8-2013

Synthetic aperture radar with compressed sensing

Yufeng Cao
University of Texas-Pan American

SYNTHETIC APERTURE RADAR WITH COMPRESSED SENSING

A Thesis

by

YUFENG CAO

Submitted to the Graduate School of the
University of Texas-Pan American
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2013

Major Subject: Mathematical Sciences

SYNTHETIC APERTURE RADAR WITH COMPRESSED SENSING

A Thesis
by
Yufeng Cao

COMMITTEE MEMBERS

Dr. Zhijun Qiao
Chair of Committee

Dr. Andras Balogh
Committee Member

Dr. Zhaosheng Feng
Committee Member

Dr. Eleftherios Gkioulekas
Committee Member

August 2013

ABSTRACT

Cao, Yufeng., Synthetic Aperture Radar with Compressed Sensing. Master of Science (MS), August, 2013, 33 pp., 11 figures, references, 24 titles.

A general synthetic aperture radar (SAR) signal model is derived from the Maxwell's equations, and compressed sensing are introduced to the signal model for SAR image reconstruction. Random Partial Fourier Matrices were applied to prove that compressed sensing can be used to this signal model from the viewpoint of mathematics. In the numerical simulation part, we show that the procedure of basis pursuit can reconstruct SAR image, based on our main results, which is shown efficient in comparison with the matched filter algorithm.

DEDICATION

This thesis is dedicated to my parents Xiaozhong Cao and Lingzhi Jia, and my sister Yuxian Cao. Thank you for all your love and support, with your love, I can conquer anything.

ACKNOWLEDGMENTS

I am grateful to my advisor Dr. Zhijun Qiao, for his guidance and patience while helping me formulate this work. I would also like to express my sincere thanks to other students in Dr. Qiao's research group.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

SYNTHETIC APERTURE RADAR

I.1 Synthetic Aperture Radar (SAR)

Synthetic aperture radar (SAR) is a remote sensing technique that uses an antenna, which is mounted on a moving platform, to image a stationary target scene. In most cases, SAR antennas are either mounted on airborne or spaceborne platforms, such as airplanes or satellites. These antennas use highly directed microwave radiation to produce images of specific patch of the Earth's surface. Unlike most remote imaging systems such as visible and infrared systems, SAR uses active imaging [9]. This means that the portion of the ground that is to be imaged must first be illuminated by a beam of microwave energy emitted by a transmitting antenna, rather than relying on passive sources of illumination such as solar radiation. SAR systems are also preferred over other types due to their ability to produce high quality images through cloud cover and at night [8].

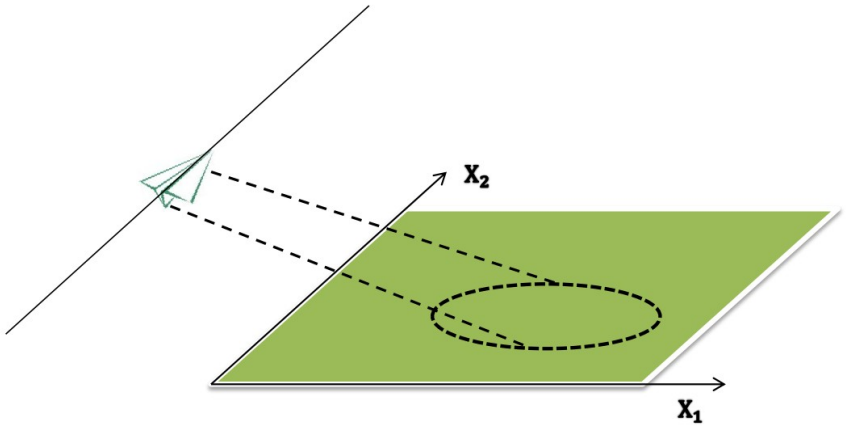

Figure I.1: Illustration of synthetic aperture radar

In the early 1950's, Carl Wiley discovered a novel method to increase the cross-range resolution of radar imaging systems, which he referred as a Doppler beam sharpening. What he discovered

was that a comparison of the returns from a series of radar pulses from an antenna mounted on a platform moving along a given flight path parallel to the ground could be used to produce a higher resolution image [5]. In this way, a synthetic aperture with length equal to that of the platform flight path is created. The method that he discovered was essentially the same as what is now called stripmap-mode SAR.

There are a number of different types of SAR imaging systems that are currently in use. Two of the most common SAR modes, which are used by modern systems, are stripmap-mode and spotlight-mode. However, it should be noted that these modes are not exclusive. That is, there are some SAR systems that can switch between imaging modes [8].

As was previously mentioned, the original SAR mode that was invented by Wiley was stripmap-mode SAR. In the case of this SAR mode, the radar antenna is mounted at a fixed angle on a moving platform. stripmap-mode SAR systems are capable of producing high resolution images over a large region of the ground. This makes it useful for terrain mapping.

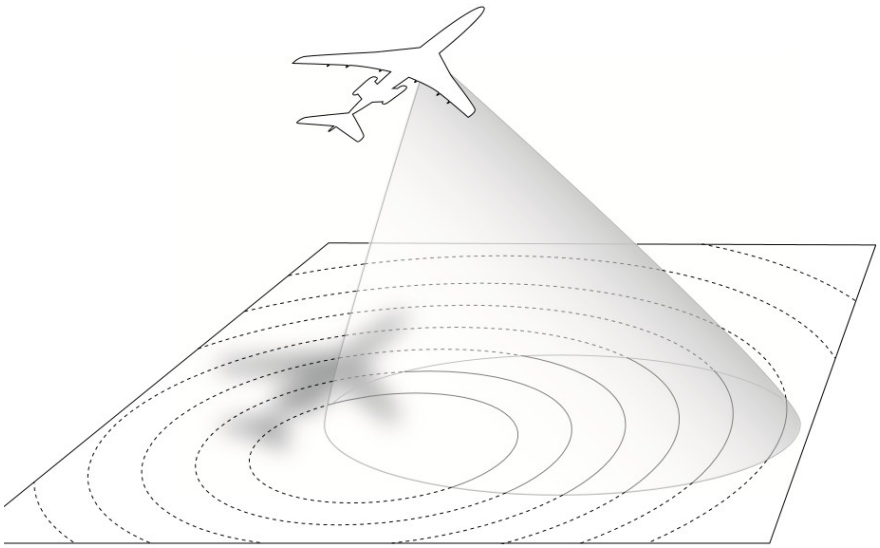

Figure I.2: Illustration of stripmap-mode SAR

In the case of Spotlight-mode SAR, the radar beam is steered so that it remains focused on a single area of the target space. Spotlight-mode SAR systems can use either electronic or mechanical beam steering. The main advantage of this mode is that it has increased image resolution

when compared to stripmap-mode SAR. However, this increase in resolution comes at the cost of decreased area coverage.

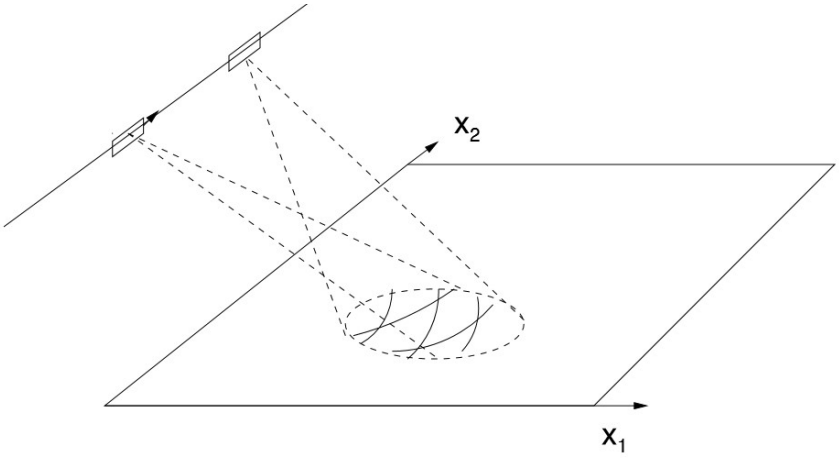

Figure I.3: Illustration of spotlight-mode SAR

The best models retain as much physics as possible, though algorithms based on such models are often computationally intensive and difficult to apply to real-time radar environments or those involving large amounts of data. One such physics-based model utilized the geometrical theory of diffraction. This model gives an estimate of the geometry, location, and response to polarization for each scatterer. Another physics-based model is based on Maxwell's equations for electromagnetism [6]. But it does not use all of Maxwell's equations. An ideal model would use the full set of Maxwell's equations, but such a model only is required systems that have antennas that take polarimetric measurements.

Most models make a linearizing assumption known as the Born approximation, which is equivalent to assuming that the wave scatters only once before returning to the antenna. For example, when imaging underneath dense vegetation, most scattering is multiple scattering and the inability to account for this degrades the quality of the image formed. In addition, this approximation may introduce unwanted effects to the image such as shadowing and it ignores polarization changes as the waves scatter.

In this thesis we will study synthetic aperture radar imaging under the wave equation model introduced by Cheney [6].

CHAPTER II
MATHEMATICAL MODEL FOR SYNTHETIC APERTURE RADAR

II.1 Wave Propagation

The signals transmitted and measured by radar antenna are electromagnetic, therefore the appropriate model for radar is Maxwell's equation [6]. For radar imaging, what is important is the wave nature of the transmitted and measured signals. Therefore, an appropriate model is given by

$$(\nabla^2 - \frac{1}{c^2(x)}\partial_t^2)u(t,x) = -j(t,x), \tag{II.1}$$

where $u(t,x)$ represents one component of the electromagnetic field due to some source $j(t,x)$ and $c(t,x)$ is local propagation speed of electromagnetic waves. Scattering causes singularities in the wave speed. In the absence of scatterers the speed of propagation is $c(t,x) = c_0$.

Scattering can be thought of as being due to perturbations in the wave speed, which we write as

$$\frac{1}{c^2(x)} = \frac{1}{c_0^2} - V(x). \tag{II.2}$$

Here $V(x)$ is the reflectivity function. Eq. (II.1) and Eq. (II.2) do not provide an entirely accurate model for electromagnetic scattering from an object. Nevertheless, this is a commonly used model for radar scattering, with the understanding that $V(x)$ does not exactly correspond to the perturbation in the electromagnetic wave speed in the material.

In this model, the electric field can be divided into two component fields

$$u = u^{in} + u^{sc}. \tag{II.3}$$

In the above equation, $u^{in}(t,x)$ is the incident field that is emitted by the antennas, and u^{sc} is the scattered field, which results from the interaction of the incident field with a target. Since the

source of the incident field is a current on the antenna, u^{in} is modeled using the non- homogeneous wave equation

$$(\nabla^2 - \frac{1}{c_0^2(x)}\partial_t^2)u^{in}(t,x) = -j(t,x). \tag{II.4}$$

A wave equation that describes the propagation of the scattered field is derived from Eq (II.1) and Eq (II.4), and it reads

$$(\nabla^2 - \frac{1}{c_0^2(x)}\partial_t^2)u^{sc}(t,x) = -V(x)\partial_t^2 u(t,x). \tag{II.5}$$

Here $V(x)$ is the reflectivity function, which is given by

$$V(x) = \frac{1}{c_0^2} - \frac{1}{c^2(x)}. \tag{II.6}$$

When the incident field comes into contact with a target, it induces a current, which causes the target to re-emit a weaker time shifted version of the same signal. However, $V(x)$ does not directly measure the intensity of the reflected signal. Instead, it indicates the level perturbation that occurs in the wave speed when the incident field comes in contact with the target plane. It will be assumed that the reflectivity function has compact support on the set of points on the target plane that have been illuminated by the antenna.

A fundamental solution of the wave equation is a generalized function satisfying

$$(\nabla^2 - \frac{1}{c_0^2(x)}\partial_t^2)g(t,x) = -\delta(t)\delta(x). \tag{II.7}$$

The solution of Eq. (II.7) is

$$g(t,x) = \frac{\delta(t - |x|/c_0)}{4\pi|x|}. \tag{II.8}$$

The Green's function enables us to solve the constant-speed wave equation with any source term. The solution for the incident field is

$$u^{in}(t,x) = \int_{R^3} dy \int_R \frac{\delta(t - \tau - |x-y|/c_0)}{4\pi|x-y|} j(\tau, y)d\tau. \tag{II.9}$$

Since the antenna current density will be modeled such that $j(t,x) = p(t)\delta(x - x_0)$. Subsisting this into Eq. (II.9) yields the following expression

$$u^{in} = \frac{p(t - |x - x_0|/c_0)}{4\pi|x - x_0|}. \tag{II.10}$$

Furthermore, it is common to represent the current wave model in the frequency domain. Such a representation can be found by considering the following Helmholtz wave equation

$$(\nabla^2 + k^2)U^{in}(x,f) = -J(x,f). \tag{II.11}$$

In the above equation, U^{in} and J are the Fourier Transform of u^{in} and j, respectively, the equation can be solved by using the Helmholtz Green's function

$$G(x) = \frac{e^{-ik|x|}}{4\pi|x|}, \tag{II.12}$$

where $k = 2\pi f/c_0$ denotes the wavenumber. Then, the solution for the frequency domain representation of the incident field is given in terms of the following equation

$$U^{in}(f,x) = \int_{R^3} \frac{e^{-ik|x-y|}}{4\pi|x-y|} J(x,f)dy, \tag{II.13}$$

since $j(t,x) = p(t)\delta(x - x_0)$, we obtain,

$$U^{in}(f,x) = p(f)\frac{e^{-ik|x-x_0|}}{4\pi|x-x_0|}. \tag{II.14}$$

II.2 The Lippmann- Schwinger Integral Equation

Since the scattered field is created as a result of the interaction of the incident field with the target scene, it would be useful if a formulation for the scattered field could be found directly in terms of the incident field. However, in general, this is not always possible. Consider that the scattered

field can be described by the following equation

$$(\nabla^2 + k^2)U^{sc}(x, f) = -V(x)U(x, f). \tag{II.15}$$

The solution to this equation can be found, in the same way as before, in the frequency domain to be

$$U^{sc}(f, x) = -\int_{R^3} \frac{e^{-ik|x-z|}}{4\pi|x-z|} V(z) f^2 U(x, f) dz, \tag{II.16}$$

and in the time domain to be

$$u^{sc}(t, x) = \int_{R^3} dz \int_R \frac{\delta(t - \tau - |x-z|/c_0)}{4\pi|x-z|}, V(z)\partial_t^2 u(\tau, z) d\tau. \tag{II.17}$$

It is clear that in both of the above expressions that the scattered field is dependent on the total field. Since in Eq. (II.16) and Eq. (II.17) the scattered field appears on both sides of the equation, it is not possible to exactly formulate the scattered field in terms of the incident field alone. In the following sections, a solution to this problem will be detailed

II.3 The Born Approximation

For radar imaging, we measure u^{sc} at the antenna, and we would like to determine V. However, both V and u^{sc} in the neighborhood of the target V are unknown, and in Eq. (II.17) these unknowns are multiplied together. This nonlinearity makes it difficult to solve for V. Consequently, almost all work on radar imaging involves making the Born approximation, which is also known as the weak-scattering or single-scattering approximation. This corresponds to replacing $u(t, x)$ on the right side of Eq. (II.17) by u^{in}. giving the following equation

$$u^{sc}(t, x) = \int_{R^3} dz \int_R \frac{\delta(t - \tau - |x-z|/c_0)}{4\pi|x-z|} V(x)\partial_t^2 u^{in}(\tau, z) d\tau. \tag{II.18}$$

In the frequency domain, the Born approximation is

$$U^{sc}(f, z) = -\int_{R^3} \frac{e^{-ik|x-z|}}{4\pi|x-z|} V(z) f^2 U^{in}(z, f) dz. \tag{II.19}$$

Since we know that

$$U^{in}(f,z) = p(f)\frac{e^{-ik|z-x_0|}}{4\pi|z-x_0|}, \tag{II.20}$$

it immediately follows that

$$U^{sc}(f,x) = -\int_{R^3} \frac{e^{-ik|x-z|}}{4\pi|x-z|} V(z) f^2 p(f) \frac{e^{-ik|z-x_0|}}{4\pi|z-x_0|} dz. \tag{II.21}$$

The Born approximation is very useful, because it makes the imaging problem linear.

II.4 Far-Field Wave Propagation Model

Often the flight path of the radar platform is designed in such a way that the maximum target distance from the origin, which is located at the target scene center, is much smaller than the distance of the antenna from the same origin. In this case, any computation of the received signal based on Eq. (II.21) can be simplified through the application of what is commonly referred to as the far-field approximation. This approximation can be understood by first noting that radar waves propagate as a spherical wave front. When the antenna is far from the target center, the curvature of this wavefront can be assumed to be negligible. Under this assumption, in the extreme far-field, radar wave propagation can be approximately represented by a plane wave. Actually, a good mathematical model is a rectangular distribution of point sources. We denote the length and width of the antenna by L and D, respectively, we denote the center of the antenna by x', thus a point on the antenna can be written $x_0 = x' + q$, where q is a vector from the center of the antenna to a point on the antenna. This assumption can be justified mathematically by applying a first order Taylor expansion to the range term $|z - x_0|$:

$$|z-x_0| = |z-x'-q| = |z-x'| - \widehat{(z-x')} \cdot q + O(\frac{L^2}{|z-x'|}). \tag{II.22}$$

Then, the phase term becomes

$$\frac{e^{-ik|z-x_0|}}{4\pi|z-x_0|} = \frac{e^{-ik|z-x'|}}{4\pi|z-x'|} \cdot e^{ik\widehat{(z-x')}\cdot q}[1 + O(\frac{|L|^2}{|z-x'|})] \cdot [1 + O(\frac{k|q|^2}{|z-x'|})]. \tag{II.23}$$

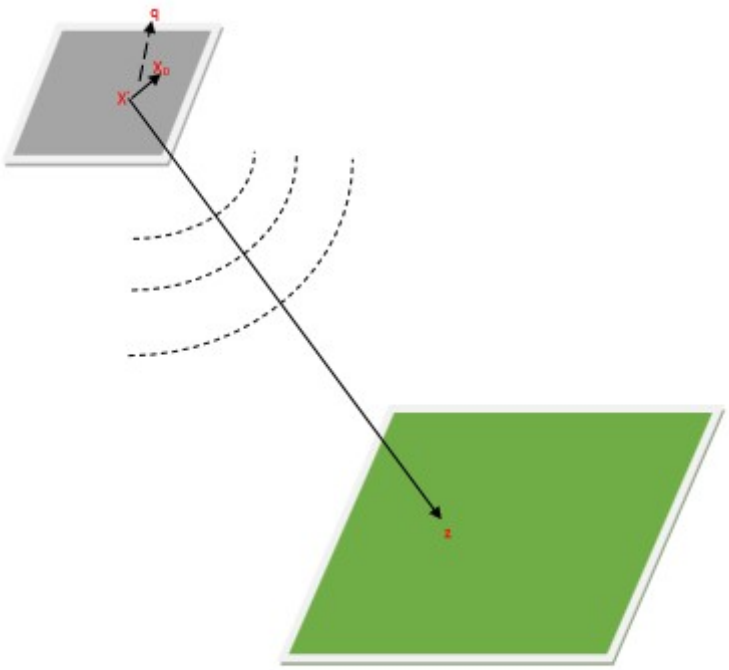

Figure II.1: Illustration of rectangular distribution of point sources

Then, we apply this approximation for the incident field, and we obtain

$$U^{in}(f,z) = p(f)\frac{e^{-ik|z-x'|}}{4\pi|z-x'|}\cdot e^{\widehat{(z-x')}\cdot q}. \tag{II.24}$$

II.5 Measured Data

We use a similar procedure to model how the receiving antenna measures the scattered field. In Eq. (II.19), we substitute Eq. (II.24) for the incident field to find the measured data.

$$U^{sc}(f,z) = -\int_{R^3}\frac{e^{-ik|x-z|}}{4\pi|x-z|}V(z)f^2p(f)\frac{e^{-ik|z-x'|}}{4\pi|z-x'|}\cdot e^{\widehat{(z-x')}\cdot q}dz. \tag{II.25}$$

If we assume that the receiving antenna is at the same location as the transmitting antenna, we find that the scalar Born model for the received signal is

$$s(f,x') = \int e^{-2ik|z-x'|}\cdot A(f,x',z)\cdot V(z)dz, \tag{II.26}$$

where

$$A(f, x', z) = \frac{f^2}{(4\pi|z - x'|)^2} \cdot F(k, \widehat{z}), \tag{II.27}$$

and

$$F(k, \widehat{z}) = p(f)e^{ik\widehat{(z-x')}\cdot q}. \tag{II.28}$$

Synthetic-aperture imaging involves a moving platform, and usually the antenna is pointed toward the earth. We denote by γ the antenna path. For a pulsed system, we assume that pulses are transmitted at times t_n and that the antenna position at time t_n is γ_n. Because the time scale on which the antenna moves is much slower than the time scale on which the electromagnetic waves propagate, we separate the time scales into a slow times, which corresponds to the n of t_n, and a fast time t. Using a continuum model for the slow time makes some of the analysis simpler but also leaves out some important effects that we will consider below. Using the continuum model for slow time, in Eq. (II.26) we replace the antenna position x' by $\gamma(s)$

$$s(f, s) = \int e^{-2ik|\gamma(s)-z|} A(f, s, z)V(z)dz. \tag{II.29}$$

With the additional assumption that the antennas are broadband and an appropriate symbol estimate of A. With this assumption, we can construct an approximate inverse operator, which we denote B. The reconstructed image I is formed by applying the inverse operator B to the data where B is of the form:

$$I(y) = B[s](y) := \int e^{-2ik|\gamma(s)-z_T|} Q(f, s, z)s(f, s)df ds, \tag{II.30}$$

where $z_T = (z, 0)$ and where Q is a filer to be determined below. The time-domain version is

$$I(y) = B[s](y) := \int e^{if(t-\frac{2|\gamma(s)-z_T|}{c})} Q(f, s, z)s(t, s)df dsdt. \tag{II.31}$$

CHAPTER III

COMPRESSIVE SENSING

III.1 Introduction

The traditional approach of reconstructing signals or images from measured data follows the well-known Shannon sampling theorem [10], which states that the sampling rate must be at least twice the highest frequency. Similarly, the fundamental theorem of linear algebra suggests that the number of collected samples of a discrete finite-dimensional signal should be at least as large as its length in order to ensure reconstruction. This principle underlies most devices of current technology, such as analog-to-digital conversion, medical imaging, or audio and video electronics. The novel theory of compressive sensing(CS)-also known under the terminology of compressed sensing, compressive sampling, or sparse recovery-provides a fundamentally new approach to data acquisition, which overcomes this common wisdom. It predicts that certain signals or images can be recovered from what was previously believed to be highly incomplete measurements.

CS relies on the empirical observation that many types of signals or images can be well approximated by a sparse expansion in terms of a suitable basis, that is, by only a small number of nonzero coefficients. This is the key to the efficiency of many lossy compression techniques such as JPEG, MP3, etc. A compression is obtained by simply storing coefficients are simply set to zero. This is certainly a reasonable strategy when full information of the signal is available. However, when the signal first has to be acquired by a somewhat costly, lengthy, or otherwise difficult measurement procedure, this seems to be a waste of resources. First, large efforts are spent in order to obtain full information on the signal, and afterward most of the information is thrown away at the compression stage. One might ask whether there is a clever way of obtaining the compressed version of the signal more directly, by taking only a small number of measurements of the signal. It is not obvious whether this is possible since measuring directly the large coefficients requires to

know a priori their location. Quite surprisingly, compressive sensing provides nevertheless a way of reconstructing a compressed version of the original signal by taking only a small amount of linear and nonadaptive measurements. The precise number of required measurements is comparable to the compressed size of the signal. Clearly, the measurements have to be suitably designed. It is a remarkable fact that all provably good measurement matrices designed so far are random matrices. It is for this reason that the theory of compressive sensing uses a lot of results from probability theory.

The first naive approach to a reconstruction algorithm consists in searching for the sparsest vector that is consistent with the linear measurements. This leads to the combinatorial l_0-problem, which unfortunately is NP-hard in general. There are essentially two approaches for tractable alternative algorithms. The first is convex relaxation leading to l_1-minimization. By now, basic properties of the measurement matrix, which ensure sparse recovery by l_1-minimization are known: the null space property (NSP) and the restricted isometry property (RIP). The latter requires that all column sub matrices of a certain size of the measurement matrix are well conditioned. This is where probabilistic methods come into play because it is quite hard to analyze these properties of deterministic matrices with minimal amount of measurements. Among the provably good measurement matrices are Gaussian, Bernoulli random matrices, and partial random Fourier matrices.

Many lossy compression techniques such as JPEG, JPEG-2000, MPEG, or MP3 rely on the empirical observation that audio signals and digital images have a sparse representation in terms of a suitable basis. Roughly speaking, one compresses the signal by simply keeping only the largest coefficients. In certain scenarios such as audio signal processing one considers the generalized situation which appears in terms of a redundant system- a so-called dictionary or frame- rather than a basis. The problem of finding the sparsest representation in terms of the given dictionary turns out to be significantly harder than in the case of sparsity with respect to a basis where the expansion coefficients are unique.

III.2 Mathematical Modeling and Analysis

This section introduces the concept of sparsity and the recovery of sparse vectors from incomplete linear and nonadaptive measurements. In particular, an analysis of l_1-minimization as a recovery method is provided. The null-space property and the restricted isometry are introduced and it is shown that they ensure robust sparse recovery. It is actually difficult to show these properties for deterministic matrices and the optimal number m of measurements, and the major breakthrough in compressive sensing results is obtained for random matrices. Example of several types of random matrices that ensure sparse recovery are given, such as Gaussian, Bernoulli, and partial random Fourier matrices.

- **Preliminaries and Notation**

This exposition mostly treats complex vectors in $\mathbb{C}^N$ although sometimes the considerations will be restricted to the real-case $\mathbb{R}^N$. The l_p-norm of a vector $x \in \mathbb{C}^N$ is defined as

$$\|x\|_p := \left(\sum_{j=1}^{N} |x_j|^p \right)^{1/p}, 0 < p < +\infty, \tag{III.1}$$

$$\|x\|_\infty := \max_{\{j \in 1, \cdots, N\}} |x_j|. \tag{III.2}$$

For $1 < p \leq \infty$, it is indeed a norm while for $0 < p < 1$ it is only a quasi-norm. When emphasizing the norm the term l_p^N is used instead of $\mathbb{C}^N$ or $\mathbb{R}^N$. The unit ball in l_p^N is $B_p^N = \{x \in \mathbb{C}^N, \|x\|_p \leq 1\}$. The operator norm of a matrix $A \in \mathbb{C}^{m \times N}$ from l_p^N to l_p^m is denoted

$$\|A\|_{p \to p} = \max_{\|x\|_p = 1} \|Ax\|_p. \tag{III.3}$$

In the important special case $p = 2$, the operator norm is the maximal singular value $\sigma_{\max}(A)$ of A.

For a subset $T \subset \{1, \cdots, N\}$. we denote by $x_T \in \mathbb{C}^N$ the vector, which coincides with $x \in \mathbb{C}^N$ on the entries in T and is zero outside T. Similarly, A_T denotes the column sub matrix of A corresponding to the columns indexed by T. Further, $T^c = \{1, \cdots, N\} \backslash T$ denotes the complement

of T and $|T|$ indicates the cardinality of T. The kernel of a matrix A is denoted by $\ker A = \{x \in \mathbb{C}^N, Ax = \mathbf{0}\}$.

- **Sparsity and Compression**

Compressive sensing is based on the empirical observation that many types of real-world signals and images have a sparse expansion in terms of a suitable basis or frame, for instance a wavelet expansion. This means that the expansion has only a small number of significant terms, or in other words, that the coefficient vector can be well approximated with one having only a small number of non vanishing entries.

This support of a vector x is denoted $supp(x) = \{j \in \{1, \cdots, N\} | x_j \neq 0\}$, and

$$\|x\|_0 := |supp(x)|. \tag{III.4}$$

It has become common to call $\|\cdot\|_0$ the l_0-norm, although it is not even a quasi-norm. A vector x is called k-sparse if $\|x\|_0 \leq k$. For $k \in \{1, 2, \cdots, N\}$,

$$\sum_k := \{x \in C^N : \|x\|_0 \leq k\}, \tag{III.5}$$

denotes the set of k-sparse vectors. Furthermore, the best k-term approximation error of a vector $x \in \mathbb{C}^N$ in l_p is defined as

$$\sigma_k(x)_p = inf_{z \in \sum_k} \|x - z\|_p. \tag{III.6}$$

If $\sigma_k(x)$ decays quickly in k then x is called compressible. Indeed, in order to compress x one way may simply store only the k largest entries. When reconstructing x from its compressed version the non-stored entries simply set to zero, and the reconstruction error is $\sigma_k(x)_p$. It is emphasized at this point that the procedure of obtaining the compressed version of x is adaptive and nonlinear since it requires the search of the largest entries of x in absolute value. In particular, the location of the non zeros is a nonlinear type of information.

The best k-term approximation of x can be obtained using the non-increasing rearrangement $r(x) = (|x_{i_1}|, \cdots, |x_{i_N}|)^T$, where i_j denotes a permutation of indices such that $|x_{i_1}| \geq |x_{i_{j+1}}|$ for $j = \{1, \cdots, N-1\}$. Then it is straightforward to check that

$$\sigma_k(x)_p = \left(\sum_{j=k+1}^{N} r_j(x)^p \right)^{1/p}, 0 < p < \infty, \tag{III.7}$$

and the vector $x_{[k]}$ derived from x by setting to zero all the $N-k$ smallest entries in absolute value is the best k-term approximation

$$x_{[k]} = \arg min_{z \in \sum_k} \|x - z\|_p, \tag{III.8}$$

for any $0 < p \leq \infty$.

The next lemma states essentially that l_p-balls with small q are good models for compressible vectors.

Lemma 1 Let $0 < q < p \leq \infty$ and set $r = 1/q - 1/p$, then

$$\sigma_k(x)_p \leq k^{-r}, k = 1, 2, \cdots, N. \tag{III.9}$$

- **Compressive Sensing**

The above outlined adaptive strategy of compressing a signal x by only keeping its largest coefficients is certainly valid when full information on x is available. If, however, the signal first has to be acquired or measured by a somewhat costly or lengthy procedure then this seems to be a waste of resources: At first, large efforts are made to acquire the full signal and then most of the information in thrown away when compressing it.

Taking m linear measurements of a signal $x \in \mathbb{C}^N$ corresponds to applying a matrix $A \in \mathbb{C}^{m \times N}$- the measurement matrix

$$y = Ax. \tag{III.10}$$

The vector $y \in \mathbb{C}^M$ is called the measurement vector. The main interest is in the vastly under sampled case $m \ll N$. Without further information, it is , of course, impossible to recover x from y since the linear system is highly under determined, and has therefore infinitely many solution. However, if the additional assumption that the vector x is k-sparse is imposed, then the situation dramatically changes as will be outlined.

The approach for a recovery procedure that probably comes first to mind is to search for the sparsest vector x, which is consistent with the measurement vector $y = Ax$. This leads to solving the l_0-minimization problem

$$\min \|z\|_0, \text{ subject to } Az = y. \tag{III.11}$$

Unfortunately, this combinatorial minimization problem is NP-hard in general [16]. In other words, an algorithm that solves for any matrix A and any right-hard side y is necessarily computationally intractable. Therefore, essentially two practical and tractable alternative to have been proposed in the literature: convex relaxation leading to l_1-minimization also called basis pursuit– and greedy algorithms, such as various matching pursuits [15]. Quite surprisingly for both types of approaches various recovery results are available, which provide conditions on the matrix A and on the sparsity $\|x\|_0$ such that the recovery solution coincides with the original x, and consequently also with the solution. This is no contradiction to the NP-hardness, since these results apply only to a subclass of matrices A and right-hand sides y.

The l_1-minimization approach considers the solution of

$$\min \|z\|_1, \text{ subject to } Az = y, \tag{III.12}$$

which is a convex optimization problem and can be seen as a convex relaxation. Various efficient convex optimization techniques apply for its solution [14]. In the real-valued case, it is equivalent to a linear program and in the complex-valued case, it is equivalent to a second-order cone program. Therefore, standard software applies for its solution.

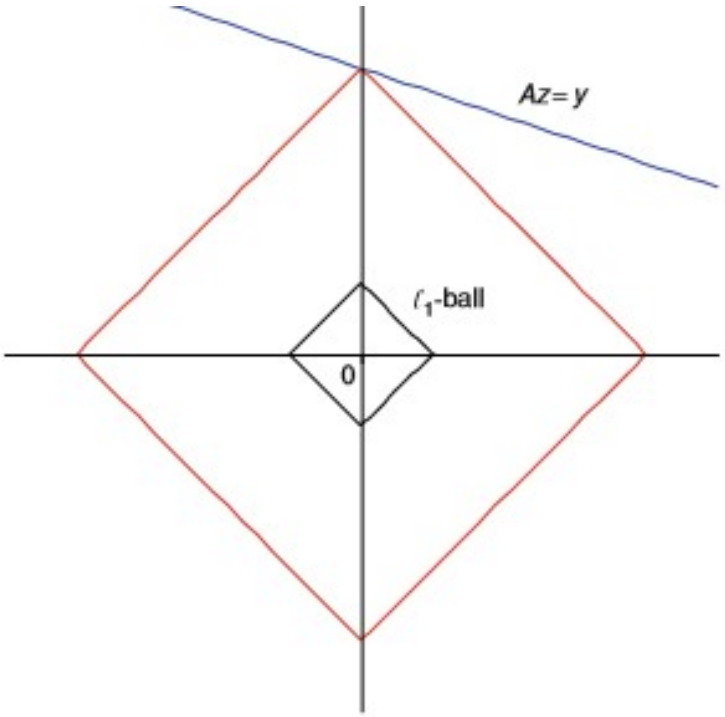

Figure III.1: The minimizer within the affine space of solutions of the linear system Az = y coincides with a sparsest solution

- **The null space property**

The null space property is fundamental in the analysis of l_1-minimization.

Definition 1 A matrix $A \in C^{m\times N}$ is said to satisfy the null space property (NSP) of order k with constant $\gamma \in (0,1)$ if and only if

$$\|\eta\|_1 \leq \gamma \|\eta_{T^c}\|_1, \tag{III.13}$$

for all sets $T \subset \{1, \cdots, N\}$, and for all $\eta \in \ker A$ The following sparse recovery result is based on this notion.

Theorem 1 Let $A \in \mathbb{C}^{m\times N}$ be a matrix that satisfies the NSP of order k with constant $\gamma \in (0,1)$. Let $x \in \mathbb{C}^N$ and $y = Ax$ and let x^* be a solution of the l_1-minimization problem, then

$$\|x - x^*\|_1 \leq \frac{2(1+\gamma)}{1-\gamma}\sigma_k(x)_1. \tag{III.14}$$

In particular, if x is k-sparse then $x = x^*$.

Proof, Let $\eta = x - x^*$, then $\eta \in \ker A$ and $\|x^*\|_1 \leq \|x\|_1$,

because x^* is a solution of the l_1-minimization problem. Let T be the set of the k-largest entries of x in absolute value. One has

$$\|x_T^*\|_1 + \|x_{T^c}^*\|_1 \leq \|x_T\|_1 + \|x_{T^c}\|_1. \tag{III.15}$$

It follows immediately from the triangle inequality that

$$\|x_T\|_1 - \|\eta_T\|_1 + \|\eta_{T^c}\|_1 - \|x_{T^c}\|_1 \leq \|x_T\|_1 + \|x_{T^c}\|_1, \tag{III.16}$$

hence

$$\|\eta_{T^c}\|_1 \leq \|\eta_T\|_1 + 2\|x_{T^c}\|_1 \leq \gamma\|\eta_{T^c}\|_1 + 2\sigma_k(x)_1, \tag{III.17}$$

or, equivalently.

$$\|\eta_{T^c}\|_1 \leq \frac{2}{1-\gamma}\sigma_k(x)_1, \tag{III.18}$$

finally

$$\|x - x^*\|_1 = \|\eta_T\|_1 + \|\eta_{T^c}\|_1 \leq (\gamma+1)\|\eta_{T^c}\|_1 \leq \frac{2(1+\gamma)}{1-\gamma}\sigma_k(x)_1, \tag{III.19}$$

and the proof is completed.

One can also show that if all k-sparse x can be recovered from $y = Ax$ using l_1-minimization then necessarily A satisfies the NSP of order k with some constant $\gamma \in (0, 1)$. Therefore, the NSP is actually equivalent to sparse l_1 recovery.

- **The restricted isometry property**

The NSP is somewhat difficult to show directly. The restricted isometry property (RIP) is easier to handle and it also implies stability under noise as stated below.

Definition 2 The restricted isometry constant δ_k of a matrix $A \in \mathbb{C}^{m\times N}$ is the smallest number such that

$$(1-\delta_k)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1+\delta_k)\|z\|_2^2. \tag{III.20}$$

A matrix A is said to satisfy the restricted isometry property of order k with constant δ_k if $\delta_k \in (0,1)$. It is easily seen that δ_k can be equivalently defined as

$$\delta_k = \max_{T \subset \{1,\cdots,N\}} \|A_T^* A_T - Id\|_2, \ I \text{ is identity}, \tag{III.21}$$

which means that all column submatrices of A with at most k columns are required to be well conditioned. The RIP implies the NSP, as shown in the following lemma.

Lemma 2 Assume that $A \in \mathbb{C}^{m \times N}$ satisfies the RIP of order $K = k + h$ with constant $\delta_K \in (0,1)$. Then A has the NSP of order k with constant $\gamma = \sqrt{\frac{k}{h}\frac{1+\delta_K}{1-\delta_K}}$.

Taking $h = 2k$ above shows that $\delta_{3k} < 1/3$ implies $\gamma < 1$. By theorem 1, recovery of all k-sparse vectors by l_1-minimization is then guaranteed. Additionally, stability in l_1 is also ensured. The next theorem shows that RIP implies also a bound on the reconstruction error in l_2.

Theorem 2 Assume $A \in \mathbb{C}^{m \times N}$ satisfies the RIP of order $3k$ with $\delta_{3k} < 1/3$. For $x \in \mathbb{C}^N$, let $y = Ax$ and x^* be the solution of the l_1-minimization problem. Then

$$\|x - x^*\|_2 \leq C \frac{\delta_k(x)_1}{\sqrt{k}}, \tag{III.22}$$

with $C = \frac{2}{1-\gamma}(\frac{\gamma+1}{\sqrt{2}} + \gamma)$, and $\gamma = \sqrt{\frac{1+\delta_{3k}}{2(1-\delta_{3k})}}$.

III.3 Applying compressed sensing to SAR

The primary interest in compressed sensing research is the inverse problem of recovering a signal $f \in \mathbb{C}^N$ from noisy linear measurements $y = Af + n \in C^N$. The focus is on underdetermined problems where the forward operator $A \in \mathbb{C}^{M \times N}$ has unit norm columns and forms an incomplete basis with $M \ll N$. The resulting ill-posed inverse problem is regularized assuming: (1) the unknown signal f is K-sparse or is compressible with K significant coefficients and (2) the noise process is bounded by $\| \ n \ \|_2 < \epsilon$. CS theory provides strong results which guarantee stable solution of the sparse signal recovery problem for a class of forward operators A that satisfies certain properties. One such class of operators is defined by bounding the singular values of the submatrices of A. Specifically, the restricted isometry constant (RIC) δ_K for forward operator A is

the smallest $\delta \in (0, 1)$, such that [3]

$$(1 - \delta_K) \parallel x \parallel_2^2 \leq \parallel Ax \parallel_2^2 \leq (1 + \delta_K) \parallel x \parallel_2^2, \tag{III.23}$$

hold for all vectors x with at most K nonzero entries.

One of the key contributions of CS is that stable recovery of compressible, noisy signals can be achieved through the solution of the computationally tractable l_1 regularized inverse problem [3]

$$\min_f \parallel f \parallel_1 \text{ subject to } \parallel Af - y \parallel_2^2 \leq \epsilon^2, \tag{III.24}$$

At present, the least conservative available bound on the reconstruction performance guarantees that if $\delta_{2K} < \sqrt{2} - 1$ and $\parallel n \parallel_2 \leq \epsilon$, then the solution $\widehat{f}$ will satisfy [4]

$$\parallel f^* - \widehat{f} \parallel_2^2 \leq C_0 K^{-1/2} \parallel f^* - f_K \parallel_1 + C_1 \epsilon, \tag{III.25}$$

where f_K is the best K-sparse approximation to the true solution f^*, C_0 and C_1 are small constants, and $\parallel \cdot \parallel_p$ represents the l_p norm. The optimization can be viewed as the convex relaxation of the NP-hard task of finding the sparsest feasible solution

$$\min_f \parallel f \parallel_0 \text{ subject to } \parallel Af - y \parallel_2^2 \leq \epsilon, \tag{III.26}$$

where $\parallel \cdot \parallel_0$ is the l_0 norm, i.e., the number of nonzero entries in the vector.In radar and other array processing applications, imperfect calibration implies that precise knowledge of A is not available. Recent work has shown that a bounded unknown additive disturbance to the matrix A still permits a RIC-based guarantee on reconstruction performance that reduces to the result as the disturbance bound approaches zero.

Consider an unknown matrix $A \in \mathbb{C}^{M \times N}$ and an orthonormal basis $(A_i)_i$ for $\mathbb{C}^{M \times N}$. Then there

exist coefficients $(s_i)_i$, such that

$$A = \sum_{i=0}^{MN-1} s_i A_i. \tag{III.27}$$

Our goal is to identify the coefficients $(s_i)_i$. Since the basis elements are fixed, identifying $(s_i)_i$ is tantamount to discovering A. We will do this by designing a test function $f = (f_0, \cdots, f_{N-1})^T \in \mathbb{C}^N$ and observing $Af \in \mathbb{C}^M$. Here, $(\cdot)^T$ denotes the transpose of a vector or a matrix. For instance, A may represent an unknown communication channel which needs to be identified for equalization purposes.

For simplicity, from now on assume that $N = M$. The observation vector can be reformulated as

$$y = \sum_{i=0}^{N^2-1} s_i A_i f = \sum_{i=0}^{N^2-1} s_i \varphi_i = \Phi s, \tag{III.28}$$

where the i-th atom $\varphi_i = A_i f$ is a column vector of length N, the concatenation of the atoms $\Phi = [\varphi_0 | \varphi_1 | \cdots | \varphi_{N^2-1}]$ is an $N \times N^2$ matrix, and $s = (s_0, s_1, \cdots, s_{N^2-1})^T$ is a column vector of length N^2. The system of equation in (III.29) is clearly highly underdetermined. If s is sufficiently sparse, then there is hope of recovering s from y.

CHAPTER IV

SAR IMAGING USING COMPRESSED SENSING

An important prerequisite of CS sparse reconstruction is that the signal must be sparse or compressible in certain representations. For SAR imaging, the true 3-D illuminated scene is projected into the 2-D range-azimuth plane, and then the projected targets are not always sparse. However, the target space can be regarded as sparse in some special applications in which only a small number of strong scatterers distribute in the illuminated scene, and the relatively few large coefficients of the scatterers can capture most of the information on scene, such as ocean ships monitoring, aircraft and spacecraft detecting, space debris imaging, and so on. Based on the features of sparse signal in these applications, the image can be reconstructed by the signal of strong scattering centers using the theory of CS, and the weak scattering centers can be regarded as noise in image.

IV.1 Random Partial Fourier Matrices

While Gaussian and Bernoulli matrices provide optimal conditions for the minimal number of required samples for sparse recovery, they are of somewhat limited use for practical applications for several reasons. Often the application imposes physical or other constraints on the measurement matrix, so that assuming B to be Gaussian may not be justifiable in practice. One usually has only limited freedom to inject randomness in the measurements. Furthermore, Gaussian or Bernoulli matrices are not structured so there is no fast matrix vector multiplication available, which may speed up recovery algorithms. Thus, Gaussian random matrices are not applicable in large-scale problems.

A very important class of structured random matrices that overcomes these drawbacks are random partial Fourier matrices, which were also the object of study in the very first papers on compressive sensing. A random partial Fourier matrix $A \in \mathbb{C}^{m\times N}$ is derived from the discrete fourier matrix

$F \in \mathbb{C}^{n\times n}$ with entries

$$F_{j,k} = \frac{1}{\sqrt{N}} e^{2\pi jk/N}, \tag{IV.1}$$

by selecting m rows uniformly at random among all N rows. Taking measurements of a sparse $x \in \mathbb{C}^N$ corresponds then to observing m of the entries of its discrete Fourier Transform $\widehat{x} = Fx$. It is important to note that the Fast Fourier Transform may be used to compute matrix vector multiplication with B and B^* with complexity $O(Nlog(N))$. The following theorem concerning the RIP constant was proven and improves slightly on the results.

Theorem 3 Let $B \in \mathbb{C}^{m\times N}$ be the random partial fourier matrix as just described. Then the restricted isometry constant of the rescaled matrix $\sqrt{\frac{NB}{m}}$ satisfy $\delta_k \leq \delta$ with probability at least $1 - N^{-\gamma log^3(N)}$ provided

$$m \geq C\delta^{-2} klog^4(N), \tag{IV.2}$$

the constants $C, \gamma > 1$ are universal.

Combining the estimate with the l_1-minimization results above shows that recovery with high probability can be ensured for all k-sparse x provided

$$m \geq Cklog^4(N). \tag{IV.3}$$

IV.2 Discrete Point Targets

For a set of stationary points the target reflectivity function is usually expressed as a sum of Dirac delta functions.

$$V(z) = \sum_j \sigma_j \delta(z - z_j). \tag{IV.4}$$

In the case Eq. (II.26) becomes

$$s(f, x') = \sum_j \int e^{-2ik|z-x'|} A(f, x', z)\sigma_j \delta(z - z_j)dz, \tag{IV.5}$$

therefore

$$s(f,x') = \sum_j \sigma_j e^{-2ik|z_j - x'|} A(f,x',z_j), \tag{IV.6}$$

where

$$A(f,x',z) = \frac{f^2}{(4\pi|z-x'|)^2} \cdot F(k,\widehat{z}), \tag{IV.7}$$

and

$$F(k,\widehat{z}) = p(f) e^{ik\widehat{(z-x')}\cdot q}. \tag{IV.8}$$

To simulate return data we let $P(f) = 1$ for $f \in (f_{min}, f_{max})$, such that the bandwidth of the system is $\beta = f_{max} - f_{min}$.

IV.3 SAR Model And Analysis

In the spirit of CS, a very small number of 'random' measurements carry enough information which can accomplish complete reconstruction for the signal. According to the feature of RIP, all submatrices of Φ are composed of k significant columns which should be nearly orthogonal. There are some well-known pairs of incoherent basis, such as randomly selected Fourier samples and random Gaussian matrix. Hence, we randomly select $M(O(Klog(N/K)) \leq M < N)$ rows of matrix B as the final measurement matrix $\Phi \in \mathbb{R}^{M\times N}$ and then the new measured signal can be expressed as

$$S = \Psi\Phi\alpha = B\alpha, \tag{IV.9}$$

where Ψ denotes a $M \times N$ matrix constructed by randomly selected M rows of the $N \times N$ identity matrix which is taken on orthogonal basis, the randomly selected matrix can be written as

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}_{M\times N}$$

In order to use CS, a linear measurement model of SAR should be created firstly. According to

Chapter II, in discrete scenarios the raw echo signal of SAR can be expressed as

$$s(f, x') = \sum_j \sigma_j e^{-2ik|z_j - x'|} A(f, x', z_j), \tag{IV.10}$$

where

$$A(f, x', z) = \frac{f^2}{(4\pi|z - x'|)^2} \cdot F(k, \widehat{z}), \tag{IV.11}$$

and

$$F(k, \widehat{z}) = p(f) e^{ik\widehat{(z-x')} \cdot q}. \tag{IV.12}$$

Thus, according to the Random Partial Fourier Matrices, the signal can be rewritten as

$$s(f, x') = F_{j,k}\alpha, \;\; j, k = 1, 2, \cdots M, \tag{IV.13}$$

where $F_{j,k} = \frac{1}{\sqrt{M}} exp(2\pi ik\varphi_j/M)$, $\varphi_j = |z_j - x'|$ and α is the other part.

IV.4 Numerical Simulations

Figure (4.1) shows an example of an image formed using the matched filter algorithm. Phase history data was simulated for point targets using Eq. $(IV.14)$

$$s(f, x') = \sum_j \sigma_j e^{-2ik|z_j - x'|} A(f, x', z_j), \tag{IV.14}$$

where $A(f, x', z) = 1$. Here, $Np = 128$ pulses were simulated with $K = 512$ frequency samples per pulse, a center frequency of $10GHz$ and a $600MHz$ bandwidth. A circular fight path was used with a 30 degree depression angle and a slant range of 10km. A 3 degree integration angle was used with a center azimuth angle of 50 degrees. The scene extent was 10 m x 10 m with 2 cm pixel spacing in each dimension.

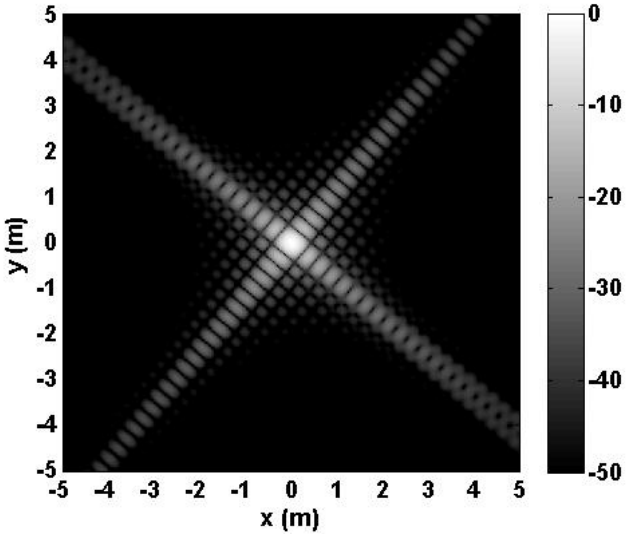

Figure IV.1: Matched filter image of one point target.

Basis pursuit will be used to reconstruct this one point, the result is following.

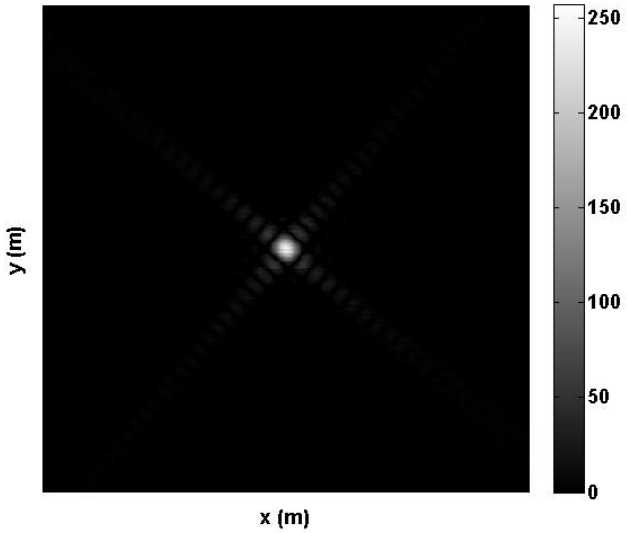

Figure IV.2: Basis pursuit image of one point,which use 1/2 of phase data.

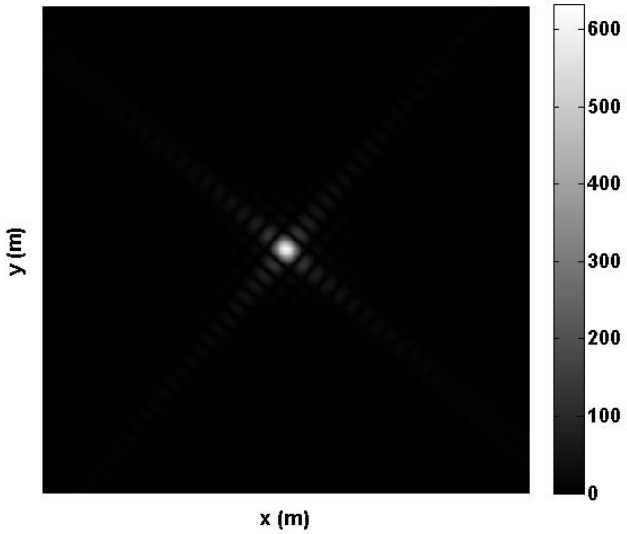

Figure IV.3: Basis pursuit image of one point, which use log^4 of phase data.

Then we want to try more points, Thus we use the rocket as a model.

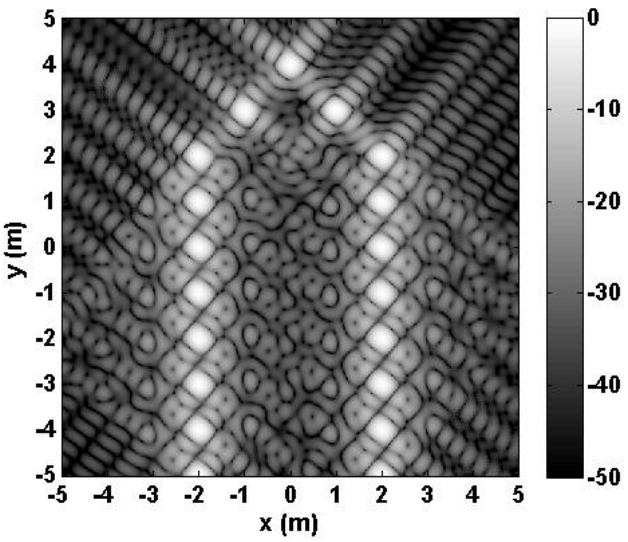

Figure IV.4: Matched filter image of a rocket

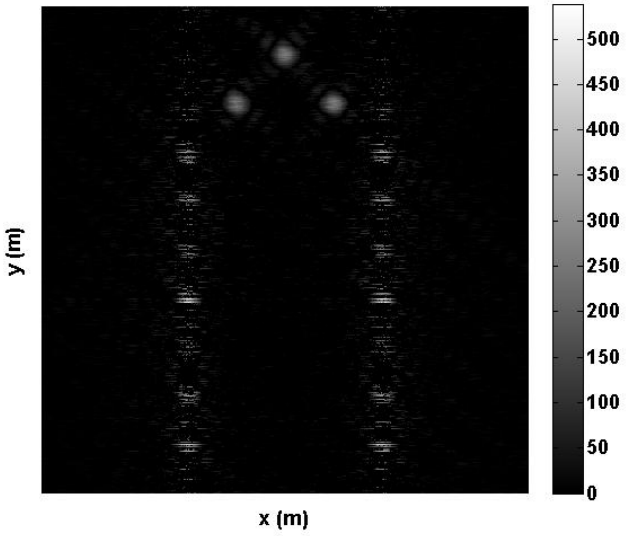

Figure IV.5: Basis pursuit image of a rocket, which use 1/2 of phase data.

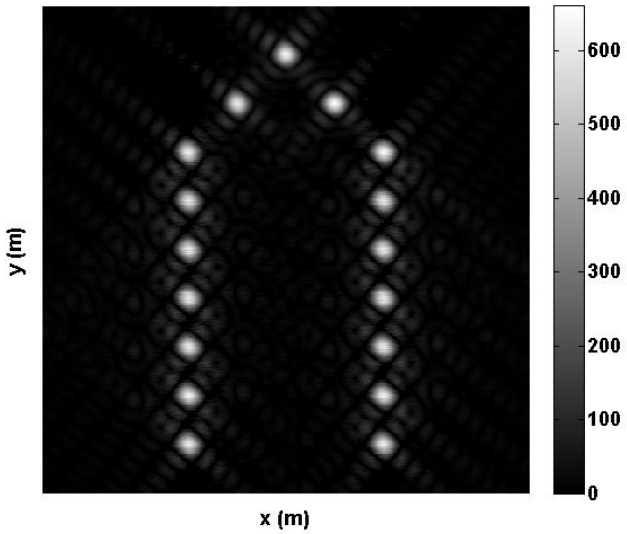

Figure IV.6: Basis pursuit image of a rocket, which use log^4 of phase data.

CHAPTER V

CONCLUSION

In this paper, we derived the SAR signal model from the scalar form of the Maxwell's equations. We gave a short introduction to compressed sensing. Then we used compressed sensing to the SAR signal model. The results have a strong theoretical foundation being derived from an inverse scattering problem in Maxwell's equations. It was showed that compressed sensing will be used to construct point targets, as efficiently as matched filter algorithms.

BIBLIOGRAPHY

[1] Cand‘es, E.J.,*The restricted isometry property and its implications for compressed sensing*.Compte Rendus de l’Academie des Sciences, Paris, Serie I, 346 589-592, 2008

[2] Borden, B., *Radar Imaging of Airborne Targets*. Philadelphia: Institute of Physics Publishing, 1999

[3] Cand‘es, E.J, Romberg, J, and Tao, T., *Robust undertainty principles: Exact signal reconstruction from highly incomplete frequency information*. IEEE Trans. Inform. Theory 52, 489-509, 2012

[4] Donoho, D. L.,*Compressed sensing* IEEE Trans. Inform. Theory,vol.52,no.4,1289-1306, 2006

[5] Chassande-Mottin, E., Flandrin, P., "On the Stationary Phase Approximation of Chirp Spectra", Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 117-120, 1998

[6] Cheney, M., Borden, B., *Fundamentals of Radar Imaging*. Philadelphia: SIAM, 2009

[7] Cheney, W., *Analysis for Applied Mathematics*. New York: Springer-Verlag, 2001

[8] Cumming, I.G., Wong, F.H., *Digital Processing of Synthetic Aperture Radar Data*. Norwood, MA: Artech House, 2005

[9] Curlander, J.C., McDonough, R.N., *Synthetic Aperture Radar: Systems and Signal Processing*. New York: Wiley & Sons, 1991

[10] Vogel, C,. *Computational Methods for inverse problem Frontier In Applied Mathematics Series*.SPIAM, New York, 2002

[11] Franceschetti, G., Lanari, R., *Synthetic Aperture Radar Processing*. FL: CRC Press,Boca Ratton, 2000

[12] Garza, G., "Mathematics of Synthetic Aperture Radar Imaging", M.S. thesis, Dept. Math., Univ. Texas-Pan American, Edinburg, TX, 2011

[13] Garza, G., Qiao, Z., "Resolution analysis of bistatic SAR", Proceedings of SPIE Vol. 8021, 802169, 2011

[14] Tropp,JA., Needell,D., "Iterative Signal Recoverry From Incomplete And Inaccurate Samples ". Appl Comput Harmon Anal 26. 301-321, 2008

[15] Boyd, S., Vandenberghe, I., *Convex Optimizational*, Cambridge University press. 2004

[16] Mallat, SG., Zhang, Z., "Matching Pursuits With Time-Frequency Dictionaries",IEEE Trans Signal Process.41(12) 3397-3415, 1993

[17] Lopez, J.X., Garza, G., Qiao, Z., "Cross-range imaging of SAR and PDE analysis", Proceedings of SPIE Vol. 7698, 76981C-1-C15, 2010

[18] Lopez, J.X., Garza, G. Qiao, Z., "Cross-Range Imaging of SAR Data", Pacific Journal of Applied Mathematics, 2(3), 65-81, 2009

[19] Natterer, F., *The Mathematics of Computerized Tomography*. Philadelphia: SIAM, 2001

[20] Natterer, F., Wubbeling, F., *Mathematical Methods in Image Reconstruction*. Philadelphia: SIAM, 2001

[21] Nicholson, K.J., Wang, C.H., "Improved Near-Field Radar Cross-Section Measurement Technique", IEEE Antennas and Wireless Propagation Letters, Vol. 8, 1103-1106, 2009

[22] Ray,T.P., Lopez, J.X., Qiao, Z., "Principles of 3D Turntable Radar Imaging", 2012 IEEE Radar Conference, 758-763, 2012

[23] Soumekh,M. *Radar Handbook*. New York: McGraw-Hill Companies, 2008

[24] Soumekh,M. *Synthetic Aperture Radar Signal Processing with MATLAB Algorithms*. New York: John Wiley & Sons, Inc., 1999

BIOGRAPHICAL SKETCH

Yufeng Cao was born in Datong, Shanxi in 1982. In 2007, he graduated with honors with a B.Sc. in Mathematics from the Shanxi normal University, Later he continued his education by earning his M.Sc. in Mathematical Sciences in 2010. From 2011-2013, he get another important M.Sc in the University of Texas-Pan America. He can be contacted at ycao1@broncs.utpa.edu.