Compressive Sensing Meets Noise Radar

Mahesh C. Shastry, Ram M. Narayanan, and Muralidhar Rangaswamy

CONTENTS

13.1 Introduction .................................................................. 430
13.1.1 State of the Art in Compressive Radar Imaging .......... 433
13.2 Basics of Compressive Stochastic Waveform Radar .......... 435
13.2.1 Compressive Radar ................................................. 435
13.2.2 Correlations in the Circulant Matrix ............................. 437
13.2.3 Experiments ...................................................... 437
13.2.4 Analysis of Experimental Data .................................... 439
13.2.5 Imaging Performance ............................................... 439
13.3 Detection Strategies for Compressive Noise Radar ............ 442
13.3.1 Compressive Sensing Detection .................................. 442
13.3.2 Statistics of the Error of Compressive Signal Recovery ...... 445
13.3.3 Threshold Estimation for Compressive Detection .......... 448
13.3.4 GPD and Compressive Sensing ................................... 450
13.3.5 Computational Complexity of GPD-Based Threshold Estimation .................................................. 451
13.3.5.1 Performance of GPD-Based Threshold Estimation .......... 452
13.4 Conclusions and Future Work ............................................ 455
13.4.1 Compressive Noise Radar Imaging and Detection .......... 455
13.4.2 Open Problems ...................................................... 456
References ..................................................................... 457

ABSTRACT In this chapter, we discuss how noise radar systems are suitable for realizing practically the promises of compressive sensing in radar imaging, in general, and in urban-sensing applications, in particular. Noise radar refers to radio frequency imaging systems that employ transmit signals that are generated to resemble random noise waveforms. Noise radar has recently been successfully applied to urban sensing applications such as through-the-wall sensing (Amin 2011). Recent advances in the field of compressive sensing provide us with techniques to overcome the challenges of waveform design, sampling, and bandwidth constraints. We review existing literature related to these problems and present new results that enable...
us to leverage compressive sensing and sparsity to improve noise radar systems. We model compressively sampled noise radar imaging as a problem of inverting linear system with a circulant random system matrix. We demonstrate the feasibility of this model by applying it to experimental data acquired using a millimeter wave ultrawideband noise radar system. Our principal contributions lie in developing theory and algorithms for imaging and detection strategies in compressively sampled noise radar imaging. We outline an approach based on extreme value statistics that works by empirically estimating the distribution of the residue of instances of the estimation algorithm. False alarms are treated as statistically rare events for estimating event probabilities in the compressive detection problem. We extrapolate the distribution of the residue from a small number of recovery instances to calibrate compressive noise radar systems. For deploying compressively sensed noise radar systems in real applications, it is necessary to develop convenient approaches to calibrate and characterize recovery performance.

13.1 Introduction

Imaging using noise radar can be cast as a linear inverse problem. 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. Information about the targets of interest is extracted by comparing the transmitted and reflected waveforms. In this chapter, 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. We concern ourselves primarily with imaging problems that involve building range profiles using noncoherent noise radar technology.

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. For many years, conventional radar systems used analog processing systems for imaging. One of the main reasons for this is the desire to maximize the bandwidth of the transmit signals. 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 (Wu and Li 1998), 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 converters (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 even more advanced signal processing techniques at the receivers to improve the functionality of radar systems. The trends in the evolution of ADC technology from 1975 to 2010 are shown in Figure 13.1 (the figure has been adapted from a 2010 survey) (Jonsson 2010). Figure 13.1 is a plot of the peak sampling frequency achieved by state-of-the-art ADCs for different effective number of bits (ENOBs). The figure represents the tradeoff between sampling frequency and the levels of quantization achievable by the ADC. As seen in Figure 13.1, the best 12-bit quantization ADCs available today can sample at a rate of less than 1 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. Traditionally, the received signals are noncoherently processed using a matched filter for range detection. The idea of using noise waveforms for radar imaging was first proposed in Horton (1959). Over the
last decade (Narayanan et al. 1998), the trend has been toward the adoption of ultrawideband signals for noise radar. Ultrawideband waveforms are defined as signals that have a bandwidth of either 500 MHz or higher or 20% of the center frequency. In our applications, we use waveforms of 500 MHz 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 \( \psi(t) \) and additive noise \( n(t) \), the reflected signal is simply given by

\[
f(t) = \psi(t) * s(t) + n(t), \tag{13.1}
\]

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

\[
f = \Psi s + n, \tag{13.2}
\]

where \( f, s, n \in \mathbb{R}^N \) and \( \Psi \in \mathbb{R}^{N \times N} \). With the orthogonality assumption, the recovered target is given by

\[
s_{\text{CR}}^* = \Psi^{-1}f = \Psi^T f. \tag{13.3}
\]

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 chapter, we propose to use a general optimization-based approach to noise radar imaging as opposed to traditional cross-correlation processing. The discrete \( l_2 \)-norm is defined for any vector \( v = (v_1, v_2, \ldots, v_N) \) as \( \|v\|_2 = \sum_{i=1}^{N} |v_i|^2 \). The discrete \( l_1 \)-norm is defined as \( \|v\|_1 = \sum_{i=1}^{N} |v_i| \). The recovery of sparse solutions to underdetermined linear systems is possible in the framework of compressive sensing. We extend the signal model such that the system matrix \( A = \Phi \Psi \in \mathbb{R}^{M \times N} \). Such a system matrix corresponds to sampling the received signal at a rate dictated by the information bandwidth rather than the signal bandwidth. We cast the problem as one of minimizing a least-squares-based cost function such as \( \min_s \|f - \Psi s\|_2 \) with an additional \( l_1 \) regularization term. While least-squares processing can potentially be more computationally expensive 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 will be shown in Section 13.2, with the least-squares formulation, we can utilize sparsity to more efficiently
sample the signals. Least-squares and sparsity-based approaches are harder to analyze due to the usage of nonlinear and iterative recovery schemes. Further, target recovery involves significant computational cost. In Section 13.3, we propose to overcome these problems with the help of an efficient data-driven detection algorithm. We validate our theoretical and empirical results by analyzing experimental noise–radar data acquired using a millimeter-wave radar system operating with a bandwidth of 500 MHz.

13.1.1 State of the Art in Compressive Radar Imaging

Random noise radar (Horton 1959) involves transmitting waveforms that are generated as stochastic processes. The technology is headed toward utilizing ultrawideband transmit waveforms (Narayanan et al. 1998) 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. Increasingly noise radar systems are using digital processing (Chen et al. 2012) for imaging in real time. Noise radar systems have been shown to be a viable technology for through-the-wall imaging in general and urban sensing in particular (Narayanan 2008). Digital noise radar systems use high rate ADC. The sampling rates of the ADC limit the maximum achievable range resolution of the system. In order to circumvent this bottleneck, we employ compressive sensing principles. The application of compressive sensing to through-the-wall imaging using stepped frequency radar was first reported in Yoon and Amin (2010). Practical compressive radar systems have been implemented for ground penetrating radar applications using stepped frequency waveforms (Gurbuz et al. 2009; Suksmono et al. 2010). Bar-Ilan and Elder (Bar-Ilan and Eldar 2014) applied the Xampling (Mishali et al. 2011) framework to develop a pulse radar prototype for joint range velocity radar imaging using Doppler focusing. Ender (Ender 2010) presented results on compressive radar imaging using 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 (Narayanan et al. 1998).

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 (Baraniuk and Steeghs 2007) 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 (Candes et al. 2006; Donoho 2006). 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 (Romberg 2009) 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. (Haupt et al. 2010) 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 (Candes et al. 2006) recovery algorithm. Rauhut et al. (Rauhut et al. 2012) derived the RIP for circulant matrices, suggesting that $O \left( \max \left( (S \log N)^{1.5}, S \log^2 S \log^2 N \right) \right)$ measurements guarantee stable recovery. Herman and Strohmer derived mutual coherence results for high-resolution radar imaging (Herman and Strohmer 2009) 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. Noise waveforms also allow for target-matched waveform design in the context of compressive sensing (Shastry et al. 2013b). Our work extends the state of the art in considering ultrawideband random waveform radar systems that 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 toward real-world applications. Simulations indicate that the number of measurements follows the optimal compressive sensing asymptotics of $O(S \log N)$ (Shastry et al. 2010).

Portions of this chapter were presented in conference publications (Shastry et al. 2010, 2012, 2013a).

The summary of this chapter is as follows:

1. We formulate compressive sensing as a problem involving the inversion of a linear system with circulant random matrices. In Section 13.2, 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 nonlinear convex optimization solvers. This complicates the task of target detection. First, while the target location is sparse,
the radar measurement is seldom sparse. Therefore, not accounting for this mismatch results in a performance loss. 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 Section 13.3, 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.

13.2 Basics of Compressive Stochastic Waveform Radar

13.2.1 Compressive Radar

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

$$f(t) = \int_{-\infty}^{\infty} \psi(\tau-t)s(\tau) d\tau + n(t). \quad (13.4)$$

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

$$f_i = \sum_k \psi_{i-k} s_k + n_i, \quad (13.5)$$

$$f = \Psi s + n, \quad (13.6)$$

where $f, s, n \in \mathbb{R}^N$ and $\Psi \in \mathbb{R}^{N \times N}$. In real situations, the vector $s$ is typically sparse. For signals of finite duration, this equation is represented exactly by a linear system with a Toeplitz system matrix $\hat{\Psi}$ so that $\hat{f} = \hat{\Psi} s + n$. We approximate this as a linear system with a circulant system matrix $\Psi$, in order to simplify computation and analysis. In our experiments, the substitution of the Toeplitz matrix with a circulant matrix is justified because the nonzero elements occur close to the beginning of the data record. Physically, this is due to the fact that we use ultrawideband waveforms. With a bandwidth of 500 MHz, as in our system, even a few microseconds of acquisition results in
information related to a much larger range than typically useful. The error in recovering accurately at the ends of the signal is intrinsic to processing signals of finite time duration (Oppenheim et al. 1999). This type of error can only be overcome by increasing the dimensionality of our problem. Experimental justification for this approximation is illustrated by the accuracy of the experimental results described in Section 13.2.3.

In sparse target scenarios, compressive sensing theory allows us the luxury to undersample \( y(t) \). We model this operation of undersampling as a premultiplication by the measurement matrix \( \Phi = R_\Omega \in \mathbb{R}^{M \times N} \), where \( R_\Omega \) consists of the rows of the identity matrix indexed by the set \( \Omega \subset \{1,2,\ldots,N\} \). Thus, we have

\[
y = R_\Omega (f + n) = R_\Omega \Psi s + R_\Omega n,
\]

with \( y \in \mathbb{R}^M \). The performance of the compressive radar ranging problem depends on the properties of the matrix \( R_\Omega \Psi \triangleq A \in \mathbb{R}^{M \times N} \). 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 \( ||s||_0 \leq S \) with \( S \ll N \). We define the quantities \( \rho \triangleq S/M \) and \( \delta \triangleq M/N \). If the matrix \( A \) satisfies certain properties for vectors of given sparsity, then the problem can be inverted by solving the following convex optimization problem, called basis pursuit denoising (BPDN):

\[
\text{BPDN}(\rho, \delta; \sigma) : \min_{s \in \mathbb{R}^N} ||s||_1 \text{ subject to } ||y - As||_2 \leq \sigma.
\]

This specific formulation of the problem is chosen over other formulations (Tropp and Wright 2010) of compressive sensing because it is useful for practical implementations. Specifically, the value \( \sigma \) 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 (Pauluzzi and Beaulieu 2000). In either case, the radar operator has access to an approximate estimate of the value \( \sigma \).

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 (Candes et al. 2006) that estimate the upper bound of residual error. Practical applications of compressive sensing require more detailed analyses of the residual error. With this in mind, the authors of the present chapter proposed using phase transition diagrams (Shastry et al. 2010) for characterizing radar systems.
13.2.2 Correlations in the Circulant Matrix

In practical compressive sensing systems, as we shall see in Section 13.2.4, transmit waveforms are not ideal. The assumption that transmit waveforms can be discretized into independent identically distributed (i.i.d.) random process is convenient for theoretical analysis of compressive recovery. In real systems, however, 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 $\tilde{\psi}(t)$ and a transfer function $h(t)$ that represents the bandlimiting nonidealities so that

$$\psi(t) = \tilde{\psi}(t) \ast h(t)$$  \hspace{1cm} (13.9)

$$\Psi = \tilde{\Psi} H.$$  \hspace{1cm} (13.10)

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 nondiagonal elements of the normalized Gram matrix $G$ to be as low as possible. We look at the error metric given by

$$\chi(G) = \sum_{i \neq j} |G_{i,j}|^2.$$  \hspace{1cm} (13.11)

In Figure 13.2, we plot the values of $\chi(G)$ for various values of $G = \Psi H_l$, where $l$ denotes the width of the power spectrum of the filters. We characterize the matrix $H_l$ 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 nearly white and subsequently the sequence of random variables modeling the waveform are 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 conforms with the experimental observations that we outline in Section 13.2.4. Even with a low-pass filter that only allows about 70% of the spectrum of the transmit waveform, we see in Figure 13.2 that $\chi$ is almost as good as the uncorrelated case.

13.2.3 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. A sample transmit waveform is shown in Figure 13.3. We operated the ADC at a rate of 1 gigasamples per second. The system
Compressive Sensing for Urban Radar

FIGURE 13.2
The effect of filtering on the TPSF. On the y-axis is $\chi(G)$. It is seen that the performance of compressive sensing is expected to deteriorate if the random transmit waveforms are highly correlated (narrow filters).

FIGURE 13.3
Time-domain plot of a transmit waveform sampled at 1 GS/s. The signal was generated using the experimental noise radar setup.

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 setup is shown in Figure 13.4. We tested the imaging capability of the system at
distances ranging from 14 m to around 33 m. We used tetrahedral corner reflectors and cylindrical scatterers as targets.

13.2.4 Analysis of Experimental Data

Compressive sensing recovery with circulant matrices is possible when the waveform that generates the circulant system matrix is i.i.d. random variables. Our experiments indicate that (a) waveforms in real systems are correlated as seen in Figure 13.5 which deviate from the normal distribution as seen in Figure 13.6, and (b) compressive signal recovery is tolerant to some extent to the 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 13.2, the degradation may not be significant as long as the correlation is not high. This is affirmed by the accuracy of recovery seen in Figures 13.7 through 13.11.

13.2.5 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 \( s_{LS}^{*} = \Psi^{-1}f \) and correlation processing \( s_{CR}^{*} \) in Equation 13.3. The scenarios and the
FIGURE 13.5
Power spectrum estimates (using a covariance estimator) of a millimeter-wave transmit waveform.

FIGURE 13.6
QQ plot of the normalized transmit waveform for millimeter-wave radar in comparison with standard normal.
related experimental analyses are described subsequently. 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 33 m from the antennas is illustrated in Figure 13.7. We chose a corner reflector because most of the energy backscatter is guaranteed to be along the antenna line of sight. This scenario is 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 14 m. The second cylinder is placed at 0.6 m and 0.3 m in the scenarios presented in Figures 13.8 and 13.9, respectively. 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 33 m and separated by distances of 0.3 m and 1 m in Figures 13.10 and 13.11, respectively.
13.3 Detection Strategies for Compressive Noise Radar

13.3.1 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 (Richards 2005). In conventional radar systems, 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. The recovered signal is given by
Two cylinders at 14 m, separated by 0.3 m

FIGURE 13.9
Two cylindrical targets at a distance of about 14 m (40 ft) from the radar and separated from each other by 0.3 m.

\[ \hat{s} = \arg \min_{s \in \mathbb{R}^N} ||s||_1 \text{ s.t. } ||y - As||_2 \leq \sigma. \]  

(13.12)

Our detection problem seeks the significance of each recovered pixel of the vector \( s \). The null hypothesis is the absence of a target at each pixel, \( s(k) \):

\[ \mathcal{H}_0^{(k)} : s(k) = 0 \]  

(13.13)

\[ \mathcal{H}_1^{(k)} : s(k) \neq 0. \]  

(13.14)

Relating this ground truth hypothesis with the recovered target vector \( \hat{s} \), we obtain the false alarm and detection probabilities for each pixel, \( \hat{s}(k) \),

\[ P_D^{(k)} = \mathbb{P}[\hat{s}(k) > \xi | \mathcal{H}_1^{(k)}] \]  

(13.15)

\[ P_{FA}^{(k)} = \mathbb{P}[\hat{s}(k) > \xi | \mathcal{H}_0^{(k)}]. \]  

(13.16)
Two corner reflectors separated by 0.9 m

Cross-correlation image $s_{CR}^*$

Least square $s_{LS}^*$

BPDN with 100% samples $s_{CS}^* (\delta, 1)$

BPDN with 25% undersampling $s_{CS}^* (\delta, 0.25)$

**FIGURE 13.10**
Two corner reflectors separated by a distance of 1 m (3 ft), located at a distance of 33 m (108 ft) from the radar.

$P_D^{(k)}$ 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. In the context of sparse radar imaging, fixing the threshold for detecting and defining significant targets is an important task. In sparse target scenes, with a large percentage of the target cells being zero, the cost of false alarms is proportionally higher.

However, the problem of casting the theory and performance of compressive radar imaging in the context of conventional radar systems remains open. The first paper to explore this problem from the perspective of signal processing was the 2010 paper on compressive signal processing in Davenport et al. (2010). Their proposal was to use a threshold detection
Two corner reflectors separated by 1.5 m

FIGURE 13.11
Two corner reflectors separated by a distance of 1.5 m (5 ft), located at a distance of 33 m from the radar.

approach based on the sufficient statistic given by $z^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}\hat{s}$. 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 publications related to the present section are the results in Anitori et al. (2013) 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 to accurate and stable solutions for problems with circulant random matrices. Thus, we develop a data-driven approach to threshold detection in compressive noise radar.

13.3.2 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 KS 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 cumulative distribution function (CDF) of data is compared with the standard normal distribution. The results of the KS test are given in Table 13.1. The qq- plots in Figures 13.12 and 13.13 compare the quantiles of empirical data with that of theoretical estimates. Deviations

**TABLE 13.1**

Table Summarizing the Kolmogorov–Smirnov Test Comparing a Few Instances of the Normalized Data with the Standard Normal Distribution

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Null Hypothesis: Standard Normal</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic</td>
<td>Reject</td>
<td>0</td>
</tr>
<tr>
<td>Data set 1</td>
<td>Reject</td>
<td>0</td>
</tr>
<tr>
<td>Data set 2</td>
<td>Reject</td>
<td>0</td>
</tr>
<tr>
<td>Data set 3</td>
<td>Reject</td>
<td>0</td>
</tr>
<tr>
<td>Data set 4</td>
<td>Reject</td>
<td>0</td>
</tr>
</tbody>
</table>

**FIGURE 13.12**

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.
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 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.

When convex optimization recovery schemes such as $l_1$-based recovery are used, the nonlinear and iterative nature of the estimator algorithm coupled with the arbitrary evolution of the iterates complicate the task of deriving theoretical expressions for the distribution of the residue. The computational cost of convex optimization solvers presents a significant barrier to employing a brute force Monte Carlo approach to determine detection thresholds. 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 with reasonable accuracy. This becomes a problem when we require the compressive radar system to compare favorably with conventional radar systems, where it is not unusual to see systems with probability of false alarm as low as $P_{fa} \sim 10^{-4}$.

There have been extensions to Monte Carlo simulations in the past to accommodate the occurrence of rare events (Broadwater and Chellappa 2010) in the context of radar signal processing. One such approach involves estimating probabilities using extreme value theory (Ozturk et al. 1996; Broadwater and Chellappa 2010). We propose applying to compressive noise radar, an approach based on limit theorems from extreme value theory.
Extreme value theory refers to the study of probabilities of rare events in stochastic systems. The basic results concern the statistics 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. (Ozturk et al. 1996). In the context of compressive sensing, our proposal is to extrapolate the probabilities of rare events from a few instances of solving the convex optimization problem. A manageable number of instances of convex optimization problem are used to generate the statistics of compressive sensing for various values of $(\rho, \delta)$ and these are used to compute thresholds for small values of $P_{fa} (<10^{-4})$.

### 13.3.3 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. Determination of the exact number of nonzero coefficients is the classical model order selection problem. This remains an open problem with respect to $l_1$-norm optimization-based signal recovery. A subjective technique for a real radar system may, for example, involve using knowledge acquired via a 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 $\{1, 2, \ldots, N\}$ about which we have knowledge that targets are absent. By observing the statistics of members of the set $Z$, across different realizations of the convex optimization solver for different $\xi$, we can determine the false alarm rate. Let $k_z^{(i)}$ denote elements of subsequence of $\{1, 2, \ldots, N\}$, such that, for each $k_z \in Z$,

$$P_{FA} = \mathbb{P}[s^*(k_z^{(i)}) > \xi]. \quad (13.17)$$

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

$$P_{FA}(\xi) = \int_{\xi}^{\infty} p_z(x) dx. \quad (13.18)$$

If we were to use Monte Carlo simulations to estimate these probability distributions, for a given $P_{fa}$, we require $Q$, the number of realizations used
for estimation to satisfy the requirement that \( Q \gg 1/P_{FA} \). This requirement implies that brute-force Monte Carlo simulations are impractical when each realization of \( s^* (k_2^{(i)}) \) is being generated from a convex optimization solver. Following past research (Ozturk et al. 1996; Broadwater and Chellappa 2010) 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

\[
G(x) \triangleq 1 - \left( 1 + \frac{\gamma x}{\zeta} \right)^{-\frac{1}{\gamma}}, \quad (13.19)
\]

with

\[
-\infty < \gamma < \infty, \quad \zeta > 0, \quad 0 > \gamma x \leq -\zeta. \quad (13.20)
\]

GPD parametrizes numerous other distributions such as the exponential distribution when \( \gamma = 0 \) and the uniform distribution when \( \gamma = -1 \). As proved by Pickands (Pickands 1975), the following result relates the GPD to the tails of general, unspecified distributions:

\[
\lim_{n \to \infty} P \left[ \sup_{0 \leq y < \infty} \left| P[Y > y + u | Y \geq u] - 1 + G(y) \right| > \epsilon \right] = 0, \quad (13.21)
\]

\[
\forall \epsilon > 0. \quad (13.22)
\]

Now we note that the conditional expectation given by

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

(13.23)

Setting \( z = u + y \), we get

\[
F(z) = F_u(z - u)(1 - F(u)) + F(u).
\]

(13.24)

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

\[
F(z) = \alpha F_u(z - u) + (1 - \alpha).
\]

(13.25)

We proceed by estimating the limit of \( F_u(z - u) \) based on Equation 13.21 so that

\[
F_u(z - u) = G(z) = 1 - \left( 1 + \frac{\gamma (z - u)}{\zeta} \right)^{-\frac{1}{\gamma}}.
\]

(13.26)
and

\[ F(z) = \alpha \left( 1 - \left( 1 + \frac{\gamma}{\zeta}(z - u) \right)^{-\frac{1}{\gamma}} \right) + (1 - \alpha) \]  

(13.27)

\[ = 1 - \alpha \left\{ 1 + \frac{\gamma}{\zeta}(z - u) \right\}^{-\frac{1}{\gamma}}. \]  

(13.28)

Thus, the general strategy for applying GPD to estimate rare-event probabilities is to first set a particular \( \alpha \) and then estimate the tail of the unknown distribution using (13.27). 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×(1−\( \alpha \))th percentile of the data. The parametric function is then fitted to the given data to arrive at values for \( \gamma \) and \( \zeta \). We follow the proposal of Ozturk et al. (1996) and employ the Nelder–Mead algorithm for solving the maximum-likelihood formulation given by

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

(13.29)

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

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

(13.30)

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

(13.31)

### 13.3.4 GPD and Compressive Sensing

With the aforementioned 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, \( s \), the noise \( \eta \), and residue \( s - \hat{s} \). Our methodology for deriving the thresholds for compressive sensing is as follows:

The threshold is computed from the probability of false alarm as

\[ \tau^{(CS)}(P_{FA}) = u + \frac{\hat{\zeta}}{\hat{\gamma}} \left( \left( \frac{P_{FA}}{\alpha} \right)^{-\hat{\gamma}} - 1 \right). \]  

(13.32)
With the relationship between $P_{FA}$ and $\tau$ established, we can proceed to derive the probability of detection from the statistics of the nonzero values and establish the receiver operating characteristics.

### 13.3.5 Computational Complexity of GPD-Based Threshold Estimation

In our work, we use convex optimization to solve the compressive signal recovery problem. While there are more computationally efficient approaches, convex optimization provides the best performance in terms of recovery accuracy (Tropp and Wright 2010). The recent algorithm of approximate message passing (AMP) (Donoho et al. 2009) 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) (van den Berg and Friedlander 2008) 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 (Gu et al. 2012) 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 $N_{mc}$, and (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 tail distribution allows us to estimate thresholds for very low values of $P_{fa}$ without significantly increasing $N_{mc}$ and $N$.

In a real system, our goal is to compute detection thresholds with as little latency as possible. An improvement of a factor of $(1/\gamma_{mc}\gamma_s) \sim 10^4$ in the computational speed of estimating thresholds represents a significant advantage when each solution of the SPGL1 algorithm takes around 300 s for a 10,000-dimensional problem, as we see in Figure 13.14. The computational expense of convex optimization-based signal recovery presents a
significant stumbling block in running extensive Monte Carlo simulations. In Figure 13.14, we present the computational cost as a function of problem size. The numbers pertain to a MATLAB® implementation of the SPGL1 algorithm (van den Berg and Friedlander 2007) on an Intel i7, 2.8 GHz system with 8 GB RAM.

A single solution of a problem of dimension 10,000 takes about 5 min. 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, it would take us in excess of 5 min 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 that can yield low false alarm rates using brute-force Monte Carlo methods.

### 13.3.5.1 Performance of GPD-Based Threshold Estimation

We demonstrate the effectiveness of Algorithm 13.1 to data generated from applying the SPGL1 convex optimization solver to synthesized compressive
Algorithm 13.1 GPD-based tail, $P_{FA}$, and threshold estimation

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

**Input:** $X$, $s$.

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

1. for $j \in Z$ do
   1. for $i = 1 \rightarrow Q$ do
      1. Solve the convex optimization problem given by BPDN($\rho$, $\delta$, $\sigma$);
      1. Use the entire recovered vector $\hat{s}$ as a training set, i.e., $T = \{1, 2, \ldots, N\}$ 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
   1. Using the points $r(j) : j \in T$, construct the distribution $\hat{p}_{r(j)}(x) = \sum_{t=1}^{Q} \mathbb{1}_{r(j)^{(t)}}(x)$;
end for

Set $\alpha = 0.1$, and select $u = r(b)$ such that $\#\{t : r(t) > r(b)\} = [\alpha Q]$, and let $T \triangleq \{t : r(t) > r(b)\}$; Let $z^{(u)}$ be a sequence such that $\forall j, z^{(u)}_j \in T$; Estimate GPD parameters $\mu$ and $\zeta$ by applying the Nelder–Mead solver to solve the maximum likelihood function optimization problem given by $\min_{\gamma, \zeta} \alpha Q \log \zeta + \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^{\alpha Q} \log \left(1 + \frac{\zeta}{\gamma} (z_i^{(u)} - u)\right)$. Call this solution $\hat{\gamma}^{(CS)}$ and $\hat{\zeta}^{(CS)}$.

noise radar problems. In Figure 13.15, 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 $\alpha$ has to be chosen carefully based on the number of reliable nonzero values in the training data. A high value of $\alpha$ will mean that the training data contain too many samples too far from the tail. If $\alpha$ 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 13.1. From these estimates, the threshold is computed for various values of the probability of false alarm from Equation 13.32. The 50 realizations are plotted as smoothed probability density function in Figure 13.16, estimated as

$$p_\tau(x) = \sum_{i=1}^{50} \mathbb{1}_\tau(x),$$
Compressive Sensing for Urban Radar

GPD estimate (red), \( \gamma = -0.3952, \zeta = 0.0319 \)

\[ F(x) \]

FIGURE 13.15
(See color insert.) Plot showing application of Algorithm 13.1 for GPD-based extrapolation of the cumulative distribution of empirical residue data. Compressive recovery with \( N = 1024 \).

\( \hat{T} \) represents the set of all numbers larger than \( \hat{\tau} \), which is the estimated threshold value. The indicator function \( \mathbb{1}_{\hat{T}}(x) \) takes the value of 1 when \( x \in \hat{T} \) and 0 otherwise. In the simulations, we used the \texttt{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 variance of the parameter estimates. The reliable median allows us to extract a meaningful estimate even when the desired probability of false alarm is as low as \( 10^{-8} \).

We applied our threshold estimation algorithm on experimental data that was acquired as described in Section 13.2. 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 nonzero 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 13.1. 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 13.17.
FIGURE 13.16
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 $SNR = 10 \text{ dB}$. 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 13.1.

13.4 Conclusions and Future Work

13.4.1 Compressive Noise Radar Imaging and Detection

In this chapter, we showed through theoretical arguments, numerical simulations, and experiments, the suitability of stochastic waveforms for
designing practical compressive radar systems. In spite of anticipated non-idealities, the performance of compressively sampled radar compares favorably with conventional radar imaging systems. 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 GPD. 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.

13.4.2 Open Problems

Going forward, we have identified two important open problems related to compressive noise radar imaging:

1. Developing an extension of the approximate message passing state evolution (Bayati and Montanari 2011) framework to circulant-matrix based compressive sensing. This provides us an approach to derive expressions for detection statistics.

FIGURE 13.17
(See color insert.) Tail estimation for real data. We observe that the empirical CDF corresponds closely to 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 Section 13.2.
2. Designing a sampling system that is suitable for high-throughput compressive sampling applications such as ultrawideband noise radar. The current system can be used to achieve higher resolutions however, for a generally applicable system, it is necessary to design a hardware system for undersampling signals. The principal challenge is in keeping track of the index and temporal location of each acquired sample.

References


