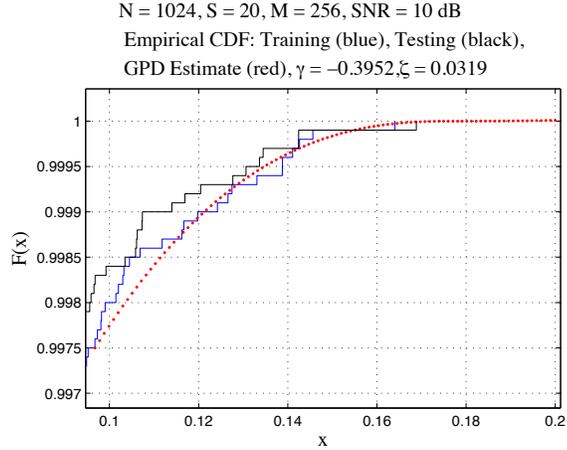


Figure 3.5: Plot showing application of Algorithm 1 for GPD-based extrapolation of the cumulative distribution of empirical residue data. Compressive recovery with $N = 1024$.

based on the number of reliable non-zero values in the training data. A high value of α will mean that the training data contain too many samples too far from the tail. If α is too low, then there will be too few samples in the training data. This GPD-based extrapolation is seen to conform with both training and test data.

The GPD estimates are obtained for 50 independent realizations of Algorithm 1. From these estimates, the threshold is computed for various values of the probability of false alarm from equation 3.30. The 50 realizations are plotted as smoothed probability density function in 3.6, estimated as

$$p_\tau(x) = \sum_{i=1}^{50} \mathbb{1}_{\hat{\tau}}(x),$$

where $\hat{\tau}$ is the estimated threshold value. In the simulations, we used the `ksdensity` function in matlab. The median threshold for each P_{FA} is indicated in the plots. As the desired P_{FA} is lower, the unknown threshold values are farther away from the training data. Thus, the uncertainty in the estimated threshold value increases, as indicated by the increasing spread in the pdf of the data. The reliable median allows us to extract a meaningful estimate even when the desired probability of false alarm is as low as 10^{-8} .

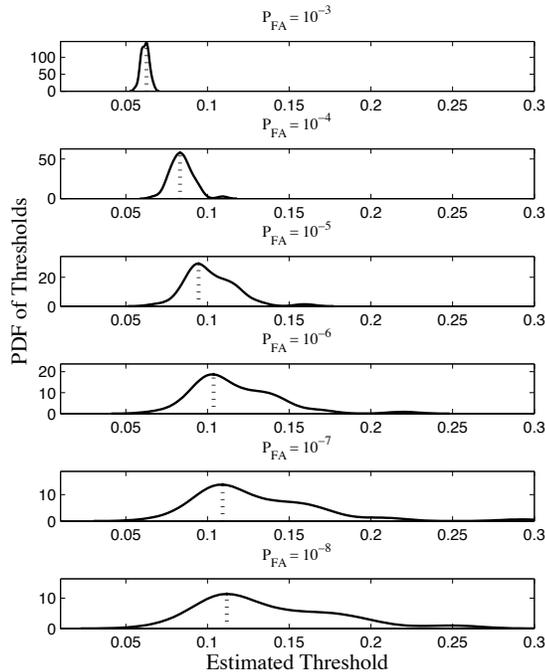


Figure 3.6: The probability distribution of the estimates of the threshold value for various desired values of P_{FA} . Compressive recovery was performed with $S = 10, M = 256, N = 1024$ and $\text{SNR} = 10$ dB. The estimate of the P_{FA} was done using $\alpha = 0.01$. The pdf of the threshold estimation itself was generated using 50 instances of algorithm 1.

3.4.4 Performance of Algorithm on Experimental Data

We applied our threshold estimation algorithm on experimental data that was acquired as described in Chapter 2. The fully sampled (1 gigasamples per second) received waveform and transmitted waveform were a record of 10^5 data samples. We divided the entire record into smaller sets of length $N = 4096$. For one of these sets, we solved the l_1 -minimization based compressive recovery algorithm to extract the radar target image. We then utilized partial knowledge of the locations of the non-zero values to construct a set that represented the points $\hat{s}_i : s_i = 0$. For the subset of locations where targets are absent, we estimated the tail using Algorithm 1. We verify the performance of this algorithm by comparing it with the estimated cumulative distribution functions of the empirical data. We combined the results of several such experiments to empirically construct the extended cumulative

distribution function. The accurate reconstruction of these values is seen in Figure 3.7.

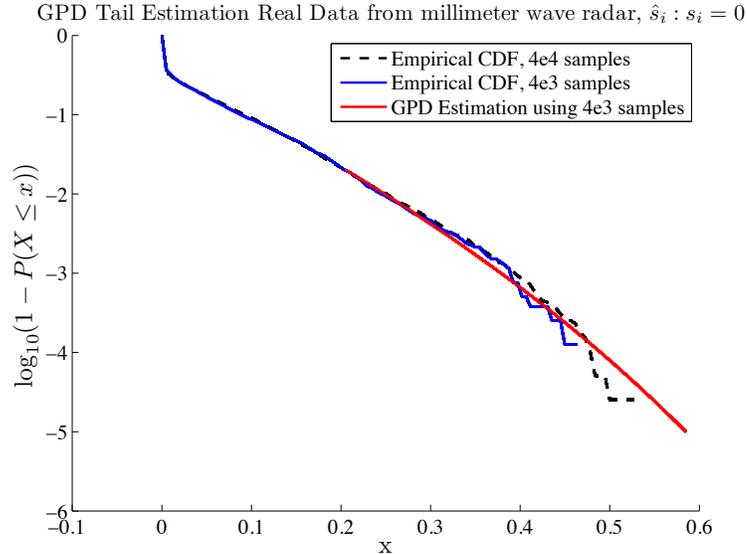


Figure 3.7: Tail estimation for real data. We observe that the empirical CDF corresponds closely with the GPD estimate for the tail. The target scenario involved one corner reflector target at 100 ft imaged multiple times using the millimeter wave radar described in Chapter 3.

3.5 Computational Complexity

In our work, we use convex optimization to solve the compressive signal recovery problem. While there are more computationally efficient algorithms for compressive recovery, convex optimization provides the best performance in terms of recovery accuracy [12]. The recent algorithm of Approximate Message Passing (AMP) [13] is an exception in that its performance is theoretically identical to convex optimization. However, this theoretical guarantee is only valid for Gaussian random system matrices. Compressive sensing is particularly sensitive to the nature of the system matrix, and it is as yet unclear how AMP-based algorithms behave for circulant matrices.

We use the spectral projected gradient algorithm- l_1 (SPGL1) [19] for solving the basis pursuit denoising problem. Each iteration of the SPGL1 algorithm involves a matrix-vector product. In our problem, this corresponds to a computational complexity of $O(MN)$,

assuming that each scalar arithmetic operation can be accomplished with complexity of $O(1)$. Multiplying with circulant system matrices is an operation that takes $O(N \log N)$ operations. While the exact estimate for the number of iterations for SPGL1 is unknown, the similar ParNes algorithm [55] requires $O(\sqrt{1/\sigma})$ iterations to converge to a solution. So we surmise that the most efficient convex-optimization based BPDN solvers have a complexity of $O(M \log N \sqrt{1/\sigma})$.

Let us assume that we use N_{mc} number of instances of SPGL1, then the cost of constructing the empirical CDF of the residue would be $O(N_{mc} N \log N \sqrt{1/\sigma})$. Our approach to estimating the CDF seeks to achieve two objectives, computationally speaking: (a) reduce the number of instances N_{mc} , (b) solve smaller problems on each instance so that we can lower the value of N . Let us assume that we desire a probability of false alarm of \tilde{P}_{fa} . In order to accurately estimate the threshold that achieves this probability, we would require around $\tilde{N}_{mc} = 10/\tilde{P}_{fa}$ instances. Thus the cost of \tilde{N}_{mc} instances of the SPGL1 algorithm would be $O(\tilde{N}_{mc} N \log N \sqrt{1/\sigma})$.

Using the GPD approach for extrapolating tails allows us to estimate thresholds for low values of P_{fa} without significantly increasing N_{mc} and N . In our experiments, we were able to extrapolate accurately, the tail distribution up to events as rare as $P_{fa} = 10^{-5}$ using around 10^3 samples. This implies that using extreme value theory to estimate the tail probabilities means we can achieve a reduction of a couple of orders of magnitude in the values of N_{mc} and N required to compute a given value P_{fa} . Let us assume that by using extreme value theory, we can compute a required \tilde{P}_{fa} with just $\gamma_{mc} \tilde{N}_{mc}$ instances of problems of $\gamma_s N$ dimensions for some constants, $\gamma_{mc}, \gamma_s \ll 1$. Then the computational cost of simulating the rare event reduces approximately by a factor of $\gamma_{mc} \gamma_s$. Based on empirical evidence, $\gamma_{mc}, \gamma_s \approx 10^{-2}$ implies that we can achieve a reduction of 10^{-4} in the cost of computing the threshold for low probabilities of false alarm. This improvement is achieved at the relatively smaller additional cost of solving the GPD parameter estimation problem, which is a two-dimensional maximum likelihood convex optimization problem.

In a real system, our goal is to compute detection thresholds with as little latency as possible. An improvement of a factor of $(1/\gamma_{mc} \gamma_s) \approx 10^4$ in the computational speed of

estimating thresholds represents a significant advantage when each solution of the SPGL1 algorithm takes around 300 seconds for a 10000 dimensional problem, as we see in Figure 3.8. Further in many instances, we may not have access to a training set that is sufficiently large to empirically estimate thresholds corresponding to low probabilities of error. Thus, we believe that an extreme-value theory based approach to threshold detection is useful for practical compressive sensing systems.

The computational expense of convex optimization based signal recovery presents a significant stumbling block in running extensive Monte Carlo simulations. We studied the computational efficiency of running compressive recovery algorithms, specifically, the spectral projected gradient l_1 (SPGL1) algorithm [19] on a quad-core Intel i7 2.8 GHz desktop with 8 GB of RAM. In Figure 3.8, we see that there is a roughly exponential increase in the computational cost as the dimension of the problem increases. We considered the case of a single target being present in the target scene. This means that the solution we seek is as highly sparse as it can get. Further, the computational efficiency is worse for real data compared to simulated data. We believe that this is due to the non-idealities of real systems. The bandlimiting properties of the filters used increases the number of non-zero pixels. This worsening of the sparsity means that the convex optimization problem takes longer to converge. Further, the rate of convergence worsens as the signal-to-noise ratio decreases.

A single solution of a problem of dimension 10,000 takes about five minutes. A problem dimension of 10,000 would provide sufficient opportunities to simulate events of likelihoods of the order of 0.01. In summary, if we resort to brute-force Monte Carlo simulations, then it would take us in excess of five minutes to even simulate events that occur with a probability between 10^{-3} and 10^{-2} . Thus computationally speaking it can get prohibitively expensive to simulate events that occur with lower probabilities. The direct implication to detection theory is that it becomes impractical to estimate thresholds which can yield low false alarm rates using brute-force Monte Carlo methods.

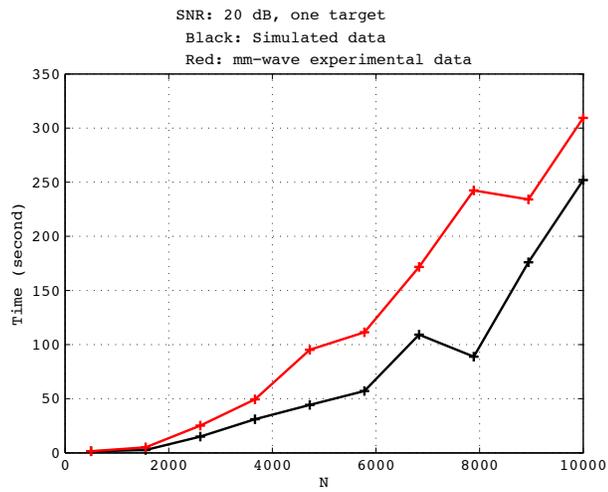


Figure 3.8: Computational complexity of SPGL 1 convex optimization signal recovery in terms of running time of algorithm on an Intel i7 2.8 GHz with 8 GB of RAM.

Chapter 4

Waveform Design and Compressively Sampled Radar

4.1 The Need for Designing Waveforms

Historically, radar signal processing has focused on developing waveforms and systems that maximize the signal-to-noise ratio. However, even as early as in the few years following Shannon’s seminal paper on information theory, there was considerable interest in casting radar imaging as a problem of maximizing the quality and quantity of *information* about the target. This approach of applying general ideas from information theory to radar was first articulated by Woodward [56] in the early 1950s. Due to constraints imposed by system-level and algorithmic complexity limited the adoption of this idea in real systems. Bell [57] extended the theory considerably in an important paper in 1993, where he formalized the idea of maximizing the mutual information of the received waveform and the target by optimizing the transmit waveforms. In the years since, there has been a considerable amount of work in this area, as applied to multiple-input multiple-output (MIMO) radar [58], adaptive radar [59], optimal detection [60], target classification [61], and knowledge-based radar [62].

Portions of this chapter were published in the paper “Waveform Design for Compressively Sampled UWB Radar,” authored by MCS, Ram M. Narayanan, and Muralidhar Rangaswamy that is to appear in the Journal of Electronic Imaging Special Issue on Com-

pressive Sensing in April 2013 [63].

Of particular relevance to our work is the idea of *knowledge-based* radar. Knowledge-based radar refers to the concept of incorporating knowledge or general information that we possess about a target scene into the detection and imaging framework. The development of knowledge-aided systems requires that we construct a framework for incorporating uncertain information concerning the target into the radar signal processing system. We propose to accomplish this task by expanding the signal in basis functions that reflect the target scene. The main benefits of this kind of waveform design are in improving the quality of imaging when *extended targets* are present in the scene. While waveforms with sharp autocorrelation functions are adequate for imaging *point targets*, for extended targets, a more elaborate approach for waveform design is necessary as we outline below.

The term *redundant dictionaries* refers to a set of basis functions that are general rather than possess attributes such as orthogonality or frame conditions. For example, a basis set that is composed as the union of the canonical basis set and the Fourier basis functions is redundant, since canonical bases are not orthogonal to Fourier bases. The idea of representing signals in such basis functions has been around since the early 1990s when they were used in the context of time-frequency analysis [16]. The motivation for using redundant bases functions was explained by Mallat and Zheng in their seminal 1993 paper, using natural language as a metaphor,

We can express a wide range of ideas and at the same time easily communicate subtle differences between close concepts, because natural languages have large vocabularies that include words with close meanings. For information processing, low level signal representations must also provide explicit information on very different properties, while giving simple cues to differentiate close patterns... Linear expansions in a single basis, whether it is a Fourier, wavelet or any other basis, are not flexible enough.

While Mallat and Zhang allude to standard bases functions such as Fourier basis functions and wavelets, we can ease this restriction as we show in this thesis. Most recently, the use of such dictionaries has been proposed in the context of face recognition where the

training sets for classifiers are used as dictionary elements[64, 65]. In their 2007 paper, Wright et al. [64] model the pixels of photographs of unknown faces as a weighted linear combination of a number of different faces from a database. If the unknown test photograph is present in the database, then the weights would be zero for all mismatched images and close to a normalized numerical value of 1 for the matching image. Here, the basis functions are formed from the pixels drawn from an arbitrary image database. These in general do not have any mathematical structure. Further, one may have many more images than the number of pixels, with these two parameters being unrelated to each other. Since there is one variable (weight) for each image, what arises is an *underdetermined* system of linear equations. The authors bring in ideas from compressive sensing to leverage the fact that a vector representing the unknown faces is bound to be sparse. They propose to pre-multiply the linear system with a random matrix to make the system suitable for compressive recovery. The resulting inverse problem is solved using l_1 optimization-based algorithms. In the context of radar imaging, Varshney et al. [66] used redundant dictionaries of wavelets to characterize anisotropy in wide-angle synthetic aperture radar (SAR). They propose a graph-structured algorithm for improving SAR images based on the redundant dictionary model of anisotropic scattering.

In a manner similar to the above, we propose to use redundant dictionaries to model radar images. We then propose to use results from the theory of compressive sensing to solve the imaging problem. Since it is impractical to implement random-matrix multiplication in the temporal acquisition of radar signals, we leverage the occurrence of circulant system matrices to make the system invertible. Circulant matrices occur naturally in linear systems as they represent the operation of circular convolution. As we shall show, a practical solution to the problem of including knowledge of the target scene would involve designing appropriate waveforms matched to the targets. The basic assumption in our approach is that we know about the existence of certain objects and radar scattering profiles in the target scene. However, these are located at unknown distances and are of unknown physical extents. We construct the dictionary from that the anticipated electromagnetic scattering features of targets and clutter. For each such basis function, we also include in

the dictionary, circularly time-shifted versions of itself. The received signal is essentially a convolution of the target model and the transmit signal. The target is modeled as a weighted linear combination of these basis functions. Radar imaging now becomes the problem of inverting this linear system to solve for the unknown weights. The linear systems that model the imaging problem are underdetermined. We leverage the occurrence of sparsity in the solutions and use l_1 -norm based optimization techniques. The invertibility of the system, in such a situation is dependent on certain properties of the system matrix. We try to achieve this invertibility by designing suitable waveforms. The algorithm we propose is agnostic to the mathematical properties of the basis set.

The basic model for the signal is based on linear convolution, as in the earlier chapters. The reflected waveform is the convolution of the transmit waveform $x(t)$ and the target $s(t)$, so that, $y(t) = \int x(t - \tau)s(\tau)d\tau + \eta(t)$. The resulting discrete linear system is $\mathbf{y} = \mathbf{X}\mathbf{s} + \boldsymbol{\eta}$. We introduce additional information into the system in the form of a redundant basis set ψ_i . Then, the system becomes $\mathbf{y} = \mathbf{AZ}\mathbf{s} + \boldsymbol{\eta}$. To reflect to diversity of basis functions used in the model, the dictionary can be written as

$$\boldsymbol{\Psi} = [\boldsymbol{\Psi}^{(1)} | \boldsymbol{\Psi}^{(2)} | \dots | \boldsymbol{\Psi}^{(D)}], \quad (4.1)$$

where each of $\boldsymbol{\Psi}^{(i)} \in \mathbb{R}^{N \times N}$ with $i = 1, \dots, D$, as we argue later is a circulant matrix.

Our motivation for the ideas proposed in this chapter can be illustrated with the help of the notion of *information* sparsity. If we incorporate more information into the imaging problem in the form of the redundant dictionary of basis functions, we can reduce the amount of information that we seek to recover. However, there is always a trade-off between sparsity and uncertainty. In gaining higher sparsity, we are introducing uncertainty due to the approximate nature of our model. Our efforts are directed towards reducing the effects of this uncertainty and improving the invertibility of the system. As we outlined earlier, the performance of compressive sensing depends on the sparsity of the recovered vector, the number of linear equations, the additive noise in the system, and mathematical properties of the matrix. In Section 4.2.1 we present the theoretical basis for this argument by looking at the restricted isometry constant for the two types of matrices.

In our work, we look at radar waveform design in the context of compressive sensing. There are two reasons to study radar waveform design from a compressive-sensing perspective. Firstly, when we expand a target scene in an arbitrary dictionary of bases, the resulting linear system is highly underdetermined. The reason for this is our uncertainty regarding the exact location and nature of the targets in the image. Secondly, with the advent of digital radar systems, it has become possible to build agile and flexible radar systems which enable us, with the use of arbitrary waveform generators and custom ASICs to efficiently generate arbitrary ultra-wideband waveforms for transmission.

In earlier chapters and in recent papers on compressive radar imaging ignore correlations and degradations in sparsity due to well-defined physical phenomena. These distortions are deterministic as opposed to the additive and multiplicative random perturbations that arise from ambient noise. In the idealized compressive radar imaging problem, the system matrix is modeled as a sequence of random statistically independent and identically distributed random variables and targets are modeled as consisting of point scatterers. We believe that such idealizations and approximations are overly restrictive in certain applications, which we outline in this chapter. Further, some papers [42] approximate transmitted and reflected waveforms as being narrowband. Our focus is on ultrawideband radar imaging, which is preferable due to the robustness and improved physical resolution that the large bandwidth offers. Further, in real systems, due to physical effects, the transmit and received waveforms are really bandlimited smoothed waveforms. The smoothing can be induced by the presence of low-pass filters at various stages in the hardware, and by the electromagnetic scattering effects of nonideal targets [67]. Our answer to mitigating the effects these *deterministic nonidealities* is to use waveform design algorithms to improve the invertibility of the imaging system. In this chapter, working within the framework of discrete systems, we propose specific models for nonideality, analyze signal recovery performance, and propose approaches to mitigating the adverse effects. We propose modeling the effect of extended targets and correlated transmit waveforms by representing target scenes in redundant dictionaries. We study how these nonidealities affect the performance of compressive signal recovery in Section 4.2.1. Then, in Section 4.3, we propose an approach for

designing optimal ultrawideband waveforms for compressive radar imaging using priors and general dictionaries to represent target scenes. We analyze the behavior of these optimized waveforms in the context of compressive signal recovery by using mutual coherence as a metric. We verify the effectiveness of optimized waveforms by performing simulations for target recovery using waveforms optimized for four types of dictionaries.

4.2 Deterministic Nonidealities in Compressive Radar

4.2.1 Partial Correlated Random Matrices

The absence of correlations among the different cells of the target scene can be represented in the fact that our measurement matrix consists of a random square partial circulant matrix with the target scene represented in the canonical basis. We use the term *partial* to denote matrices formed by selecting a subset of the rows of the *full* square matrix. We enforce the condition that the measurement matrix should be circulant, as it represents what we believe is the simplest practical approach in terms of the hardware complexity to implement the compressive sensing of ultra-wideband time-domain signals that have a dense support in the frequency domain. We consider the compressive noise radar imaging problem in the presence of two types of nonidealities: (i) when the transmit waveform is distorted by a bandlimiting filter of known response, and (ii) when the target scene consists of extended scatterers. Compared to the non-idealities caused by the presence of additive and multiplicative noise, the above non-idealities allow us to utilize deterministic approaches to mitigate their effect. Hardware non-idealities have the effect of making the random variables generating the circulant matrix to be correlated. The presence of extended targets and clutter reduces the sparsity of the target scene. As we show subsequently, mathematically, the two non-idealities can be treated in a similar manner. We define the idealized received signal $y_0(t)$, such that its recoverably sampled form can be represented sparsely in the canonical basis. The transmit waveform $x(t)$ then defines the measurement basis of the compressive sensing problem, so that, $\Phi = \mathbf{X}$. The recoverability of the signal in the above idealized problem has been theoretically and empirically shown in literature [68, 69, 70, 26].

Bandlimited Waveforms

Let $h(t)$ represent a known filter that models the bandlimiting characteristics of radar hardware. Then, the transmitted signal will be

$$\tilde{x}(t) = x(t) * h(t), \quad (4.2)$$

where $x(t)$ is a fully random transmit waveform, and $*$ represents convolution. The effect of the bandlimiting filter is to *smudge* point targets. The smudged point targets however, can be recovered by representing the signal $\tilde{x}(t)$ in a more general dictionary than the canonical basis. While the smudging worsens the sparsity of the target scene as represented in the canonical domain, we can improve imaging by expanding the dictionary to include basis functions which can represent smooth signals. We can then write

$$\tilde{x}(t) = \sum_{k=1}^D x(t) * \psi_k(t) \quad (4.3)$$

$$\tilde{\mathbf{X}} = \mathbf{X}\Psi. \quad (4.4)$$

The matrix Ψ is a circulant matrix generated from the coefficients of the bandlimiting filter $\psi(t)$. For deriving restricted isometry property (RIP) results, we incorporate this nonideality by modifying the system equation thus,

$$\mathbf{z} = \mathbf{R}_\Omega \tilde{\mathbf{X}} \mathbf{s} + \mathbf{R}_\Omega \boldsymbol{\eta} \quad (4.5)$$

$$= \mathbf{R}_\Omega \mathbf{X} \Psi \mathbf{s} + \mathbf{R}_\Omega \boldsymbol{\eta}. \quad (4.6)$$

The recovery of the target scene and the solution to the convex optimization problem in Equation (4.7) both depend upon the restricted isometry property of the matrix $\mathbf{R}_\Omega \mathbf{X} \Psi$.

Extended Targets

We can compensate for the correlations represented by the band-limiting filter and extended targets by expanding the signal space to the dictionary given by the matrix Ψ . The matrix

is designed to represent correlations in the system. Since correlations smudge the target-scene vector, we synthesize a dictionary based on smooth wavelet-type bases functions. These basis functions are not necessarily orthogonal. As we show, we can always design waveforms that make the system matrix $\mathbf{X}\Psi$ convenient for compressive sensing.

In the matrix $\Psi = [\Psi^{(1)}|\Psi^{(2)}|\dots|\Psi^{(D)}]$, each sub-matrix $\Psi^{(k)}$ has circulant structure and is generated from the prior information of target profiles. The uncertainty is accounted for by increasing D . High uncertainty would require us to search through a larger dictionary, thus increasing the number of basis functions in our dictionary. The compressive radar imaging problem can be written as

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathbb{R}^{DN}} \|\mathbf{s}\|_{l_1} \text{ s.t. } \|\mathbf{z} - \mathbf{R}_\Omega \mathbf{X} \Psi \mathbf{s}\|_{l_2} \leq \sigma. \quad (4.7)$$

The matrix Ψ is designed to encode available information about the target scene. Since we are dropping the requirement that the dictionary be orthogonal, we can include prior information arising from a training set into the dictionary. In the context of radars, this prior information may include the shape of the targets as given by an arbitrary function $f(r)$ where r is a variable representing the range. The variable spatial location of the radar transmitter and receiver will result in uncertainty in fitting the received signal to this model. The spatial location, as represented by the aspect angle and the variable distance to the target will result in an expanded dictionary with basis functions given by $\sigma(r) = f(cr + a)$, with parameters a and c representing the location and extent of the target, respectively. The dictionary can thus be extended to include the parameters c and a . If we discretize the parameter space defined by (c, a) into a grid of size $D \times N$, we then have a dictionary that contains, ND elements to choose from. Thus, the general expression for the received signal is

$$s(t) = \sum_{j=1}^N \sum_{k=1}^D s_j s_k f(c_k t + a_j) \quad (4.8)$$

Discretizing, we get that the dictionary vectors, $\psi^{(k)}(t) = f(c_k t)$. Each matrix, $\Psi^{(k)}$ is the circulant convolution matrix associated with the vector $\psi^{(k)}$. It is important to note

that the function $f(\cdot)$ can be arbitrary without any requirement of orthogonality. This flexibility allows us to include prior information based on target scene \mathbf{s} . Since we are allowing for a dictionary with no special properties, we need to make sure that the system matrix $\mathbf{X}\Psi$ is suitable for compressive sensing. The design of suitable system matrices entails transmission of waveforms optimized for the dictionary. This problem of designing waveforms is addressed in Section 4.3.

4.2.2 Degradation in Restricted Isometry Constants

The matrix $\mathbf{R}_\Omega \mathbf{X}\Psi$ represents an underdetermined system. The existence of the solution to the compressive sensing optimization problem in Equation (4.7) is decided by the *restricted isometry property*. The matrix \mathbf{A} is said to satisfy the restricted isometry property of order S if for a small δ_S ,

$$(1 - \delta_S)\|\mathbf{v}\|_{l_2}^2 \leq \|\mathbf{A}\mathbf{v}\|_{l_2}^2 \leq (1 + \delta_S)\|\mathbf{v}\|_{l_2}^2 \quad \forall \mathbf{v} \text{ with } \|\mathbf{v}\|_{l_0} \leq S. \quad (4.9)$$

4.2.3 RIP for Circulant Matrices

In this section, we adopt the approach of Rauhut et al. [69] to understand the restricted isometry behavior of correlated circulant matrices. The restricted isometry property can be rewritten and simplified as

$$(1 - \delta_S)\mathbf{v}^H \mathbf{p} \leq \mathbf{v}^H \mathbf{A}^H \mathbf{A} \mathbf{v} \leq (1 + \delta_S)\mathbf{v}^H \mathbf{v} \quad (4.10)$$

$$-\delta_S \leq \mathbf{v}^H (\mathbf{A}^H \mathbf{A} - \mathbf{I}) \mathbf{v} \leq \delta_S \quad (4.11)$$

$$|\mathbf{v}^H (\mathbf{A}^H \mathbf{A} - \mathbf{I}) \mathbf{v}| \leq \delta_S. \quad (4.12)$$

Since \mathbf{A} is a fixed matrix, the above expression can be equivalently written as

$$\sup_{\mathbf{v} \in \mathbb{R}^N, \|\mathbf{v}\|_{l_0} \leq S} |\mathbf{v}^H (\mathbf{A}^H \mathbf{A} - \mathbf{I}) \mathbf{v}| = \delta_S. \quad (4.13)$$

Thus, we proceed by considering the matrix \mathbf{A} . In our problem, $\mathbf{A} = \mathbf{R}_\Omega \mathbf{X} \Psi$. We introduce the $N \times N$ permutation matrix \mathbf{S} ,

$$\mathbf{S} \triangleq \begin{pmatrix} 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix} \quad (4.14)$$

Permutation matrices \mathbf{S}^k perform the operation of circularly shifting vectors by k places. The matrices \mathbf{S}^k are constructed from permutations of the identity matrix. Permutation matrices satisfy the property that for any integer k , $(\mathbf{S}^k)^H = \mathbf{S}^{-k}$. With h representing the bandlimiting filter characterizing the hardware, we set $\Psi = \mathbf{H} \in \mathbb{R}^{(N \times N)}$, with $D = 1$. Let Γ denote the set containing the indices of non-zero coefficients of the representation of \mathbf{H} . The matrix $\mathbf{X}\mathbf{H}$ can be written as a power series as follows:

$$\mathbf{X} = \sum_{k=1}^N x_k \mathbf{S}^k, \mathbf{H} = \sum_{k=1}^N \psi_k \mathbf{S}^k, \quad (4.15)$$

$$\mathbf{A} = \mathbf{R}_\Omega \left(\sum_{k=1}^N x_k \mathbf{S}^k \right) \left(\sum_{k=1}^N \psi_k \mathbf{S}^k \right), \quad (4.16)$$

$$\mathbf{A}^H \mathbf{A} = \sum_{k1, k2 \in \Gamma \times \Gamma} \sum_{l1, l2 \in [N] \times [N]} \psi_{k1} \psi_{k2} x_{l1} x_{l2} \mathbf{S}^{-k2-l2-k1-l1} \mathbf{R}_\Omega^H \mathbf{R}_\Omega \mathbf{S}^{k2+l2+k1+l1}. \quad (4.17)$$

Alternately, we can also write $\mathbf{A}^H \mathbf{A}$ as

$$\mathbf{A}^H \mathbf{A} = \sum_{k, l \in [N] \times [N]} g_k g_l \mathbf{S}^{-k} \mathbf{R}_\Omega^H \mathbf{R}_\Omega \mathbf{S}^l, \quad (4.18)$$

$$g_i \triangleq \langle \mathbf{x}_i, \psi_i \rangle, \quad (4.19)$$

where, \mathbf{x}_i and ψ_i refer to rows of the \mathbf{X} and Ψ matrices. Following the analysis and notations of Rauhut et al. [69], the RIP analysis for correlated random matrices involves

evaluating the mean and tail bounds on the process given by

$$\delta_S = \sup_{\mathbf{v} \in \mathbb{R}^N, \|\mathbf{v}\|_{l_0} \leq S} |\mathbf{v}^H (\mathbf{A}^H \mathbf{A} - \mathbf{I}) \mathbf{v}|. \quad (4.20)$$

Rauhut et al. [69] showed that the expression for the restricted isometry constant of random circulant matrices can be written as $\tilde{\delta}_S = \sup_{\mathbf{v} \in T} |\sum_{k,l} g_k g_l' Z_{\mathbf{v}}(k,l)|$, where g_k and g_l' represent independent identically distributed random processes indexed by the subscript. We use the superscript $\tilde{\cdot}$ to identify the restricted isometry constant for random circulant matrices. For convenience, we define the set, $T \triangleq \{\mathbf{v}, \text{ such that } \mathbf{v} \in \mathbb{R}^N \text{ with } \|\mathbf{v}\|_{l_0} \leq S\}$. In order to study the restricted isometry property, we need to compare the following two quantities with the circulant random case:

$$\mathbb{E} \delta_S = \mathbb{E} \sup_{\mathbf{v} \in T} |\mathbf{v}^H (\mathbf{A}^H \mathbf{A} - \mathbf{I}) \mathbf{v}|, \quad (4.21)$$

$$\text{Prob} [\delta_S - \mathbb{E} \delta_S > \lambda] = \text{Prob} \left[\sup_{\mathbf{v} \in T} |\mathbf{v}^H (\mathbf{A}^H \mathbf{A} - \mathbf{I}) \mathbf{v}| - \mathbb{E} \delta_S > \lambda \right]. \quad (4.22)$$

We proceed [69] by defining $Z_{\mathbf{v}}(k,l) = \mathbf{v}^H \mathbf{S}^{-k} \mathbf{R}_{\Omega}^H \mathbf{R}_{\Omega} \mathbf{S}^l \mathbf{v}$. We plug in the expressions from (4.19) into (4.21) and (4.22) to get

$$\mathbb{E} \delta_S = \mathbb{E} \sup_{\mathbf{v} \in T} \left| \sum_{k,l} g_k g_l Z_{\mathbf{v}}(k,l) \right| \quad (4.23)$$

$$= \mathbb{E} \sup_{\mathbf{v} \in T} \left| \sum_{k,l} \langle \mathbf{x}_k, \boldsymbol{\psi}_k \rangle \langle \mathbf{x}_l, \boldsymbol{\psi}_l \rangle Z_{\mathbf{v}}(k,l) \right| \quad (4.24)$$

$$> C_1 \mathbb{E} \sup_{\mathbf{v} \in T} \left| \sum_{k,l} x_k x_l' Z_{\mathbf{v}}(k,l) \right|, \quad (4.25)$$

We further have that the tail of the distribution of the restricted isometry constant for correlated matrices will always be larger than the corresponding tail for uncorrelated random

circulant matrices. We can formalize this as

$$\text{Prob}[\delta_S - \mathbb{E}\delta_S > \lambda] > C_2 \text{Prob} \left[\sup_{\mathbf{v} \in T} \left| \sum_{k,l} x_k x'_l Z_{\mathbf{v}}(k,l) \right| > \mathbb{E} \sup_{\mathbf{v} \in T} \left| \sum_{k,l} x_k x'_l Z_{\mathbf{v}}(k,l) \right| + \lambda \right], \quad (4.26)$$

where x'_l are independent copies of x_l , and $C_1, C_2 > 1$ are constants. The above results follow by expanding the terms $\langle \mathbf{x}_i, \boldsymbol{\psi}_i \rangle$ and then applying decoupling inequalities [71]. The right hand sides of Equations (4.25) and (4.26) are the estimates for RIP constants of partial uncorrelated random circulant matrices [69]. The larger restricted isometry constants imply that the recovery results with correlated random circulant system matrices will be worse than the uncorrelated case, for a given level of sparsity and fixed number of measurements. The above derivations show that the restricted isometry constants are probabilistically larger. This suggests a degradation in the performance of compressive sensing. A more accurate picture of the degradation in the performance due to bandlimited waveforms can be derived in the following manner. We introduce the parameter $\gamma < 1$ such that the support of the autocorrelation function is γN .

Proposition 1. *Let the support of the autocorrelation function of the transmitted waveform be $S_{ac} = 2\gamma N$, with $\gamma < 1$. Assume that compressive recovery is guaranteed for circulant matrices generated from i.i.d random samples with the acquisition of $M > C_1 S \log N$ for some constant $C_1 > 0$. Then, in the presence of correlation, the lower limit on the number of samples required for recovery increases to $M = \min(N, C_1 \gamma S N \log N)$.*

Proof. Let \mathbf{h} denote the filter that causes the correlation in the random waveform. We can write the correlated waveform as $\tilde{\mathbf{x}} = \mathbf{x} * \mathbf{h}$, where \mathbf{x} is an i.i.d random process. The convolution between the correlated waveform and the target scene can be written as $\mathbf{y} = \mathbf{x} * \mathbf{h} * \mathbf{s} = \mathbf{X} \mathbf{H} \mathbf{s}$, where \mathbf{X} and \mathbf{H} are circulant matrices associated with the vectors \mathbf{x} and \mathbf{h} respectively. Then, $\mathbf{y} = \mathbf{X} \tilde{\mathbf{s}}$, where $\tilde{\mathbf{s}} = \mathbf{H} \mathbf{s}$. We note that the support of \mathbf{h} is γN . It follows that $\text{supp}(\tilde{\mathbf{s}}) = \text{supp}(\mathbf{H} \mathbf{s}) = \gamma N S$. If $\gamma S N \log N$ is larger than N , then it trivially follows that N equations will be sufficient for recovery. \square

The existence of extended targets similarly implies that the support of the target scene even in the absence of transmit waveform correlations is larger, leading to the requirement of acquiring more samples. Fortunately, random variables comprising the system matrix are correlated we can improve compressive sensing recovery by: (i) expanding the set of bases in the dictionary representing the signal and (ii) optimizing the transmit waveform such that the resulting system matrix is invertible based on the dictionary of basis functions. In the following sections, we describe approaches to designing such optimal waveforms.

4.3 Waveform Design Algorithm

The problem of designing optimal circulant sensing matrices was considered in a recent paper by Xu et al. [72]. The approach is to minimize the cost function that is constructed to represent the mutual coherence [39]

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X} \text{ circulant}} \|\Psi^H \mathbf{X}^H \mathbf{X} \Psi - I\|_F. \quad (4.27)$$

We use \mathbf{F} to denote the DFT matrix. Xu et al. define $\mathbf{p} = \text{abs}(\mathbf{x})^2$ and show that the algebraic simplification [72] of the cost function results in

$$\|\Psi^H \mathbf{X}^H \mathbf{X} \Psi - I\|_F^2 = \mathbf{p}^H \text{abs}(\mathbf{F}^H \Psi \Psi^H \mathbf{F})^2 \mathbf{p} - 2\mathbf{p}^H \text{diag}(\mathbf{F}^H \Psi \Psi^H \mathbf{F}) - N. \quad (4.28)$$

Xu et al. [72] thus propose solving the following optimization problem to design circulant matrices:

$$\mathcal{C}(\mathbf{p}) = \mathbf{p}^H \text{abs}(\mathbf{F}^H \Psi \Psi^H \mathbf{F}) \mathbf{p} - 2\mathbf{p}^H \text{diag}(\mathbf{F}^H \Psi \Psi^H \mathbf{F}), \quad (4.29)$$

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p} \in \mathbb{R}^N} \mathcal{C}(\mathbf{p}) \quad (4.30)$$

$$\text{subject to } \mathbf{p} \geq 0. \quad (4.31)$$

Following precedent in literature, we refer to a collection of basis functions (atoms) as molecules. The structure of the molecules allows us to simplify the cost function, and con-

sequently the computation involved in solving the optimization problem. The simplification arises from the assumption that the molecules Ψ_i are circulant. We can write each molecule as, $\Psi^{(i)} = \mathbf{F}\mathbf{\Lambda}^{(i)}\mathbf{F}^H$. We then have

$$\mathbf{B} \triangleq \mathbf{F}^H [\Psi^{(1)} | \Psi^{(2)} | \dots | \Psi^{(D)}] [\Psi^{(1)H} | \Psi^{(2)H} | \dots | \Psi^{(D)H}]^H \mathbf{F} \quad (4.32)$$

$$= \mathbf{F}^H \left(\sum_{i=1}^D \Psi^{(i)} \Psi^{(i)H} \right) \mathbf{F} \quad (4.33)$$

$$= \mathbf{F}^H \left(\sum_{i=1}^D \mathbf{F}\mathbf{\Lambda}^{(i)}\mathbf{F}^H\mathbf{F}\mathbf{\Lambda}^{(i)H}\mathbf{F}^H \right) \mathbf{F} \quad (4.34)$$

$$= \sum_{i=1}^D \mathbf{F}^H\mathbf{F}\mathbf{\Lambda}^{(i)H}\mathbf{\Lambda}^{(i)}\mathbf{F}^H\mathbf{F} \quad (4.35)$$

$$= \sum_{i=1}^D \mathbf{\Lambda}^{(i)H}\mathbf{\Lambda}^{(i)}. \quad (4.36)$$

The minimizer for the cost function will satisfy the following properties:

$$\nabla \mathcal{C}(\hat{\mathbf{p}}) = 0 \quad (4.37)$$

$$\text{abs} \left(\sum_{k=1}^D \hat{\mathbf{\Lambda}}^{(k)H} \hat{\mathbf{\Lambda}}^{(k)} \right) \hat{\mathbf{p}} = \text{diag} \sum_{k=1}^D \hat{\mathbf{\Lambda}}^{(k)H} \hat{\mathbf{\Lambda}}^{(k)} \quad (4.38)$$

$$\hat{p}_i = \frac{1}{\sum_{k=1}^D \hat{\mathbf{\Lambda}}_i^{(k)H} \hat{\mathbf{\Lambda}}_i^{(k)}}. \quad (4.39)$$

For reconstructing the waveform, we first modulate $\hat{\mathbf{p}}$ with a random phase, so that with r_i is a random variable drawn from an appropriate probability distribution,

$$u_i = \sqrt{\hat{p}_i} e^{jr_i}. \quad (4.40)$$

We wish the matrix \mathbf{X} to be real. For enforcing this condition, it is required that the eigenvalues must exist in conjugate pairs. Thus, we construct \hat{u}_i such that $\hat{u}_i = u_i, i \leq \lfloor N/2 \rfloor$, $\hat{u}_i = u_{i-\lfloor N/2 \rfloor}^H, i > \lfloor N/2 \rfloor$. The complete algorithm is summarized as Algorithm 2.

With the eigenvalues of \mathbf{X} , thus constructed, the transmit signal itself is given by

$$\mathbf{X} = \mathbf{F} \text{diag}(\hat{u}_i) \mathbf{F}^H. \quad (4.41)$$

Algorithm 2 Waveform design algorithm based on minimizing the Frobenius norm distance between the Gram matrix and the identity matrix.

Waveform Design for Ultrawideband Radar Imaging

Input: Dictionary $\Psi = [\Psi^{(1)} | \Psi^{(2)} | \dots | \Psi^{(D)}]$ describing uncertain prior information about the target scene.

Output: Optimized waveform \mathbf{x} .

- 1: Normalize the dictionary such that for each column of Ψ , $\|\psi_j^{(k)}\|_2 = 1$.
 - 2: **for** $k = 1 \rightarrow D$ **do**
 - 3: Use FFT algorithm to compute the eigenvalues $\lambda_i^{(k)}$ ($1 \leq k \leq D$ and $1 \leq i \leq N$) for each matrix $\Psi^{(k)}$.
 - 4: **end for**
 - 5: **for** $i = 1 \rightarrow N$ **do**
 - 6: Compute the vector $\hat{\mathbf{p}}$ with $\hat{p}_i = \frac{1}{\sum_{k=1}^D \Lambda_i^{(k)H} \Lambda_i^{(k)}}$.
 - 7: Randomize the phase of $\sqrt{\hat{p}_i}$, such that $u_i = \sqrt{\hat{p}_i} e^{jr_i}$ where r_i is a random number and $j = \sqrt{-1}$.
 - 8: Discard half of the values u_i so as to enforce even symmetry, and define the vector \tilde{u}_i such that $\tilde{u}_i = u_i$ for $i < \lfloor N/2 \rfloor$ and $\tilde{u}_i = u_{-i \bmod (\lfloor N/2 \rfloor)}$ for $i > \lfloor N/2 \rfloor$. This is necessary to ensure that the optimized transmit waveform is real [72] following the next step.
 - 9: **end for**
 - 10: The optimized transmit waveform is given by $\mathbf{x}^H = \mathbf{F} \tilde{\mathbf{u}}$.
-

4.3.1 Performance of Compressive Recovery with Optimized Transmit Waveform

We approach the analysis of the compressive sensing algorithm for optimized waveforms by looking at the mutual coherence of the pair of bases given by \mathbf{X} and Ψ . We continue to use the property that the matrices \mathbf{X} and Ψ are circulant. The system matrix \mathbf{A} will be the product of \mathbf{X} and Ψ . Following our notation, we recall that \mathbf{x} represents the transmit waveform generating the matrix \mathbf{X} and $\psi^{(k)}$ represents the vector generating the matrix $\Psi^{(k)}$. The symbol λ represents the Fourier co-efficients of the vector \mathbf{x} and $\lambda^{(k)}$ represent

the Fourier coefficients of the matrix $\Psi^{(k)}$. The mutual coherence between the *circulant* optimized waveform matrix \mathbf{X} and $\Psi^{(k)}$ is

$$\mu^{(k)} = \frac{|\mathbf{x}^H \Psi^{(k)}|}{\|\mathbf{x}^H\|_2 \|\Psi^{(k)}\|_2} \quad (4.42)$$

$$= \frac{|\lambda^H \mathbf{F}^H \mathbf{F} \lambda^{(k)}|}{\|\mathbf{F}\lambda\|_2 \|\mathbf{F}\lambda^{(k)}\|_2} \quad (4.43)$$

$$= \frac{\sum_{i=1}^N \frac{|\lambda_i^{(k)}|}{\sqrt{\sum_{k=1}^D \lambda_i^{(k)H} \lambda_i^{(k)}}}}{\left(\sqrt{\sum_{i=1}^N \lambda_i^{(k)H} \lambda_i^{(k)}} \right) \left(\sum_{i=1}^N \frac{1}{\sqrt{\sum_{k=1}^D \lambda_i^{(k)H} \lambda_i^{(k)}}}} \right)} \quad (4.44)$$

4.4 Simulations

In order to validate the above theory, simulations were performed to solve the one-dimensional radar imaging problem using optimized waveforms. The optimized waveform consists of 1024 samples. The target scene is represented in a redundant dictionary designed to encode the presence of extended targets. We consider that the reflected signal has an SNR of 40 dB and 25% of the samples ($M = 0.25 \times N = 256$) were acquired. The dictionary consists of eleven *molecules*. Ten of the *molecules* represent rectangular pulses of different extents. The basis functions that comprise each *molecule* matrix are time-shifted and inverted copies of the generating basis function, thus encoding the uncertainty in the locations of the targets. The last molecule is the identity matrix. It is generated from the canonical basis, and is used to represent point scatterers in the target scene. A few examples of the basis functions used are plotted in Figure 4.1. The functions shown in Figure 4.1 represent shifted and scaled copies, with no conditions of orthogonality imposed on the basis functions.

In order to quantify the improvement in compressive recovery using the waveform optimization approach, we look at the statistics of the intercolumn coherence of the system matrix. Our approach to the analysis is to first define the matrix $\mathbf{G} = \mathring{\mathbf{A}}^H \mathring{\mathbf{A}}$, where $\mathring{\mathbf{A}}$ is constructed by normalizing each column of \mathbf{A} with respect to its l_2 -norm. We then observe the statistics of the coherence across different columns by plotting the empirical cumulative distribution function (CDF) of the values. We follow precedent [73] for the

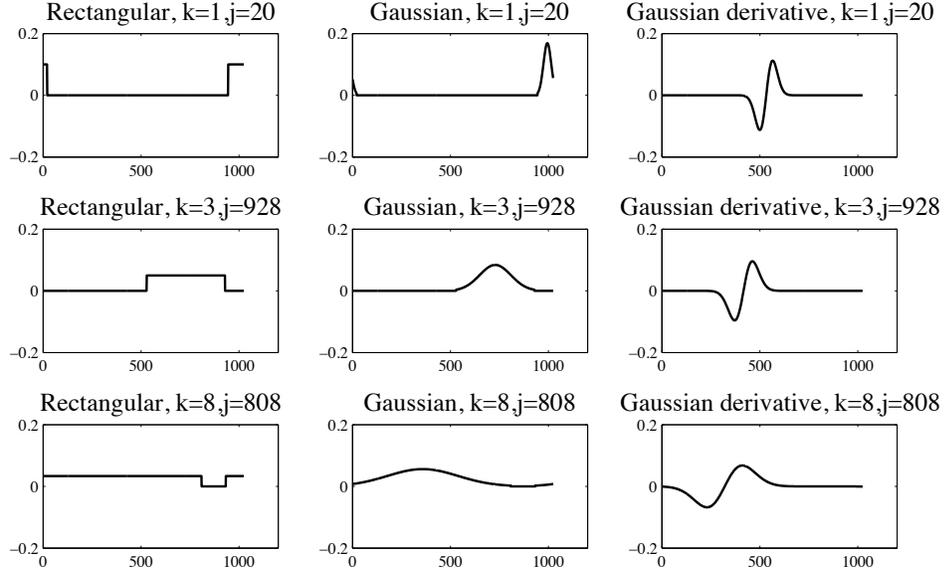
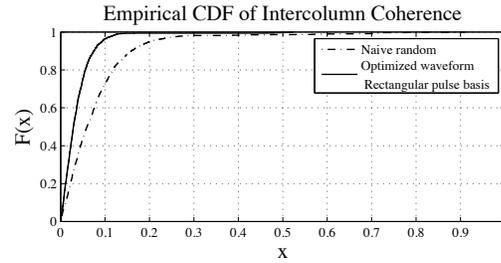


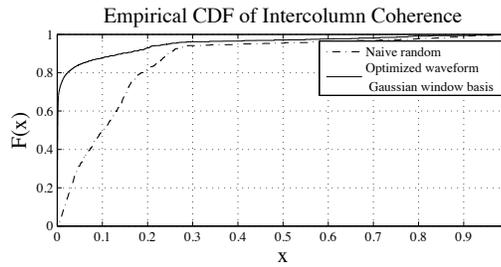
Figure 4.1: Examples of specific basis functions used in Section 4.4. These are the columns of the matrix Ψ . The index j refers to the index of the column of the basis within the matrix $\Psi^{(k)}$.

statistical evaluation of the mutual coherence metric, and construct the CDF using the formula, $F(x) = \frac{\sum_{i \neq j} \mathbb{1}_{G(i,j) < x}}{N \times (N-1)}$. The plot of these CDFs for different bases functions are shown in Figures 4.21, 4.22, and 4.23. It is observed that circulant random sensing matrices are worse in terms of the mutual coherence than optimized circulant sensing matrices. The intercolumn coherence is more likely to be higher in the case of random circulant matrices. This conforms with the theoretical arguments of Section 4.2.2. Thus when sparsity occurs in arbitrary bases functions, optimized waveforms will outperform random waveforms.

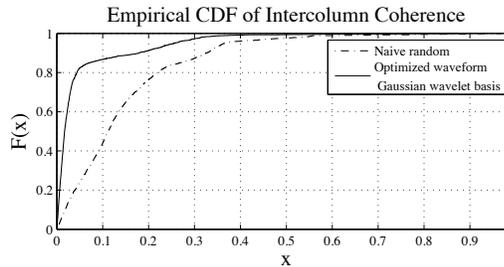
The target scene consisted of extended targets and point targets. Compressive sensing recovery was performed using the spectral projected gradient algorithm for l_1 -norms [74] to solve the optimization problem, $\min_{\mathbf{s} \in \mathbb{R}^{DN}} \|\mathbf{s}\|_{l_1}$ subject to $\|\mathbf{z} - \mathbf{R}_\Omega \mathbf{X} \Psi \mathbf{s}\|_{l_2} < \sigma$. Only 25% of the samples of the reflected waveform were used for recovery. The reflected waveform was sampled uniformly. Since the target is not sparse in the canonical basis, a naive compressive sensing approach that does not use optimized waveforms fails to recover the target scene. The recovery performance when we use a transmit waveform that is optimized for the



(1)



(2)



(3)

Figure 4.2: CDF plots comparing the inter-column coherence of system matrices corresponding to naive random and optimized waveforms.

redundant dictionary model is shown in Figure 4.4. It is seen that the naive Gaussian random waveforms are unable to simultaneously recover extended and point targets. The reconstructed target scene is shown in Figure 4.4 and the corresponding vectors $\hat{\mathbf{s}}$ are also plotted. It is seen that optimized waveforms significantly outperform the naive approach.

We repeat the simulations for target scenes represented by two other dictionaries. We looked at the dictionary comprised of molecules which are circulant matrices generated from the Gaussian window basis function. This is useful in representing smooth extended targets. The dictionary is appended with a molecule representing the canonical basis function. It is seen in Figure 4.5 that in this case too, the optimized waveform outperforms the naive

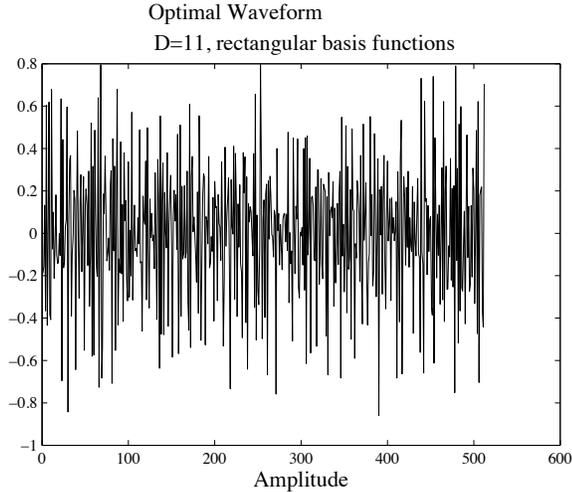


Figure 4.3: A typical transmit waveform, optimized in this simulation for a dictionary of rectangular basis functions and the identity matrix, with $D = 11$.

transmit waveform. In Figure 4.6, we simulate the signal recovery results with the assumption that the smooth target can be represented by the Gaussian-derivative wavelet function. The Gaussian-derivative wavelet function comes endowed with useful properties that are characteristic to wavelet functions- namely limited support and tightness of frames. We see that for target scenes that are sparse in such a dictionary too, recovery performance is better when optimized transmit waveforms are used. In all the above cases, we observe that the naive transmit waveform fails to pick out point targets. The reason for this is that point targets cannot be accurately represented by the continuous limited support basis functions that are placed in the dictionary to pick out extended targets. This problem is overcome by suitably designing the transmit waveform so that mutual coherence and RIP properties are retained for the arbitrary medley of basis functions that the dictionary represents. Figure 4.3 shows a typical result of the waveform optimization algorithm designed for a dictionary of rectangular basis functions and the identity matrix, with $D = 11$. For a statistical validation of the performance gain, we define the *miss-rate*, as the fraction of elements of $\hat{\mathbf{s}}$ that are recovered within $\nu = 10\%$ of the actual value. We can write this as

$$\xi = \frac{1}{N} \mathbb{1}_{|\hat{s}(k) - s(k)| \leq (\nu) |s(k)|}(k). \quad (4.45)$$

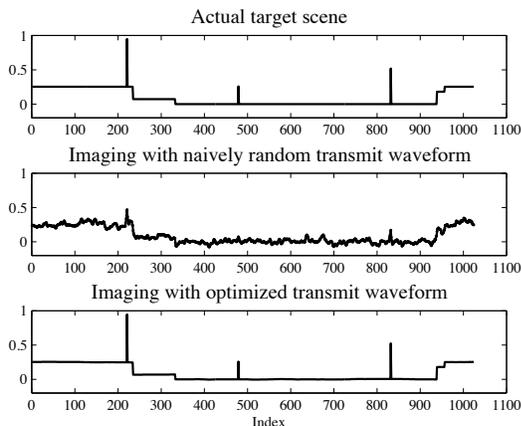


Figure 4.4: Target recovery performance (left- $\Psi \hat{\mathbf{s}}$, right- $\hat{\mathbf{s}}$) with naively random waveforms and waveforms optimized for the dictionary model generated from rectangular window basis function.

The recovery is repeated for 20 realizations. Each realization of the experiment consists of a target scene that is randomly generated using the modeling dictionary generated from the rectangular basis function. The error metric ξ is computed for each trial. This is plotted in Figure 4.7. It is seen that optimizing the waveform consistently yields an improvement in the performance of compressive recovery.

We apply the waveform design algorithm to simulated images that mimic the mathematical structure of extended targets in radar imaging scenarios. The target recovery along each point along the ‘crossrange’ represents an independent scene. Thus, the two dimensional image can be viewed as representative of a variety of target scenes. Optimized waveforms generated using Algorithm 1 work better than naive waveforms in each independent realization of the imaging problem. The targets in the example are shown in Figure 4.8. The imaging problem was simulated by performing target recovery along each of the 100 cross-range locations. The recovery results for naive transmit waveforms are shown in Figures 4.9, 4.12, 4.15, 4.11, and 4.14. The enhancement in the target recovery performance is seen in Figures 4.10, 4.13, and 4.16. While the solid targets represent mathematically smooth features, the lines in the simulated target scene are characterized by the non-smooth transitions they represent. The lines can be due to the presence of corner reflectors and sharp scatterers

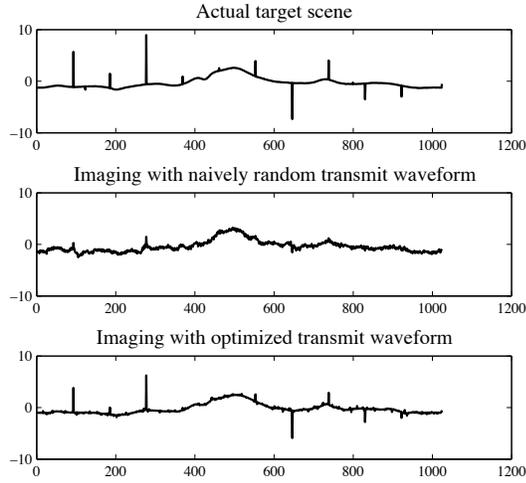


Figure 4.5: Target recovery performance (left- $\Psi \hat{\mathbf{s}}$, right- $\hat{\mathbf{s}}$) with naively random waveforms and waveforms optimized for the dictionary model generated from a Gaussian window basis function.

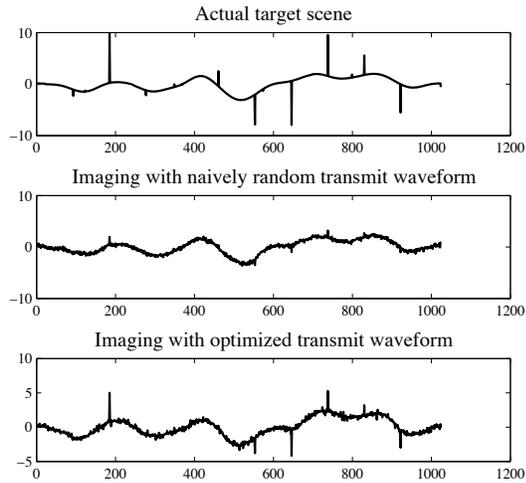


Figure 4.6: Target recovery performance (left- $\Psi \hat{\mathbf{s}}$, right- $\hat{\mathbf{s}}$) with naively random waveforms and waveforms optimized for the dictionary model generated from Gaussian derivative-wavelet basis function.

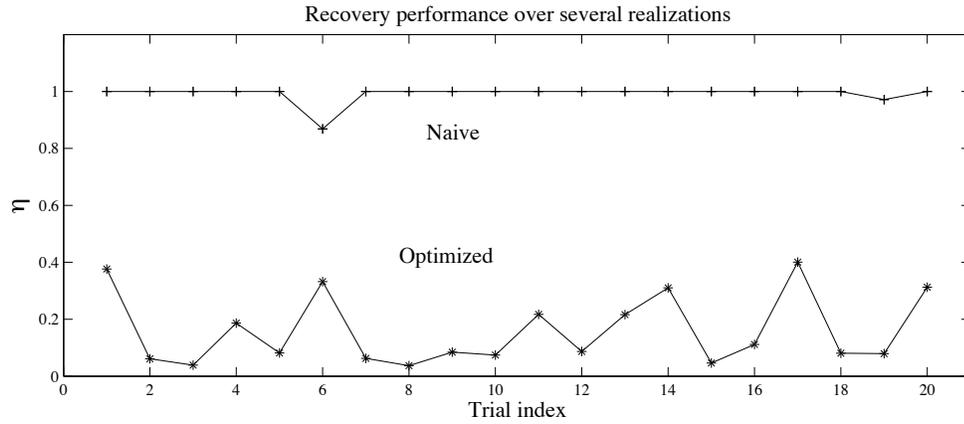


Figure 4.7: Error metric ξ for 20 realizations of the rectangular bases function case, $\nu = 0.1$.

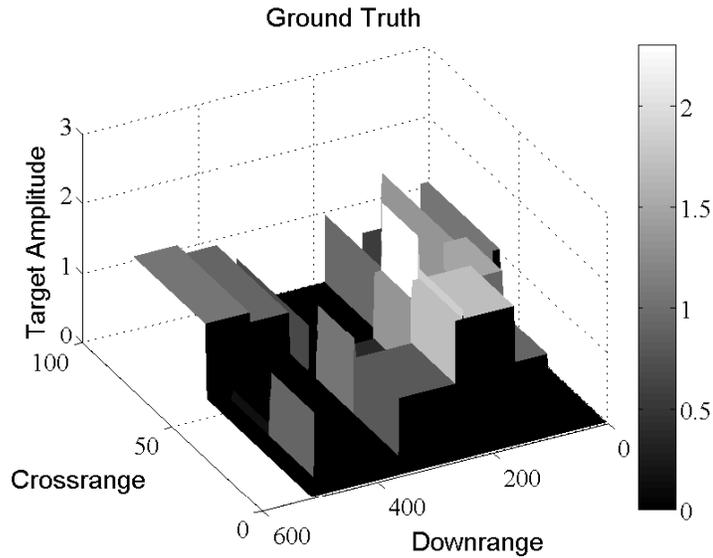


Figure 4.8: Ground truth for simulated target scene, targets representable in basis set of rectangular functions.

in the target scene. Optimized waveforms enable us to recover jointly, smooth (extended targets) as well as non-smooth (point and line scatterers) features in the target scene. We note from Figures 4.9, 4.12, and 4.15 that point scatterers are not recovered as well with naive waveforms. The optimized waveforms were designed for rectangular, Gaussian, and Gaussian wavelet basis functions so as to mimic the expected target scenes.

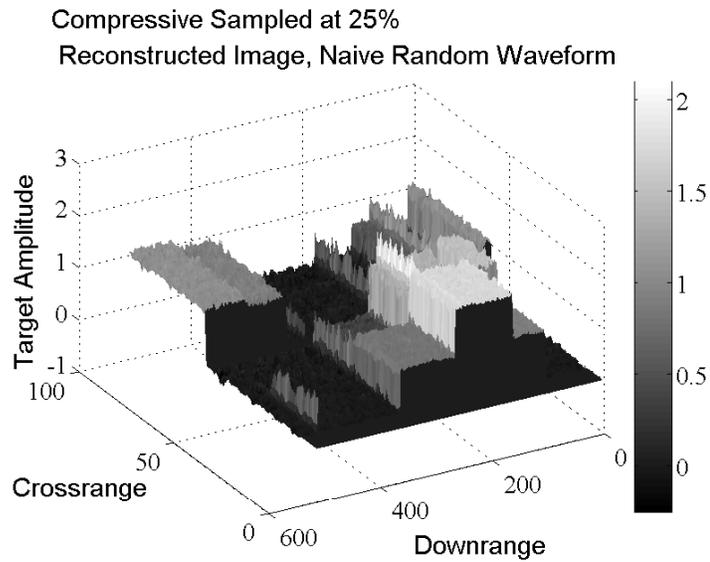


Figure 4.9: Compressive imaging using naive random waveforms, targets representable in basis set of rectangular functions.

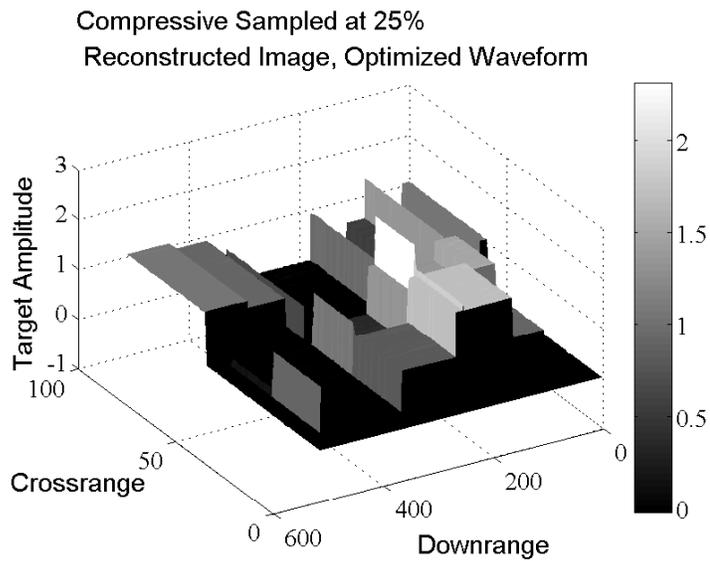


Figure 4.10: Compressive imaging using optimized random waveforms, targets representable in basis set of rectangular functions.

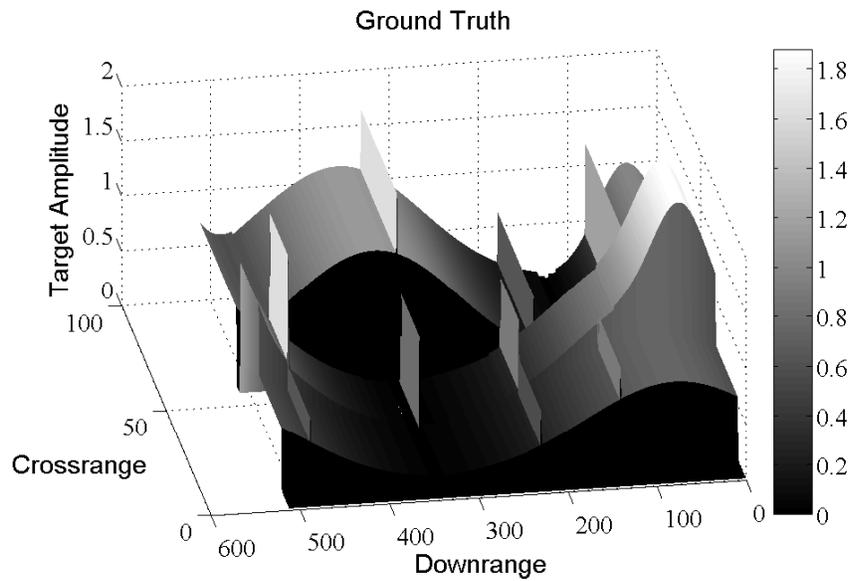


Figure 4.11: Ground truth for simulated image, targets representable in basis set of Gaussian functions.

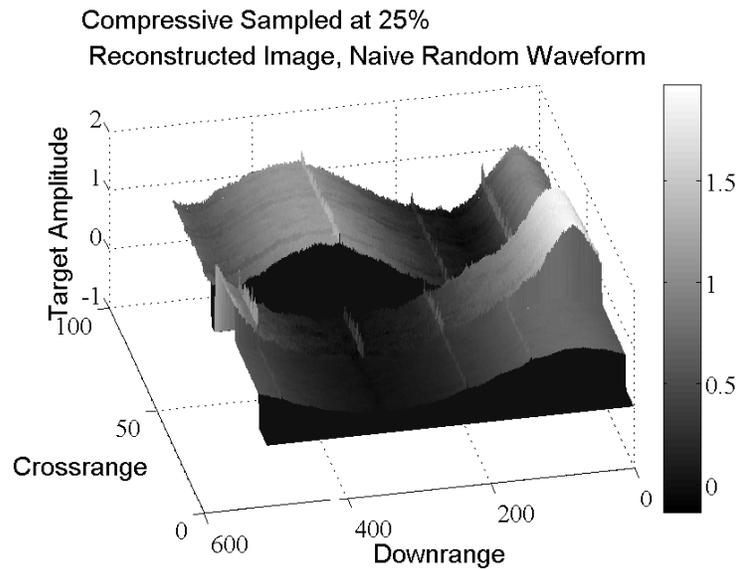


Figure 4.12: Compressive imaging using naive random waveforms, targets representable in basis set of Gaussian functions.

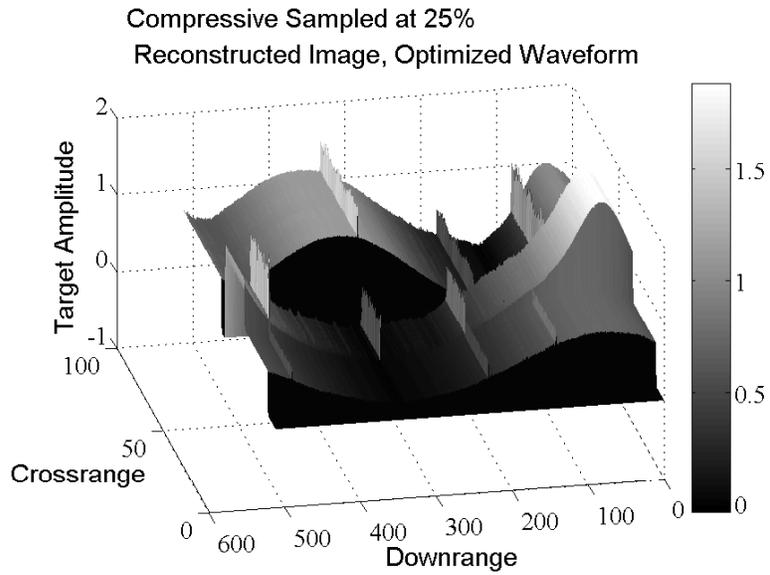


Figure 4.13: Compressive imaging using optimized random waveforms, targets representable in basis set of Gaussian functions.

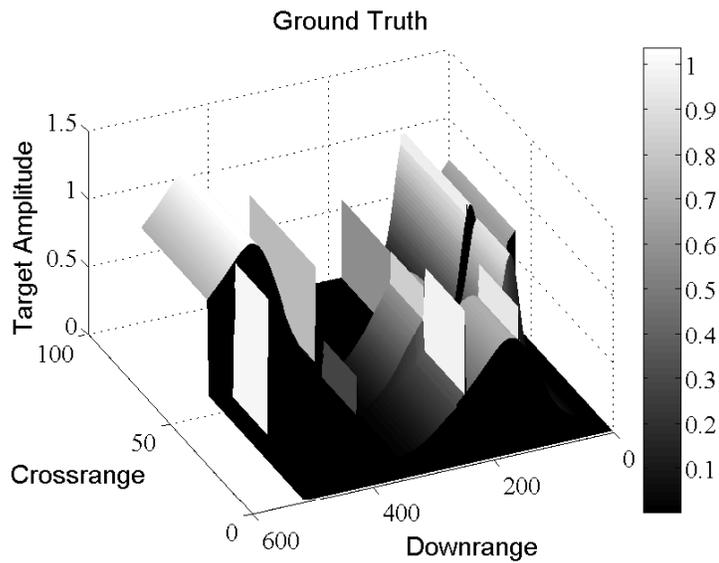


Figure 4.14: Ground truth for simulated image, targets representable in basis set of Gaussian wavelets.

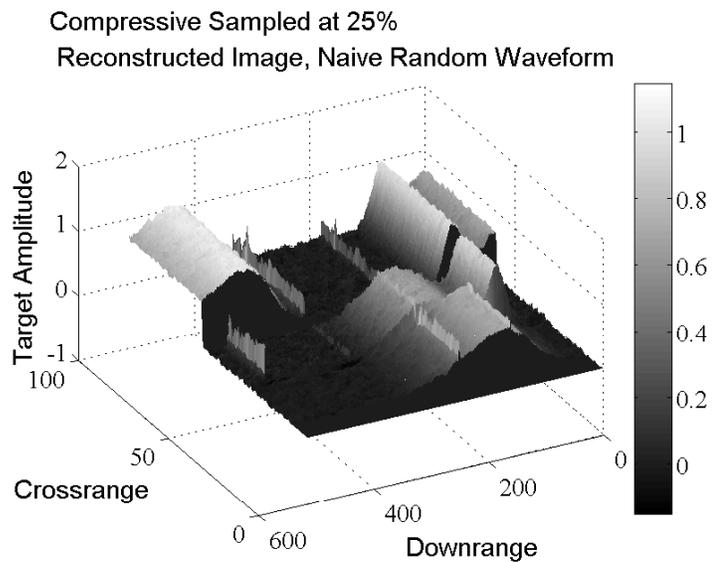


Figure 4.15: Compressive imaging using naive random waveforms, targets representable in basis set of Gaussian wavelets.

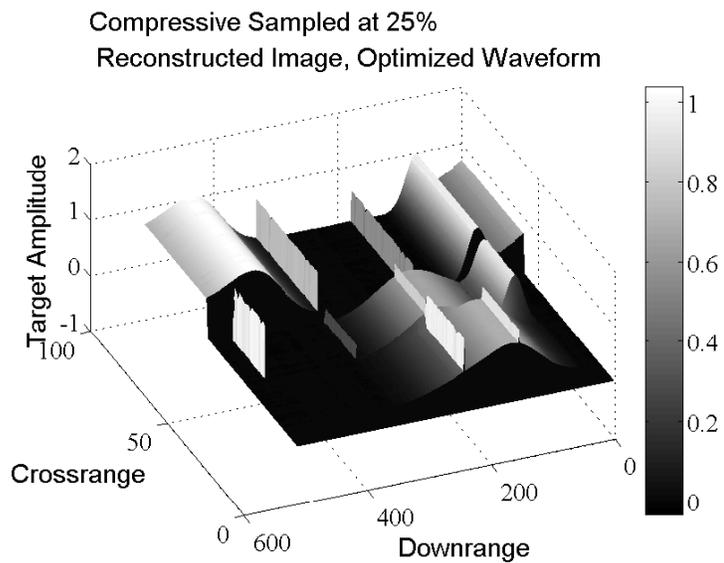


Figure 4.16: Compressive imaging using optimized random waveforms, targets representable in basis set of Gaussian wavelets.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

5.1.1 Compressive Noise Radar Imaging and Detection

In this thesis, we showed through theoretical arguments, numerical simulations, and experiments, the suitability of stochastic waveforms for designing practical compressive radar systems. We analyzed the performance of compressive noise radar systems using phase transitions. Phase transitions are feasible tools for calibrating compressive radar systems. We analyzed the importance of the uniform norm of residual error and proposed a measure of performance based on this quantity. For the first time in literature, we experimentally demonstrate the applicability of compressive sensing to ultrawideband noise radar imaging. In spite of anticipated non-idealities, the performance of compressively sampled radar compares favorably with conventional radar imaging systems. The ability of compressive sensing to resolve two closely spaced targets is also comparable with that of conventional matched-filter based radar imaging. We developed an approach based on extreme value theory to estimate the tail of compressive sensing recovery residues. The closed form of the distribution of the tail thus obtained belongs to the family of Generalized Pareto Distribution. We successfully tested our algorithms on experimental noise radar data proving that compressive noise radar imaging is a feasible technology that could replace or augment conventional noise radar systems.

5.2 Open Problems

Circulant Random Matrices in Compressive Sensing: An important problem in the area of compressive sensing as applied to radar imaging is the characterization of the performance of circulant random system matrices. Consistently, in our simulations, we observed that the performance of circulant random matrices is comparable to that of random matrices. There is strong evidence to suggest that the number of samples required for accurate compressive signal recovery with random circulant system matrices scales as $O(S \log N)$. However, existing analyses of the RIP and mutual coherence properties of circulant random matrices [26, 40, 69] only provide results that indicate suboptimal performance. Phase-transition diagrams indicate the equivalence of the behavior of circulant random and fully random matrices in the context of compressive sensing. This suggests that a theoretical analysis that takes the route of phase transitions may be used to prove optimal behavior. However, phase transitions have only been shown to theoretically work for compressive sensing problems with random system matrices.

Theory of Compressive Detection for Noise Radar: The approximate message passing algorithm [45] provides an approach to derive the distribution of the residue of compressive sensing recovery. The theoretical analyses proceeds by first showing the exact equivalence of the approximate message passing algorithm with the convex optimization formulation. The residue is hypothesized to be drawn from the normal distribution. Then, the expectation and variance of the residue are derived for the non-linear recursion that represents the approximate message passing algorithm. However, the key theorems in this analysis use results about the asymptotic behavior of random matrices. The problem of extending this theory to random circulant matrices remains open.

5.2.1 Waveform Design

In Chapter 4, we extended the theory of compressive radar imaging to account for correlated waveforms arising from extended targets and hardware non-idealities. We showed analytically the extent to which correlations degrade the performance of compressive radar

imaging. We proposed incorporating redundant dictionaries into recovering targets in the context of ultra-wideband radar imaging. By suitably optimizing transmit waveforms, we showed that it is possible to image targets that are sparsely represented in arbitrary basis. In particular, waveform optimization was useful in accurately imaging scenarios where both point and extended targets are present. Thus, the optimization of the waveforms mitigates the correlations induced by the arbitrary dictionary representation. The waveform design algorithm only involves algebraic operations on the eigenvalues of the *molecule*-matrices. This attribute of computational efficiency makes it attractive for practical compressive sensing. Simulation confirmed the theory in demonstrating that the optimization approach improves signal recovery performance by compensating for nonidealities.

Clutter Suppression: A Redundant Dictionaries Perspective In radar imaging, *clutter* refers to unwanted reflections occurring in an imaging problem. An example of radar clutter is the unwanted reflections arising from the ground on which a target of interest is located. An important area where our approach to knowledge-based compressive sensing has potential application is in clutter suppression. Imaging with redundant dictionaries opens up the possibility of including clutter models in the target imaging algorithm. In the past, statistical models have been successfully developed for characterizing clutter. Redundant dictionaries in conjunction with compressive signal recovery present an alternative deterministic approach to clutter suppression. In conventional radar systems, the reflected waveform is given by $\mathbf{z} = \mathbf{X}\mathbf{s}$. By using redundant dictionaries for imaging, we expand the signal model to $\mathbf{z} = \mathbf{X}\Psi\mathbf{s}$. Let T denote the subset of $\{1, 2, \dots, N\}$ that index the basis functions associated with the target. Let C denote the subset of $\{1, 2, \dots, N\}$ that index the basis functions associated with the clutter. We use the $\Psi^{(T)}$ and $\Psi^{(C)}$ to denote the matrix formed from basis functions indexed by T and C . The clutter and target models are separated by writing $\Psi\mathbf{s} = \Psi^{(T)}\mathbf{s}^{(T)} + \Psi^{(C)}\mathbf{s}^{(C)}$. By dividing the coefficients to \mathbf{s} into

targets and clutter, we can write the equation as,

$$\Psi \mathbf{s} = \Psi^{(T)} \mathbf{s}^{(T)} + \Psi^{(C)} \mathbf{s}^{(C)} \quad (5.1)$$

$$\mathbf{z} = \mathbf{X} \Psi \mathbf{s} = \mathbf{X} \Psi^{(T)} \mathbf{s}^{(T)} + \mathbf{X} \Psi^{(C)} \mathbf{s}^{(C)} \quad (5.2)$$

$$= \mathbf{X} [\Psi^{(T)}, \Psi^{(C)}] \begin{bmatrix} \mathbf{s}^{(T)} \\ \mathbf{s}^{(C)} \end{bmatrix} \quad (5.3)$$

$$= \mathbf{X} \Psi \mathbf{s}. \quad (5.4)$$

This gives us a way of separating features of the image resulting from the target and the clutter. By associating the co-efficients with the respective subsets T and C , one can separate targets and clutters based on arbitrary linear models. Clutter models are often probabilistic in nature [4], reflecting the uncertain nature of prior knowledge about the target scene. In the linear model we propose above, the uncertainty is quantified by the number of bases functions. Such models are appropriate in applications where the clutter does not display extreme variability. For example, in through-the-wall sensing applications [75], the main form of clutter is reflections from objects such as chairs and from phenomena such as multi-path. Since the structures in the environment remain mostly stationary, they show up as extended scatterers in radar images.

A note on basis mismatch: We expect the recovery performance of our algorithm to be sensitive to mismatch in the assumed target models and actual target scene. A corpus of training sets for target profiles in various scenarios can be generated from real world experiments and electromagnetic modeling. Such a training set can be used to enhance the imaging resolution by adopting the approach described in this Chapter 4. In the future we thus intend to deploy a practical system that uses the proposed approach. In the general context of compressive sensing, the robustness of such circulant-measurement-matrix design approaches to additive noise and basis mismatch remain exciting open problems.

5.2.2 Compressive Noise Radar Hardware

In our work, we collected fully sampled records of transmit and received waveforms and then performed the undersampling in software. However, given the incoherent nature of random noise waveforms, we believe that it should be convenient to design sampling hardware to implement sub-Nyquist sampling. In such a scheme, the random transmit waveform would be generated digitally and then converted to the analog domain for transmission. Records of the digitally generated transmit waveform can be used for performing the signal recovery. If we design a system with the ability to digitally generate UWB transmit waveforms, it can also be used to realize the waveform design approach presented in Chapter 4.

Bibliography

- [1] Jonsson, B.E., “A survey of A/D-Converter performance evolution,” *Electronics, Circuits, and Systems (ICECS), 2010 17th IEEE International Conference on*, 2010
- [2] Candes, E.J., Romberg, J., and Tao, T., “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, volume 52, no. 2, pp. 489–509, ISSN 0018-9448, February 2006
- [3] Baraniuk, Richard G, “More is less: Signal processing and the data deluge,” *Science*, volume 331, no. 6018, pp. 717–719, 2011
- [4] Richards, Mark A., *Fundamentals of Radar Signal Processing*, McGraw-Hill, 2005
- [5] Wu, Yuanbin and Li, Jinwen, “The design of digital radar receivers,” *Aerospace and Electronic Systems Magazine IEEE*, volume 13, no. 1, pp. 35–41, 1998
- [6] Horton, B. M., “Noise Modulated Distance Measuring System,” *Proceedings of the IRE*, volume 47, no. 5, pp. 821–828, May 1959
- [7] Narayanan, Ram M., Xu, Yi, Hoffmeyer, Paul D., and Curtis, John O., “Design, performance, and applications of a coherent ultra-wideband random noise radar,” *Optical Engineering*, volume 37, no. 6, pp. 1855–1869, June 1998
- [8] Candès, Emmanuel J and Wakin, Michael B, “An introduction to compressive sampling,” *Signal Processing Magazine IEEE*, volume 25, no. 2, pp. 21–30, 2008
- [9] Cotter, Shane F, Rao, Bhaskar D, Engan, Kjersti, and Kreutz-Delgado, Kenneth, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *Signal Processing IEEE Transactions on*, volume 53, no. 7, pp. 2477–2488, 2005
- [10] Lustig, Michael, Donoho, David, and Pauly, John M, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magnetic Resonance in Medicine*, volume 58, no. 6, pp. 1182–1195, 2007
- [11] Donoho, D.L. and Tanner, J., “Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing,” *Philosophical Transactions of the Royal Society A*, volume 367, no. 1906, pp. 4273–4293, Nov. 2009

- [12] Tropp, Joel A and Wright, Stephen J, “Computational methods for sparse solution of linear inverse problems,” *Proceedings of the IEEE*, volume 98, no. 6, pp. 948–958, 2010
- [13] Donoho, David L, Maleki, Arian, and Montanari, Andrea, “Message-passing algorithms for compressive sensing,” *Proceedings of the National Academy of Sciences*, volume 106, no. 45, pp. 18914–18919, September 2009
- [14] Ji, Shihao, Xue, Ya, and Carin, Lawrence, “Bayesian compressive sensing,” *Signal Processing IEEE Transactions on*, volume 56, no. 6, pp. 2346–2356, 2008
- [15] Chartrand, Rick and Staneva, Valentina, “Restricted isometry properties and nonconvex compressive sensing,” *Inverse Problems*, volume 24, no. 3, p. 035020, 2008
- [16] Mallat, Stéphane G and Zhang, Zhifeng, “Matching pursuits with time-frequency dictionaries,” *Signal Processing IEEE Transactions on*, volume 41, no. 12, pp. 3397–3415, 1993
- [17] Pati, Yagyensh Chandra, Rezaifar, Ramin, and Krishnaprasad, PS, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, IEEE, 1993
- [18] Needell, Deanna and Tropp, Joel A, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, volume 26, no. 3, pp. 301–321, 2009
- [19] van den Berg, E. and Friedlander, M. P., “Probing the Pareto frontier for basis pursuit solutions,” *SIAM Journal on Scientific Computing*, volume 31, no. 2, pp. 890–912, 2008
- [20] Boyd, Stephen and Vandenberghe, Lieven, *Convex Optimization*, Cambridge University Press, 2006
- [21] Pauluzzi, D.R. and Beaulieu, N.C., “A comparison of SNR estimation techniques for the AWGN channel,” *Communications IEEE Transactions on*, volume 48, no. 10, pp. 1681 – 1691, ISSN 0090-6778, oct 2000
- [22] Duarte, Marco F, Davenport, Mark A, Takhar, Dharmpal, Laska, Jason N, Sun, Ting, Kelly, Kevin F, and Baraniuk, Richard G, “Single-pixel imaging via compressive sampling,” *Signal Processing Magazine IEEE*, volume 25, no. 2, pp. 83–91, 2008
- [23] Herrmann, Felix J, Friedlander, Michael P, and Yilmaz, Ozgur, “Fighting the curse of dimensionality: compressive sensing in exploration seismology,” *Signal Processing Magazine IEEE*, volume 29, no. 3, pp. 88–100, 2012
- [24] Romberg, Justin and Neelamani, Ramesh, “Sparse channel separation using random probes,” *Inverse Problems*, volume 26, no. 11, pp. 115015.1–25, November 2010

- [25] Paredes, Jose L, Arce, Gonzalo R, and Wang, Zhongmin, "Ultra-wideband compressed sensing: Channel estimation," *Selected Topics in Signal Processing IEEE Journal of*, volume 1, no. 3, pp. 383–395, 2007
- [26] Haupt, J., Bajwa, W.U., Raz, G., and Nowak, R., "Toeplitz Compressed Sensing Matrices With Applications to Sparse Channel Estimation," *IEEE Transactions on Information Theory*, volume 56, no. 11, pp. 5862–5875, ISSN 0018-9448, November 2010
- [27] Berger, Christian R, Wang, Zhaohui, Huang, Jianzhong, and Zhou, Shengli, "Application of compressive sensing to sparse channel estimation," *Communications Magazine IEEE*, volume 48, no. 11, pp. 164–174, 2010
- [28] Khan, Osama Ullah, Chen, S-Y, Wentzloff, David D, and Stark, Wayne E, "Impact of Compressed Sensing With Quantization on UWB Receivers With Multipath Channel Estimation," *Emerging and Selected Topics in Circuits and Systems IEEE Journal on*, volume 2, no. 3, pp. 460–469, 2012
- [29] Cheng, Peng, Gui, Lin, Rui, Yun, Guo, Y Jay, Huang, Xiaojing, and Zhang, Wenjun, "Compressed sensing based channel estimation for two-way relay networks," *Wireless Communications Letters IEEE*, volume 1, no. 3, pp. 201–204, 2012
- [30] Meng, Jia, Yin, Wotao, Li, Yingying, Nguyen, Nam Tuan, and Han, Zhu, "Compressive sensing based high-resolution channel estimation for OFDM system," *Selected Topics in Signal Processing IEEE Journal of*, volume 6, no. 1, pp. 15–25, 2012
- [31] Liang, Quanquan, Liu, Mingyan, and Yuan, Dongfeng, "Channel estimation for opportunistic spectrum access: uniform and random sensing," *Mobile Computing IEEE Transactions on*, volume 11, no. 8, pp. 1304–1316, 2012
- [32] Baraniuk, R. and Steeghs, P., "Compressive Radar Imaging," *2007 IEEE Radar Conference*, Waltham, MA, April 2007
- [33] Chen, Pin-Heng, Shastry, Mahesh C., Lai, Chieh-Ping, and Narayanan, R. M., "A Portable Real-Time Digital Noise Radar System for Through-the-wall Imaging," *IEEE Transactions on Geosciences and Remote Sensing*, volume 50, no. 10, 2012, in press
- [34] Gurbuz, Ali Cafer, McClellan, James H, and Scott, Waymond R, "A compressive sensing data acquisition and imaging method for stepped frequency GPRs," *Signal Processing IEEE Transactions on*, volume 57, no. 7, pp. 2640–2650, 2009
- [35] Suksmono, Andriyan Bayu, Bharata, Endon, Lestari, Andrian Andaya, Yarovoy, Alexander G, and Ligthart, Leo P, "Compressive stepped-frequency continuous-wave ground-penetrating radar," *Geoscience and Remote Sensing Letters IEEE*, volume 7, no. 4, pp. 665–669, 2010
- [36] Bar-Ilan, Omer and Eldar, Yonina C, "Sub-Nyquist Radar via Doppler Focusing," *arXiv preprint arXiv12110722*, 2012

- [37] Mishali, Moshe, Eldar, Yonina C, Dounaevsky, Oleg, and Shoshan, Eli, “Xampling: Analog to digital at sub-Nyquist rates,” *Circuits Devices Systems IET*, volume 5, no. 1, pp. 8–20, 2011
- [38] Ender, Joachim H G, “On compressive sensing applied to radar,” *Signal Processing*, volume 90, no. 5, pp. 1402–1414, ISSN 0165–1684, special Section on Statistical Signal and Array Processing, 2010
- [39] Donoho, D.L., “Compressed sensing,” *IEEE Transactions on Information Theory*, volume 52, no. 4, pp. 1289–1306, ISSN 0018–9448, April 2006
- [40] Romberg, Justin, “Compressive Sensing by Random Convolution,” *SIAM Journal on Imaging Sciences*, volume 2, no. 4, pp. 1098–1128, 2009
- [41] Rauhut, Holger, Romberg, Justin, and Tropp, Joel A., “Restricted isometries for partial random circulant matrices,” *Applied and Computational Harmonic Analysis*, , no. 0, ISSN 1063-5203, 2011
- [42] Herman, Matthew and Strohmer, Thomas, “High-resolution radar via compressed sensing,” *IEEE Transactions on Signal Processing*, volume 57, no. 6, pp. 2275–2284, June 2009
- [43] Oppenheim, Alan V., Schafer, Ronald W., and Buck, John R., *Discrete-time signal processing (2nd ed.)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, ISBN 0-13-754920-2, 1999
- [44] Tropp, J.A. and Wright, S.J., “Computational Methods for Sparse Solution of Linear Inverse Problems,” *Proceedings of the IEEE*, volume 98, no. 6, pp. 948–958, June 2010
- [45] Bayati, Mohsen and Montanari, Andrea, “Dynamics of message passing on dense graphs, with applications to compressed sensing,” *IEEE Transactions on Information Theory*, volume 57, no. 2, pp. 764–785, February 2011
- [46] Montanari, Andrea and Bayati, Mohsen, “The LASSO risk for Gaussian matrices,” *IEEE Transactions on Information Theory*, volume PP, no. 99, p. 1, ISSN 0018–9448, 2011
- [47] Donoho, D.L., Maleki, A., and Montanari, A., “The Noise-Sensitivity Phase Transition in Compressed Sensing,” *Information Theory IEEE Transactions on*, volume 57, no. 10, pp. 6920–6941, Oct. 2011
- [48] Anitori, Laura, Maleki, Arian, Otten, Matern, Baraniuk, Richard G, and Hoogetboom, Peter, “Design and Analysis of Compressed Sensing Radar Detectors,” , 2013
- [49] Tang, Gongguo and Nehorai, Arye, “Verifiable and computable performance analysis of sparsity recovery,” *Pre print*, 2012

- [50] Davenport, Mark A, Boufounos, Petros T, Wakin, Michael B, and Baraniuk, Richard G, “Signal processing with compressive measurements,” *Selected Topics in Signal Processing IEEE Journal of*, volume 4, no. 2, pp. 445–460, 2010
- [51] Duarte, Marco F., Davenport, Mark A., Wakin, Michael B., and Baraniuk, Richard G., “Sparse signal detection from incoherent projections,” in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), III*, 2006
- [52] Broadwater, J.B. and Chellappa, R., “Adaptive Threshold Estimation via Extreme Value Theory,” *IEEE Transactions on Signal Processing*, volume 58, no. 2, pp. 490–500, ISSN 1053-587X, Feb. 2010
- [53] Ozturk, A., Chakravarthi, P. R., and Weiner, D. D., “On determining the radar threshold for non-Gaussian processes from experimental data,” *IEEE Transactions on Information Theory*, volume 42, no. 4, pp. 1310–1316, ISSN 0018-9448, Jul 1996
- [54] Pickands III, James, “Statistical Inference Using Extreme Order Statistics,” *The Annals of Statistics*, volume 3, no. 1, pp. 119–131, ISSN 00905364, Jan. 1975
- [55] Gu, M., Lim, L.-H., and Wu, C. J., “PARNES: A rapidly convergent algorithm for accurate recovery of sparse and approximately sparse signals,” *ArXiv e prints*, November 2009
- [56] Woodward, Philip M, *Probability and information theory, with applications to radar*, volume 3, Pergamon Press London, 1953
- [57] Bell, Mark R., “Information theory and radar waveform design,” *Information Theory IEEE Transactions on*, volume 39, no. 5, pp. 1578–1597, 1993
- [58] Yang, Yang and Blum, Rick S, “MIMO radar waveform design based on mutual information and minimum mean-square error estimation,” *Aerospace and Electronic Systems IEEE Transactions on*, volume 43, no. 1, pp. 330–343, 2007
- [59] Goodman, Nathan A, Venkata, Phaneendra R, and Neifeld, Mark A, “Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors,” *Selected Topics in Signal Processing IEEE Journal of*, volume 1, no. 1, pp. 105–113, 2007
- [60] Kay, Steven, “Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation,” *Selected Topics in Signal Processing IEEE Journal of*, volume 1, no. 1, pp. 31–41, 2007
- [61] Sowelam, Sameh M. and Tewfik, Ahmed H., “Waveform selection in radar target classification,” *Information Theory IEEE Transactions on*, volume 46, no. 3, pp. 1014–1029, 2000
- [62] Capraro, Gerard T, Farina, Alfonso, Griffiths, Hugh, and Wicks, Michael C, “Knowledge-based radar signal and data processing: a tutorial review,” *Signal Processing Magazine IEEE*, volume 23, no. 1, pp. 18–29, 2006

- [63] Shastry, Mahesh C., Narayanan, Ram M., and Rangaswamy, Muralidhar, “Waveform design for compressively sampled UWB radar,” *Journal of Electronic Imaging SPIE Special Issue on Compressive Sensing for Imaging*, April 2013
- [64] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., and Ma, Yi, “Robust Face Recognition via Sparse Representation,” *Pattern Analysis and Machine Intelligence IEEE Transactions on*, volume 31, no. 2, pp. 210–227, ISSN 0162-8828, feb. 2009
- [65] Srinivas, U., Chen, Yi, Monga, V., Nasrabadi, N.M., and Tran, T.D., “Exploiting Sparsity in Hyperspectral Image Classification via Graphical Models,” *Geoscience and Remote Sensing Letters IEEE*, volume 10, no. 3, pp. 505–509, ISSN 1545-598X, may 2013
- [66] Varshney, Kush R, Çetin, Müjdat, Fisher, John W, and Willsky, Alan S, “Sparse representation in structured dictionaries with application to synthetic aperture radar,” *Signal Processing IEEE Transactions on*, volume 56, no. 8, pp. 3548–3561, 2008
- [67] Fannjiang, Albert C., “Compressive Imaging of Subwavelength Structures,” *SIAM Journal on Imaging Sciences*, volume 2, no. 4, pp. 1277–1291, 2009
- [68] Romberg, Justin, “Compressive Sensing by Random Convolution,” *SIAM Journal on Imaging Sciences*, volume 2, no. 4, pp. 1098–1128, 2009
- [69] Rauhut, Holger, Romberg, Justin, and Tropp, Joel A., “Restricted isometries for partial random circulant matrices,” *Applied and Computational Harmonic Analysis*, volume 32, no. 2, pp. 242–254, ISSN 1063-5203, 2012
- [70] Shastry, Mahesh C, Narayanan, Ram M, and Rangaswamy, Muralidhar, “Compressive radar imaging using white stochastic waveforms,” *Waveform Diversity and Design Conference*, Niagara Falls, Canada, August 2010
- [71] de la Peña, Victor H. and Giné, Evarist, *Decoupling: From Dependence to Independence*, Springer, 1999
- [72] Xu, Yangyang, Yin, Wotao, Chen, Susan, and Osher, Stanley, “Learning a Circulant Sensing Matrix,” , 2012
- [73] Ertin, Emre, Potter, Lee C, and Moses, Randolph L, “Sparse target recovery performance of multi-frequency chirp waveforms,” *19th European Signal Processing Conference (EUSIPCO2011)*, 2011
- [74] van den Berg, E. and Friedlander, M. P., “SPGL1: A solver for large-scale sparse reconstruction,” [Http://www.cs.ubc.ca/labs/scl/spgl1](http://www.cs.ubc.ca/labs/scl/spgl1), June 2007
- [75] Ahmad, Fauzia and Amin, Moeness G, “Multi-location wideband synthetic aperture imaging for urban sensing applications,” *Journal of the Franklin Institute*, volume 345, no. 6, pp. 618–639, 2008

- [76] Candès, Emmanuel J, “The restricted isometry property and its implications for compressed sensing,” *Comptes Rendus Mathématique*, volume 346, no. 9, pp. 589–592, 2008

Appendix

Mathematical Derivations

A.1 Compressive Signal Recovery: A Model-Selection Perspective

Consider the problem of estimating the variable $s \in \mathbb{R}^n$ from its projections $y \in \mathbb{R}^m$. In our problem, s is the vector representing the radar target scene and y is the undersampled reflected waveform. If $\boldsymbol{\eta}$ represents the uncorrelated noise corrupting the system,

$$\mathbf{z} = \mathbf{A}\mathbf{s} + \boldsymbol{\eta} \tag{A.1}$$

Let the set $D \subset \{1, 2, \dots, n\}$ and $\text{cardinality}(D) = d$ denote an optimal model for estimating the signal. This means that the vector s is zero on D^c .

Further we will assume that the projection \mathbf{A} preserves the energy of any vector $v \in \mathbb{R}^d$ up to a factor of δ ,

$$(1 - \delta)\|\mathbf{v}\|_{l_2} \leq \|\mathbf{A}_D\mathbf{v}\|_{l_2} \leq (1 + \delta)\|\mathbf{v}\|_{l_2} \tag{A.2}$$

Note that A.2 is equivalently, a bound on the eigenvalues of $\mathbf{A}_D^T\mathbf{A}$, i.e., $1 - \delta \leq \lambda(\mathbf{A}_D^T\mathbf{A}_D) \leq 1 + \delta$.

Best Model Known *a priori*: Let us now assume that somehow, we possess prior knowledge of the best model, i.e., knowledge of the set D . In a practical system this would imply that we measure only the subset D . Then,

$$y = \mathbf{A}_D s_D + \mathbf{A}_{D^c} s_{D^c} + \eta \quad (\text{A.3})$$

Let \mathbf{A}_D represent the $m \times d$ vector that measures only on the best model. Then, the estimated vector s , optimal in the least square sense will be,

$$\mathbf{s}_D^* = (\mathbf{A}_D^T \mathbf{A}_D)^{-1} \mathbf{A}_D^T y \quad (\text{A.4})$$

This estimated variable is non-zero only on the set of indices given by D . We can write the estimation error as,

$$\|\mathbf{e}\|_{l_2} = E\{\|\mathbf{s}_D - \mathbf{s}_D^*\|^2\} + \sum_{i \notin D} s_i^2 \quad (\text{A.5})$$

Now, let us consider the term $\mathbf{s}_D - \mathbf{s}_D^*$, we can write it as follows:

$$\mathbf{s}_D - \mathbf{s}_D^* = (\mathbf{A}_D^T \mathbf{A}_D)^{-1} \mathbf{A}_D^T (\mathbf{A}_D s_D + \mathbf{A}_{D^c} s_{D^c} + \eta - \mathbf{A}_D s_D) \quad (\text{A.6})$$

$$= (\mathbf{A}_D^T \mathbf{A}_D)^{-1} \mathbf{A}_D^T (\mathbf{A}_{D^c} s_{D^c} + \eta) \quad (\text{A.7})$$

By taking the auto-correlation, the mean square error can be written as,

$$E\{\|\mathbf{s}_D - \mathbf{s}_D^*\|_{l_2}^2\} = (\mathbf{A}_D^T \mathbf{A}_D)^{-1} \mathbf{A}_D^T (\mathbf{A}_{D^c} s_{D^c} + \sigma^2 \text{Trace}[(\mathbf{A}_D^T \mathbf{A}_D)^{-1}]) \quad (\text{A.8})$$

$$\geq \sigma^2 \text{Trace}[(\mathbf{A}_D^T \mathbf{A}_D)^{-1}] \quad (\text{A.9})$$

$$\geq \sigma^2 \frac{d}{1 + \delta} \quad (\text{A.10})$$

Thus, since $\delta < 1$

$$\|\mathbf{e}\|_{l_2}^2 \geq \sum_{i \in D^c} s_i^2 + \frac{1}{2} d \sigma^2 \quad (\text{A.11})$$

When we choose the best model, D , we would have the bound on the error being the minimum of the error over all sets D , i.e.,

$$\|e\|_{l_2}^2 \geq \frac{1}{2} \min_D (\|s - s^*\| + d\sigma^2) \quad (\text{A.12})$$

$$\|e\|_{l_2}^2 \geq \sum_i \min(s_i^2, \sigma^2) \quad (\text{A.13})$$

Blind Reconstruction from Undersampled Points: Compressive sensing estimates s *blindly* in the following way, by minimizing the l_1 norm.

$$\mathbf{s}^* = \arg \min_{\mathbf{s} \in \mathbb{R}^n} \|\mathbf{s}\|_{l_1} \text{ s.t. } \|\mathbf{z} - \mathbf{A}\mathbf{s}\| \leq \lambda\sigma \quad (\text{A.14})$$

According to [76], if $\lambda = \sqrt{2 \log p}$, with a high probability, the above minimization problem results in an error bound of,

$$\|e\|_{l_2}^2 \leq O(\log p) \left(\sum_i \min(s_i^2, \sigma^2) \right) \quad (\text{A.15})$$

Therefore, comparing A.13 and A.15, the penalty of using the blind reconstruction method characteristic of compressive sensing is that the error bound is scaled up by a function that asymptotically varies as the logarithm of the noise level estimate.

Vita of Mahesh C. Shastry

My main interests are in the theory and applications of linear algebra, optimization, harmonic analysis, electromagnetics, numerical methods, and probability theory in electrical engineering. Secondly, I am interested in discrete nonlinear dynamics, statistical inference, and statistical physics.

Select Publications:

1. *MC Shastry, RM Narayanan, M Rangaswamy, **Waveform Design for Compressive Radar Imaging Using Redundant Dictionaries***, accepted for publication in Journal of Electronic Imaging (SPIE) Special Section on Compressive Sensing for Imaging.
2. *MC Shastry, RM Narayanan, M Rangaswamy, **Performance Characterization of Compressively Sampled Noise Radar***, in preparation.
3. *MC Shastry, RM Narayanan, M Rangaswamy, **Compressively Sampled Noise Radar***, invited book chapter- in preparation.
4. *S Jolad, A Roman, MC Shastry, **A Bounded Divergence Measure Based on The Bhattacharyya Coefficient***, Arxiv pre-print
5. *PH Chen, MC Shastry, CP Lai, RM Narayanan, **Design and Analysis of Portable Through-Wall Radar***, IEEE Transactions on Geosciences and Remote Sensing, Vol. 50, No. 10, 4123-4134, Oct. 2012.
6. *RM Narayanan, MC Shastry, PH Chen, M Levi, **Through-the-Wall Detection of Stationary Human Targets Using Doppler Radar***, Progress in Electromagnetics Research B, Vol. 20, pp. 147-166, **2010**
7. *N Nagaraj, MC Shastry, PG Vaidya, **Increasing Average Period Lengths by Switching of Robust Chaos Maps in Finite Precision***, *European Physical Journal Special Topics, Nonlinear Dynamics and Chaos: Selected Problems, Vol. 165, pp 73 - 83, 2008*