Jacobs University Bremen

Department of Mathematics

Bachelor Thesis

Discrete Compressive Sensing:
Null Space Properties and Convex Optimization Methods

Submitted by: Tianlin Liu

Matriculation number: 20331165

Date of submission: May 27, 2016

Field of study: Mathematics

Supervisor: Prof. Dr.Götz Pfander, Dr. Dae Gwan Lee

Acknowledgements

I would like to express my gratitude to my supervisors Prof. Dr. Götz Pfander and Dr. Dae Gwan Lee for their invaluable guidance.

I am grateful to my family, for everything.

I would like to thank my friends for being always close to me.

Last but not least, my sincere thanks goes to my host-father, Mr. Jens Pieper, who provides me with unfailing support and continuous encouragement throughout my years in Bremen, Germany.

Abstract

Compressed Sensing (CS), also known as *Compressed Sensing*, or *Compressive Sampling*, is a technique of reconstructing signals from reduced number of measurements, under the assumption of sparsity, or compressibility, of the signals. It is known that under certain conditions, CS can reconstruct sparse signals at a rate significantly lower than the Shannon-Nyquist-Rate. This thesis aims to study CS under the constraint of discreteness of signals. Our contribution in this thesis is two-fold: first, we developed the Discrete Null Space property, which guarantees the unique reconstruction of discrete signals; second we propose the optimization methods of BFW-SAV(Binary Fixed-Weights Sum of Norms) and BRSN(Binary Reweighed Sum of Norms), which are two novel convex optimization methods that provide accurate binary signal reconstruction.

Contents

Chapter 1

Introduction

The topic of *Compressive Sensing* (CS) was firstly introduced by David Donoho, Emmanuel Candes, Justin Romberg and Terence Tao in 2006. They showed that under certain conditions, sparse signals, i.e., signals with a limited amount of non-zero entries, can be recovered accurately and uniquely from a small set of data in the measurement system [1, 2, 3, 4, 5]. The technique of CS is significant and far-reaching as it is a scheme of reconstructing sparse signals with only a few measurements. What is more, the CS techniques allow precise reconstruction of sparse signals with a sub *Shannon-Nyquist Sampling rate*.

Much research in recent years has focused on using CS techniques to reconstruct discrete signals. This paradigm of purposefully uniting CS and discrete signal reconstruction is called Discrete Compressive Sensing. The development of Discrete CS is motivated by grayscale image processing [6], symbol detection [7], control engineering [8], multi-user detection and communications [9, 10], to name a few.

This thesis aims to investigate discrete CS in two perspectives: (1) the unique guarantee of discrete signal reconstruction, and (2) convex optimization methods for discrete signals. The rest of the chapters are organized as follows: in the second chapter, a short survey on General CS is given; in the third chapter, the condition of unique reconstruction guarantee is developed under the Discrete CS scheme, and the optimization methods for discrete signals are presented; in the fourth chapter, numerical experiments are provided; and in the last chapter, the thesis is summarized and some open questions are raised.

Notation

Throughout out this thesis, Blackboard bold $\mathbb{F}$ is used to represent the underlying field $\mathbb{R}$ or $\mathbb{C}$.

Vectors and matrices are in bold. Vectors $\mathbf{v} \in \mathbb{F}^n$ are column vectors, denoted by

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

The 1-vector of dimension N is defined by

$$\mathbf{1}_N = (\underbrace{1, 1, \ldots, 1}_{N})^T.$$

The sign function of a real number x is defined as follows:

$$\operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}.$$

The sign function of a vector $\mathbf{v} = (v_1, \ldots, v_n)^T \in \mathbb{R}^n$ is defined as follows:

$$\operatorname{sgn}(\mathbf{v}) = (\operatorname{sgn}(v_1), \ldots, \operatorname{sgn}(v_n)).$$

When $1 \leq p < \infty$, the *p-norm* of vector $\mathbf{v} \in \mathbb{F}^n$ is defined by:

$$||\,\mathbf{v}\,||_p = \left[\sum_{i=1}^{n} |v_i|^p \right]^{\frac{1}{p}}.$$

When $0 < p < 1$, the *p-quasi-norm* is defined in the same fashion.

When $p = 0$, the l_0 "norm" of $\mathbf{v} \in \mathbb{R}^n$ is defined by

$$||\,\mathbf{v}\,||_0 = ||(v_1, v_2, \ldots, v_n)^T||_0 = \{ \text{ the number of } v_i \; |1 \leq i \leq n, v_i \neq 0\}.$$

Please note that in spite of the name, the l_0 norm is not really a norm. The name follows from the convention.

The *support* of a vector $\mathbf{v}$, denoted by $\mathrm{supp}(\mathbf{v})$, is defined by

$$\mathrm{supp}(\mathbf{v}) = \{i | \, \mathbf{v}(i) \neq 0\}.$$

Matrices $\mathbf{A} \in \mathbb{F}^{m \times n}$ are denoted by

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

The Kronecker product of two matrices $\mathbf{A} \in \mathbb{R}^{M \times N}$ and $\mathbf{B} \in \mathbb{R}^{K \times L}$ is denoted by $\mathbf{A} \otimes \mathbf{B}$, which is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{1,1}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1,N}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,1}\mathbf{B} & a_{M,2}\mathbf{B} & \cdots & a_{M,N}\mathbf{B} \end{pmatrix} \in \mathbb{R}^{MK \times NL}.$$

Chapter 2

Compressive Sensing: an Overview

2.1 Underdetermined linear system and K-sparse model

The mean goal of Compressive Sensing is to find sparse solutions to the underdetermined linear system, which is written as $\mathbf{y} = \Phi\,\mathbf{x}$:

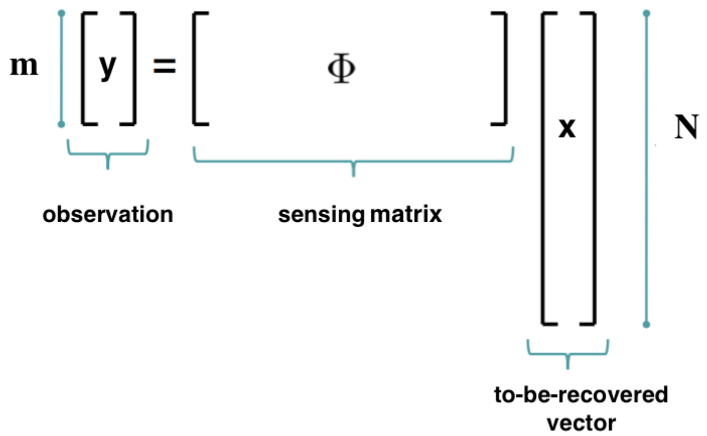

Figure 2.1: Underdetermined linear system

where $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^N$, $\Phi \in \mathbb{R}^{m\times N}$, $m \ll N$. Note that, in the CS Scheme, we assume that Φ the sensing matrix "reduces the dimension" in the sense that it maps the high dimensional vector to the low dimensional one. From basic linear algebra, we know that the underdetermined linear system has infinitely many solutions.

We now consider the *K-sparse Model*, by imposing the sparsity condition on the vector $\mathbf{x}$. By saying "$\mathbf{x}$ is K-sparse," we mean that $\mathbf{x}$ has at most K non-zero entries.

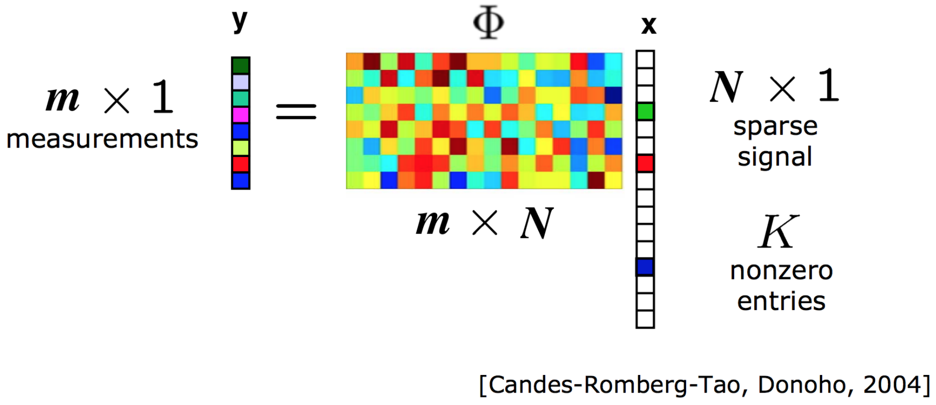

Figure 2.2: K-Sparse Model

The K-sparse model is meaningful for two reasons: first, it preserves structure and information in the linear system; second, under certain conditions that we will specify later, the model is invertible with high probability, and in this case we are able to tackle the previously ill-posed inverse problem.

We call a triple $(\mathbf{y}, \mathbf{A})$ a *problem instance* [11]. Given a problem instance, Compressive Sensing is a technique to reconstruct the underlying $\mathbf{x}$ in the K-sparse model.

2.2 Reconstruction of Sparse Signal

There are several classical techniques to reconstruct the underlying solution vector $\mathbf{x}$. In this section, we introduce two well-known optimization methods, l_0 and l_1 minimization. We conclude that although l_0 embodies our commitment to the K-sparse model, it is more practical to work with l_1 minimization.

2.2.1 l_0 minimization

Given the problem instance $(\mathbf{y}, \mathbf{A})$, we wish to penalize the number of non-zero entries of the to-be-recovered vector $\mathbf{x}$. This leads to the l_0 minimization:

$$\text{minimize} \, || \, \mathbf{z} \, ||_0 \text{ subject to } \Phi \, \mathbf{z} = \mathbf{y} \qquad (P_0)$$

However, directly solving the l_0 minimization is NP-hard. The proof is provided by [12]. To see the validity of the claim intuitively, suppose $N = 1000$, $K = 50$. As a result, there will be

$$\binom{1000}{50} \approx 10^{95}$$

possible supporting sets. Even if we check 10^{12} subsets per second, the computing time will last for 10^{65} years.

2.2.2 l_1 minimization

To get rid of NP-hardness, we replace l_0 norm by l_1 norm. The method is called the l_1 minimization, or the *Basis Pursuit*, which was firstly introduced in [13].

$$\text{minimize} \, || \, \mathbf{z} \, ||_1 \text{ subject to } \Phi \, \mathbf{z} = \mathbf{y} \tag{P_1}$$

The Basis Pursuit has two advantages: it is not NP-hard, and therefore it is practical to solve; what is more, it is a convex optimization problem and therefore we have handy tools to study it.

Naturally, one wants to ask the question that, under which conditions the solutions provided by (P_1), the practical case, coincide with those of (P_0), the ideal case. Work reported in [2, 3, 4] showed that the (P_0) and (P_1) lead to the same result, given that the vector is sparse enough. The work also provides easily-satisfied bounds on the required sparsity that guarantee such equivalence.

2.2.3 Greedy Methods

Based on Basis Pursuit, numerous efficient algorithms are developed, for example, MP(Matching Pursuit) [14], OMP(Orthogonal Matching Pursuit) [15, 16], IH(Iterative Threshold) [17], etc. These greedy algorithms are known for their low computational complexity, though they have the drawbacks of low reconstruction accuracy.

Among all greedy algorithms, one of the most heavily studied and widely used algorithms is the OMP. As shown in the box below, OMP contains two steps in each iteration: the index-selecting step (OMP_1) and the projection step (OMP_2). The two-step iteration indicates its greedy nature: the index is chosen to minimize the residual in $|| \cdot ||_2$ locally at each iteration.

Orthogonal Matching Pursuit(OMP)

Input: measurement matrix $\mathbf{A}$, measurement vector $\mathbf{y}$
Initialization: $S^0 = \emptyset$, $\mathbf{x}^0 = 0$
Iteration: repeat until a stopping criterion is met at $n = \overline{n}$

$$S^{n+1} = S^n \cup \{j_{n+1}\}, \; j_{n+1} := \underset{j \in [N]}{\operatorname{argmax}} \{ |(\mathbf{A}^*(\mathbf{y} - \mathbf{A}\,\mathbf{x}^n))_j| \} \tag{OMP$_1$}$$

$$\mathbf{x}^{n+1} = \underset{\mathbf{z} \in \mathbb{C}^N}{\operatorname{argmin}} \{ ||\,\mathbf{y} - \mathbf{A}\,\mathbf{z}\,||_2, \operatorname{supp}(\mathbf{z}) \subset S^{n+1} \} \tag{OMP$_2$}$$

Output: the $\overline{n}$ sparse vector $\mathbf{x}^\sharp = \mathbf{x}^{\overline{n}}$

2.3 Null Space Property

We know that Basis Pursuit can accurately recover signals given that the signals are sparse enough, but there is no reason to believe the reconstructed vector is unique. In this section, we introduce the *Null Space Property*, which guarantees the unique reconstruction of the $k-$sparse vector $\mathbf{x}$ via Basis Pursuit.

Definition 2.3.1 (Definition 4.1 in [18])**.** Let $\Phi \in \mathbb{F}^{m \times N}(\mathbb{F} = \mathbb{C} \text{ or } \mathbb{R})$ and $1 \leq k \leq N$. Then Φ is said to satisfy the Null Space Property (NSP) relative to a set $K \subset [N]$ if

$$\|\mathbf{v}_K\|_1 < \|\mathbf{v}_{K^c}\|_1 \quad \text{for any } \mathbf{v} \in \ker\,\Phi \backslash \{0\}.$$

where

$$\mathbf{v}_K(i) := \begin{cases} \mathbf{v}(i) & \text{if } i \in K, \\ 0, & \text{otherwise.} \end{cases}$$

It is said to satisfy the NSP of order k if it satisfies the NSP relative to any set $K \subset [N]$ with $\operatorname{Card}(K) \leq k$,

We now indicate the link between the Null Space Property and exact recovery of sparse vector via Basis Pursuit.

Theorem 2.3.2 (Theorem 4.4 in [18])**.** Let $\Phi \in \mathbb{F}^{m \times N}(\mathbb{F} = \mathbb{C}$ *or* $\mathbb{R})$, $K \subset [N]$, every vector $\mathbf{x} \in \mathbb{F}$ supported on a set K is the unique solution of (P_1) with $\mathbf{y} = \mathbf{A}\mathbf{x}$ if and only if Φ satisfies the null space property relative to K.

Theorem 2.3.3 (Theorem 4.5 in [18])**.** Let $\Phi \in \mathbb{F}^{m\times N}(\mathbb{F} = \mathbb{C}$ or $\mathbb{R})$, $K \subset [N]$, every k-sparse vector $\mathbf{x} \in \mathbb{F}$ is the unique solution of (P_1) with $\mathbf{y} = \mathbf{A}\mathbf{x}$ if and only if Φ satisfies the null space property of order k.

The above two theorems specify the conditions that we required for unique and exact recovery under the Basis Pursuit approach.

2.4 Summary

Compressive Sensing aims to reconstruct sparse vectors from the underdetermined linear system. There are several reconstruction approaches, such as l_0 minimization, l_1 minimization, and the greedy algorithms. l_0 minimization is proved to be NP-hard and therefore not practical to conduct, though it is of theoretical importance; l_1 minimization, or Basis Pursuit, viewed as a practical substitute of l_1, is heavily studied by researchers; based on the analysis of l_1, efficient greedy algorithms such as MP, OMP, and IHT, are developed. The Null Space Property specifies a condition that guarantees the unique and exact recovery of sparse vectors in the Basis Pursuit approach.

Chapter 3

Discrete Compressive Sensing

In the previous chapter, we introduced the techniques of CS, which enable sparse signals reconstruction at a rate significantly lower than the Shannon-Nyquist-Rate. The price we pay for applying the technique is the "sparseness" of signals. In this chapter, we introduce paradigm of *Discrete Compressive Sensing*, which exploits not only the sparseness of the signals but also the "discreteness" of the signals.

3.1 Problem Setup

To formulate the discrete Compressive Sensing paradigm, we introduce the *Discrete K-Sparse Model* by modifying the K-sparse model under the discrete constraint. The modification is shown intuitively in Figure 3.1.

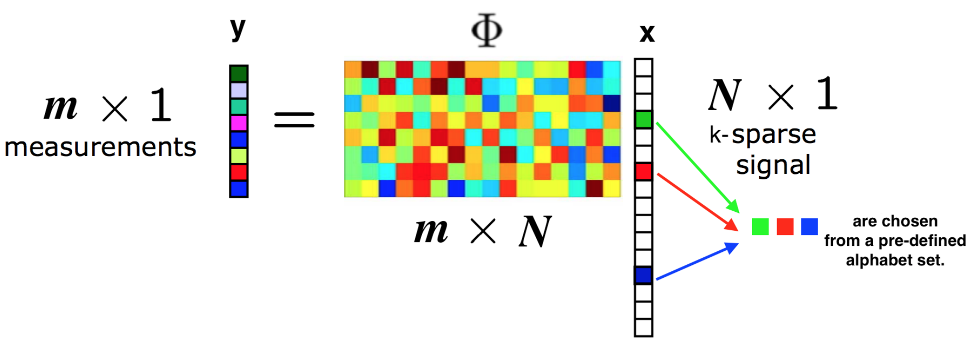

Figure 3.1: Discrete K-sparse Model

As before, $\mathbf{A}$ is a $m \times N$ matrix, $y \in \mathbb{R}^m$, and $\mathbf{x}$ is K-sparse. Further, we assume that $\mathbf{x}$ takes values from a finite set or alphabet. That is to say, we assume that we have the prior knowledge that entries of $\mathbf{x}$ are selected from some known finite set, or alphabet, α.

3.2 Discrete Null Space Property

Our aim is to find the conditions of unique reconstruction guarantee of Discrete K-sparse model. Our previous experience in the general K-sparse model suggests that we shall polish the Null Space Property in the discrete setting. Indeed, since the Discrete K-sparse model is a sub-case of General K-sparse model, we expect a sharper condition for the unique and exact reconstruction. To this end, we relax the condition of Null Space Property and define the *Discrete Null Space Property* as the following:

Definition 3.2.1. A matrix $\mathbf{A} \in \mathbb{R}^{m\times N}$ is said to satisfy the discrete null space property relative to the alphabet $\alpha \subset \{0, \pm 1, \pm 2, \cdots \pm L\}$ and the set $K \subset [N]$ if

$$|\langle \mathbf{v}_K, \operatorname{sgn}(\mathbf{x})_K\rangle| < || \mathbf{v}_{\overline{K}} ||_1. \tag{3.2.1}$$

$\mathbf{v} \in \mathbf{ker}\,\mathbf{A} \setminus \{0\}$ and for $\mathbf{x} \in \alpha^N \cap \Sigma_K$.

It is said to satisfy the discrete null space property of order k if it satisfies the discrete null space property relative to any set $K \subset [N]$.

Remark 3.2.2. To check the discrete null space property of order k, it suffices to check the largest k entries in absolute value. Notice that for $\alpha = \{0, \pm 1, \cdots, \pm L\}^N \cap \Sigma_K$ where L is a positive integer and $K \subset [N]$,

$$\max_{\mathbf{x}_0 \in \mathcal{C}} |\langle \mathbf{v}_K, \operatorname{sgn}(\mathbf{x}_0)_K\rangle| = \sum_{i\in K} |\mathbf{v}(i)| = \|\mathbf{v}_K\|_1.$$

Remark 3.2.3. Suppose the alphabet $\alpha = \{0, 1, \cdots, L\}$, $x \in \{0, 1, \cdots, L\}^N \cap \Sigma_K$ where $K \subset [N]$, then the condition 3.2.1 is written as

$$\left|\sum_{i\in K} \mathbf{v}(i)\right| < \| \mathbf{v}_{\overline{K}} \|_1 \quad \text{for any } \mathbf{v} \in \ker A \setminus \{0\}. \tag{3.2.2}$$

Now we indicate the link between Discrete Null Space property and the unique and exact recovery for the linear system in the Discrete K-sparse model.

Theorem 3.2.4. Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times N}$, every vector $\mathbf{x} \in \alpha^N \cap \Sigma_K$ is the unique solution of P_1 with $\mathbf{y} = \mathbf{A}\,\mathbf{x}$ if and only if $\mathbf{A}$ satisfies the discrete null space property with respect to α.

Proof.

Essentially we mimic the proof of Null Space Property of individual vectors (Theorem 4.26 & Theorem 4.30 in [18]).

Suppose $\mathbf{A}$ satisfies the the discrete null space property relative to the alphabet $\alpha \subset \{0, \pm 1, \pm 2, \cdots \pm L\}$ and the set $K \subset [N]$, i.e., $\forall\, \mathbf{v} \in \mathbf{ker}\,\mathbf{A} \setminus \{0\}$ and $\mathbf{z} \in \alpha^N \cap \Sigma_K$, we have

$$|\,\langle \mathbf{v}_K, \operatorname{sgn}(\mathbf{z})_K\rangle\,| < ||\,\mathbf{v}_{\overline{K}}\,||_1.$$

We want to show $\mathbf{z} \in \alpha^N \cap \Sigma_K$ is the unique solution of P_1.

For $\mathbf{x} \neq \mathbf{z}$ such that $\mathbf{A}\,\mathbf{x} = \mathbf{A}\,\mathbf{z}$, write $\mathbf{v} = \mathbf{z} - \mathbf{x} \in \ker \mathbf{A} \setminus \{0\}$

$$\begin{aligned}
||\,\mathbf{x}\,||_1 &= ||\,\mathbf{z} - \mathbf{v}\,||_1 \\
&= ||(\mathbf{z} - \mathbf{v})_K||_1 + ||(\mathbf{z} - \mathbf{v})_{\overline{K}}||_1 \\
&= \langle (\mathbf{z} - \mathbf{v})_K, \operatorname{sgn}(\mathbf{z} - \mathbf{v})_K\rangle + ||\,\mathbf{v}_{\overline{K}}\,||_1 \\
&\geq |\langle (\mathbf{z} - \mathbf{v})_K, \operatorname{sgn}(\mathbf{z})_K\rangle| + ||\,\mathbf{v}_{\overline{K}}\,||_1 \\
&> |\langle (\mathbf{z} - \mathbf{v})_K, \operatorname{sgn}(\mathbf{z})_K\rangle| + |\,\langle \mathbf{v}_K, \operatorname{sgn}(\mathbf{z})_K)\rangle\,| \\
&\geq |\,\langle (\mathbf{z})_K, \operatorname{sgn}(\mathbf{z})_K\rangle\,| \\
&= ||\,\mathbf{z}\,||_1.
\end{aligned}$$

Hence $\mathbf{z}$ is the unique minimizer.

Conversely, suppose $\mathbf{z}$ is the unique solution of P_1 and let $\mathbf{v} \in \mathbf{ker}\,\mathbf{A} \setminus \{0\}$. For any $t > 0$, $\mathbf{x}_t \coloneqq \mathbf{z} - t\,\mathbf{v}$ satisfies $\mathbf{A}\,\mathbf{x} = \mathbf{A}\,\mathbf{z}$. Then

$$\begin{aligned}
||\,\mathbf{z}\,||_1 &< ||\,\mathbf{x}_t\,||_1 \\
&= ||(\mathbf{x}_t)_K||_1 + ||(\mathbf{x}_t)_{\overline{K}}||_1 \\
&= ||(\mathbf{x}_t)_K||_1 + ||(\mathbf{z} - t\,\mathbf{v})_{\overline{K}}||_1 \\
&= ||(\mathbf{x}_t)_K||_1 + t||\,\mathbf{v}_{\overline{K}}\,||_1.
\end{aligned}$$

Minus $||(\mathbf{x}_t)_K||_1$ on each side, we get:

$$\begin{aligned} t||\,\mathbf{v}_{\overline{K}}\,||_1 &> ||\,\mathbf{z}\,||_1 - ||(\mathbf{x}_t)_K||_1 \\ &= \langle(\mathbf{z})_K, \operatorname{sgn}(\mathbf{z})_K\rangle - \langle(\mathbf{z} - t\,\mathbf{v})_K, \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K\rangle \\ &= \langle(\mathbf{z})_K, \operatorname{sgn}(\mathbf{z})_K - \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K\rangle + t\langle\mathbf{v}_K, \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K\rangle. \end{aligned}$$

Divide t on each side,

$$||\,\mathbf{v}_{\overline{K}}\,||_1 > \langle(\mathbf{z})_K, \frac{1}{t}(\operatorname{sgn}(\mathbf{z})_K - \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K)\rangle + \langle\mathbf{v}_K, \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K\rangle. \tag{3.2.3}$$

Note that, $\forall k \in K$, $\operatorname{sgn}(\mathbf{v}_k) = 1$ or -1, and

$$\lim_{t\to 0^+} \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K = \operatorname{sgn}(\mathbf{z})_K.$$

Let $t \to 0^+$ in (2.2),

$$\begin{aligned} ||\,\mathbf{v}_{\overline{K}}\,||_1 &= \lim_{t\to 0^+} ||\,\mathbf{v}_{\overline{K}}\,||_1 \\ &> \lim_{t\to 0^+} \langle(\mathbf{z})_K, \frac{1}{t}(\operatorname{sgn}(\mathbf{z})_K - \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K)\rangle + \lim_{t\to 0^+} \langle\mathbf{v}_K, \operatorname{sgn}(\mathbf{z} - t\,\mathbf{v})_K\rangle \\ &= \langle\mathbf{v}_K, \operatorname{sgn}(\mathbf{z})_K\rangle. \end{aligned}$$

Replace $\mathbf{v}$ to $-\,\mathbf{v}$ if necessary, we have

$$||\,\mathbf{v}_{\overline{K}}\,||_1 > -\langle\mathbf{v}_K, \operatorname{sgn}(\mathbf{z})_K\rangle,$$

and we are done. □

3.3 Optimization methods in Discrete CS

Recall that in the Discrete CS paradigm, we aim to reconstruct the discrete sparse signal $\mathbf{x}$ from the underdetermined linear system

$$\mathbf{y} = \mathbf{A}\,\mathbf{x}, \tag{3.3.1}$$

where $\mathbf{x} \in \{\alpha_1, \ldots, \alpha_q\}^N,\ \operatorname{supp}(\mathbf{x}) = K,\ \Phi \in \mathbb{R}^{m\times N}, \mathbf{y} \in \mathbb{R}^m, m \ll N.$

Recent research in Discrete CS has suggested that some novel optimization methods have better reconstruction performances compared with the renowned l_1 minimization, which is widely used in the General CS paradigm. In this section, we give a short survey on those optimization methods tailored for Discrete CS, and then we propose two new optimization methods.

3.3.1 Boxed Basis Pursuit

In [11] the authors proposed the Feas, or Boxed BP optimization to reconstruct *k-simple* vectors, where *k-simple* means that the vector contains only k entries that are not 0 or 1. If we fix the binary alphabet elements as $\alpha = \{0, 1\}$, the binary signal under the Discrete CS scheme coincides with the 0-simple vector.

Boxed-BP

$$\text{minimize} \,||\, \mathbf{z} \,||_1$$

$$\text{subject to } \Phi\, \mathbf{z} = \mathbf{y}, \text{ where } \mathbf{A}\, \mathbf{z} = \mathbf{y}, \; 0 \leq z_i \leq 1$$

3.3.2 l_∞ minimization

The l_∞ minimization, also known as max-norm minimization or Democratic Representation, is introduced to reconstruct the discrete signal with entries $+1$ and -1 in [19].

l_∞ minimization

$$\text{minimize} \,||\, \mathbf{z} \,||_\infty \text{ subject to } \Phi\, \mathbf{z} = \mathbf{y} \tag{l_0}$$

Note that

- It is possible to apply the same optimization method to the other alphabet sets with cardinality of two, i.e., $\alpha = \{\alpha_1, \alpha_2\}$, by simple a translation.
- The solutions of l_∞ minimization are in fact of the boundary points of hypercube $[-1, 1]^N$. This implies that it is not possible to extend the optimization method to the case $\alpha \coloneqq \{\alpha_1, \ldots, \alpha_q\}$ with $q > 2$.

3.3.3 Regularized l_1 minimization (Rl_1)

The Regularized l_1 minimization is proposed in [20]. It is a sparse-seeking optimization method, aiming to force the signal entries that are not in the alphabet set to be 0.

Regularized l_1 minimization(Rl_1)

$$\text{minimize} \sum_{i=1}^{q} ||\mathbf{z} - \alpha_i \mathbf{1}_N^T||_1 \text{ subject to } \Phi\,\mathbf{z} = \mathbf{y}$$

$\alpha = \{\alpha_1, \ldots, \alpha_q\}$ is the alphabet set

The Regularized l_1 minimization has the following advantages:

- Theorem 1 in [20] provides the theoretical guarantee that if we fix $p = 2$, Rl_1 minimization yields the unique solution to 3.3.1 with high probability. What is more, if $q = 2$, Rl_1 and l_∞ yield the same result.
- According to in the analysis provided by section IV in [20], Rl_1 is slightly less computationally expensive than l_∞.
- Unlike l_∞ optimization method, which provides satisfactory reconstruction result only if $q = 2$, we can apply Rl_1 to reconstruct signals with the alphabet $\alpha = \{\alpha_1, \ldots, \alpha_q\}$, $q > 2$.

3.3.4 Sum of Absolute Values minimization (SAV)

The following optimization method is introduced in [21]. The method is proposed to reconstruct the discrete signals under finite alphabet $\alpha = \{\alpha_1, \ldots, \alpha_q\}$ that does not necessarily contain 0. Moreover, the author assumes that the elements of a signal are chosen from a finite alphabet with a known probability distribution, which is given by

$$p_i = \mathbf{P}(x_j = \alpha_i),\ \forall j = 1,\ \ldots, N, i = 1, \ldots, q,$$

where

$$p_i > 0,\ p_1 + \cdots + p_q = 1$$

Sum of Absolute Values minimization(SAV)

$$\text{minimize} \sum_{i=1}^{q} p_i || \mathbf{z} - \alpha_i \mathbf{1}_N ||_1 \text{ subject to } \Phi \mathbf{z} = \mathbf{y}$$

$$\text{where } p_i = \mathbf{P}(z_j = \alpha_i), \forall j \in [N]$$

Example 3.3.1. Suppose $\alpha = \{0, 1\}$, we can re-write our SAV penalty function as:

$$\text{minimize } p|| \mathbf{z} ||_1 + (1-p)|| \mathbf{z} - \mathbf{1}_N ||_1 \text{ subject to } \Phi \mathbf{z} = \mathbf{y},$$

$$\text{where } p = \mathbf{P}(z_j = 0), \forall j \in N.$$

Note that

- SAV is a refined version of Rl_1. They are different only on the weights of summons of l_1 norms. The former chooses the weights according to the probability distribution of the alphabet set, while the later fixes the the uniform 1 as weights.
- We expect the reconstruction performance of SAV as symmetric about $p = 0.5$. That is to say, for a fixed p_0, SAV's reconstruction performance of the signals of sparsity p_0 is supposed to be the same as that of $(1 - p_0)$.

3.3.5 Sum of Norms minimization (SN)

The SN optimization is proposed in [22] to reconstruct the binary sparse signal with only 0 and 1 as entries. That is to say, in the Discrete CS scheme, we fix $\alpha = \{\alpha_1, \alpha_2\} = \{0, 1\}$.

The authors in [22] state that SN is the first optimization method that brings sparse representation [1, 5, 25] and democratic representation [19, 26, 27] together to solve Binary compressive Sensing problem. They show that l_∞ and l_1 can be combined to utilize the underlying binary sparsity. This is achieved by adding l_1 norm and l_∞ norm together, up to a scaling factor λ and a shifting factor vector c.

Sum of Norms minimization(SN)

$$\text{minimize} \, ||\mathbf{z}||_1 + \lambda ||\mathbf{z} - c\,\mathbf{1}_N||_\infty \text{ subject to } \Phi\,\mathbf{z} = \mathbf{y} \qquad (l_{1,\infty})$$

Numerical experiments in [22] are performed to choose the well-performed parameter, $c = \frac{1}{2}$ and $\lambda = 100$.

3.3.6 Binary Fixed-Weights Sum of Absolute Values minimization (BFW-SAV)

The first optimization we propose, the BFW-SAV, is a refined version of SAV in reconstructing binary signal. The motivation of defining this optimization method will be given in 4.2.1.

Binary Fixed Weights Sum of Absolute Value minimization(BFW-SAV)

$$\text{minimize } \rho||\mathbf{z}||_1 + (1-\rho)||\mathbf{z} - \mathbf{1}_N||_1 \text{ subject to } \Phi\,\mathbf{z} = \mathbf{y}$$

$$\text{where } \rho = \text{sgn}(p - \frac{1}{2}) \cdot \epsilon + \frac{1}{2}, \epsilon > 0,\ p = \mathbf{P}(z_j = 0), \forall j \in [N]$$

3.3.7 Binary Reweighted Sum of Norms (BRSN)

The second optimization method we propose, BRSN, combines SAV and SN to solve Binary Compressive Sensing problems:

Binary Reweighted Sum of Norms (BRSN)

$$\text{minimize } (1-p)||\mathbf{z}||_1 + p||\mathbf{z} - \mathbf{1}_N^T||_1 + \lambda\,||\mathbf{z} - \frac{1}{2}\,\mathbf{1}_N||_\infty$$

$$\text{subject to } \Phi\,\mathbf{z} = \mathbf{y},\ \text{where } p = \mathbf{P}(z_j = 0), j \in [N].$$

BRSN is proposed due to the following considerations:

- Motivated by SN, we combine l_1 norm and l_∞ norm to exploit the underlying binary sparsity.
- Motivated to SAV, we use the known probability distribution over alphabet to determine the weights of summons of l_1 minimization. Notice that in the binary CS problem, $p = \mathbf{P}(z_j = 0)$ corresponds to the sparsity level of the vector. In the case of p is close to 1, the to-be-recovered vector is assumed to be sparse; in the case of p close to 0, the to-be-recovered vector is sparse is assumed to be dense. In our K-sparse model, if we assume that the to-be-recovered vector is identically distributed, then $p = 1 - \frac{K}{N}$.

3.4 Summary

The paradigm of Discrete CS aims to reconstruct discrete signals from incomplete linear measurements. Based on the priors of sparsity as well as discreteness of signals, we developed the Discrete Null Space Property, which is a unique-reconstruction-guarantee condition that is sharper compared with the usual Null Space Property introduced in Chapter 2.

Recently, several optimization methods are proposed to solve Discrete CS problems. We expect these optimization methods, tailored for different alphabet sets, to perform better than the classical l_1 minimization. The methods are summarized as below:

Optimization Method	Abbreviation	Alphabet set that applies
Boxed Basis Pursuit	Boxed-BP	$\alpha = \{0, 1\}$
Infinity Norm	l_∞	$\alpha = \{\alpha_1, \alpha_2\}$ $\forall \alpha_i \in \mathbb{R}, i \in 1, 2$
Reweighted l_1	Rl_1	$\alpha = \{\alpha_1, \dots, \alpha_q\}$ $\forall \alpha_i \in \mathbb{R},\ i = 1, \dots, q,\ q \in \mathbb{N}$
Sum of Absolute Values	SAV	$\alpha = \{\alpha_1, \dots, \alpha_q\}$ $\forall \alpha_i \in \mathbb{R},\ i = 1, \dots, q,\ q \in \mathbb{N}$
Sum of Norms	SN	$\alpha = \{0, 1\}$
Binary Fixed Weights Sum of Absolute Values	BFW-SAV	$\alpha = \{0, 1\}$
Binary Reweighed Sum of Norms	BRSN	$\alpha = \{0, 1\}$

Chapter 4

Numerical Experiments on Binary CS

In this section, we present the numerical experiments that evaluate the different optimization methods we introduced in the last chapter under the Binary Compressive Sensing Scheme. First, we conduct experiments on SAV optimization and polish the results published in [21], and consequently we will see BFW-SAV is a product of such a polishment; second, we compare our proposed BFW-SAV and BRSN with other methods.

4.1 Experiment Environment

Recall that in Binary CS scheme, our goal is to find binary sparse solutions to the underdetermined systems $\mathbf{y} = \mathbf{A}\,\mathbf{x}$. Throughout this chapter, we fix $\mathbf{A}$ as 100×200 Gaussian matrices. We use Gaussian matrices because they satisfy Null Space Property with high probability(section 9.4 of [18]), and therefore the underdetermined linear system omits unique solution with high probability; we use the 100×200 dimension because [21] used the same dimensions, and we use them for the convenience in comparison.

In each iteration, we generate the k-sparse binary $\mathbf{x}$. Then we produce $\mathbf{y}$ by computing $\mathbf{A}\,\mathbf{x}$. Next, we perform optimizations to reconstruct $\mathbf{x}$ on the problem instance $(\mathbf{y}, \mathbf{A})$. As a result, we get a (not necessarily correct) solution vector $\mathbf{z}$. Comparing with the genuine solution $\mathbf{x}$, we shall find how satisfactory our reconstruction method is.

There are several ways to define reconstruction accuracy. The most common way in the literature of general CS is the averaged SNR(Signal-to-noise ratio), where the SNR for each iteration i is defined as

$$\mathrm{SNR}(i) = \frac{||\,\mathbf{x} - \mathbf{z}^i\,||_2}{||\,\mathbf{x}\,||_2} \tag{4.1.1}$$

where $\mathbf{z}^i$ is the reconstructed signal in the i-th experiment unit.

Given M iterations, the (averaged) SNR over M units is defined as

$$\text{(averaged) SNR} = \frac{1}{M}\sum_{i=1}^{M} \mathrm{SNR}(i) \tag{4.1.2}$$

However, the averaged SNR is not really the appropriate reconstruction-accuracy measurement if we have the prior knowledge that the signal is discrete. This is because:

- In most cases, the elements of the alphabet set are chosen to be integers, for instance, $\alpha = \{0, 1\}$. Hence, the round-to-integer operations are usually applied to the output in the reconstruction scheme. After the rounding, SNR loses its potential to measure the closeness between reconstructed vector and genuine solution.
- In the discrete CS scheme, we are interested in the cases that optimization methods perfectly reconstruct the signals in each and every entry. SNR fails to provide this "exact-reconstruction" information.

With the above considerations, instead of using averaged SNR, we use the averaged RFP(Reconstruction Failure Probability), which is defined as the following

$$\mathrm{RFP}(i) = \begin{cases} 1 & \exists j,\ 1 \leq j \leq N, \text{such that } z_j^i \neq x_j^i \\ 0 & \text{else} \end{cases}$$

where N is the dimension of the signal $\mathbf{x}$ and $\mathbf{z}$.

Given M experiments, the (averaged) RFP over M is defined as

$$\text{(averaged) RFP} = \frac{1}{M}\sum_{i=1}^{M} \mathrm{RFP}(i) \tag{4.1.3}$$

The experiments throughout this chapter are aided by **CVX** toolbox [23, 24].

4.2 Numerical Experiments on SAV

Recall the SAV minimization from section 3.3.4:

Sum of Absolute Value minimization(SAV)

$$\text{minimize} \sum p_i || \mathbf{z} - \alpha_i \mathbf{1}_N ||_1 \text{ subject to } \Phi \mathbf{z} = \mathbf{y}, \text{ where } p_i = \mathbf{P}(z_j = \alpha_i), \forall j \in [N]$$

Restricting ourself in the Binary CS scheme, we fix the alphabet set as $\alpha = \{\alpha_1, \alpha_2\} := \{0, 1\}$. Then the penalty function can be re-written as:

$$\text{minimize } p|| \mathbf{z} ||_1 + (1-p)|| \mathbf{z} - \mathbf{1}_N||_1 \text{ subject to } \Phi \mathbf{z} = \mathbf{y} \text{ where } p = \mathbf{P}(z_i = 0) \quad (4.2.1)$$

4.2.1 How to Choose Weights for SAV

We ask the following question: to reconstruct binary signals accurately based on SAV, do we really have to know the probability distribution over the alphabet set $\alpha = \{0, 1\}$?

As a comparison to the weights p and $(1-p)$ in the optimization in 4.2.1, we do the following optimization,

$$\text{minimize } w|| \mathbf{z} ||_1 + (1-w)|| \mathbf{z} - \mathbf{1}_N||_1 \text{ subject to } \Phi \mathbf{z} = \mathbf{y} \quad (4.2.2)$$

where w is a fixed parameter which is no longer related to the probability distribution over the alphabet.

We want to check that if it is indeed the case that the RFP has the lowest curve when p and w coincides. To realize this, we perform a two loop programming: in the outer loop, we increase $p = \mathbf{P}(z_i = 0)$ from 0 to 1 and generate the genuine signal $\mathbf{x}$ according to p; in the inner loop, for each of p, we increase w from 0 to 1 and perform the optimization 4.2.2. The experiment result is shown in Figure 4.1 below.

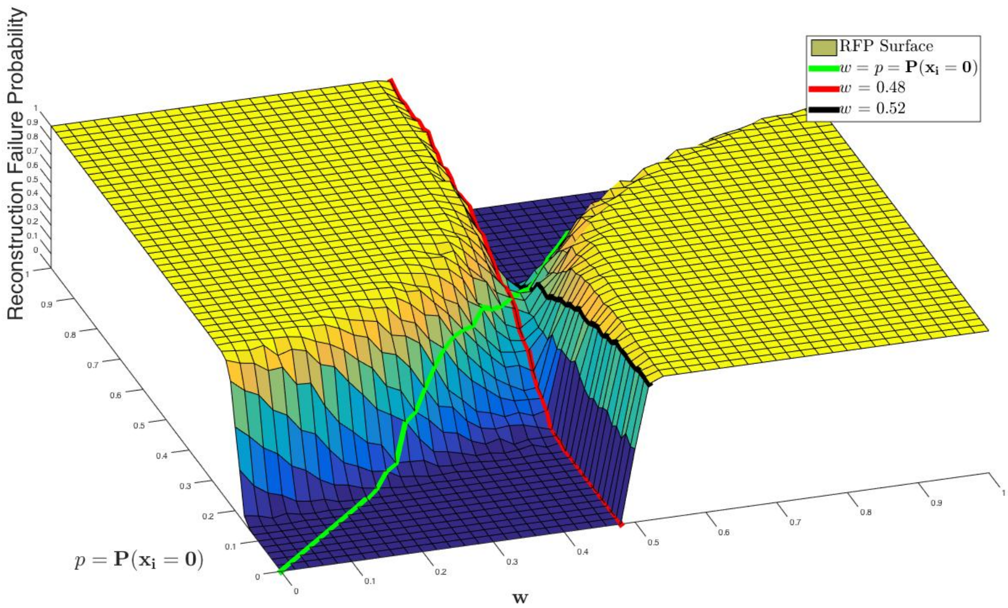

Figure 4.1: RFP with respect to $p = \mathbf{P}(x_i = 0)$ and w

There is one surface and three curves in Figure 4.1:

- The surface: the RFP surface of w associated with p, where w is the weight in 4.2.2 and $p = \mathbf{P}(z_i = 0)$.
- Green curve: the RFP in the case of $w = p = \mathbf{P}(z_j = 0)$ in 4.2.2, i.e., the SAV.
- Red curve: the RFP in the case of $w = \frac{1}{2} - \epsilon$ in 4.2.2.
- Black curve: the RFP in the case of $w = \frac{1}{2} + \epsilon$ in 4.2.2.

In the following figure, we compare the three curves in two dimensional space.

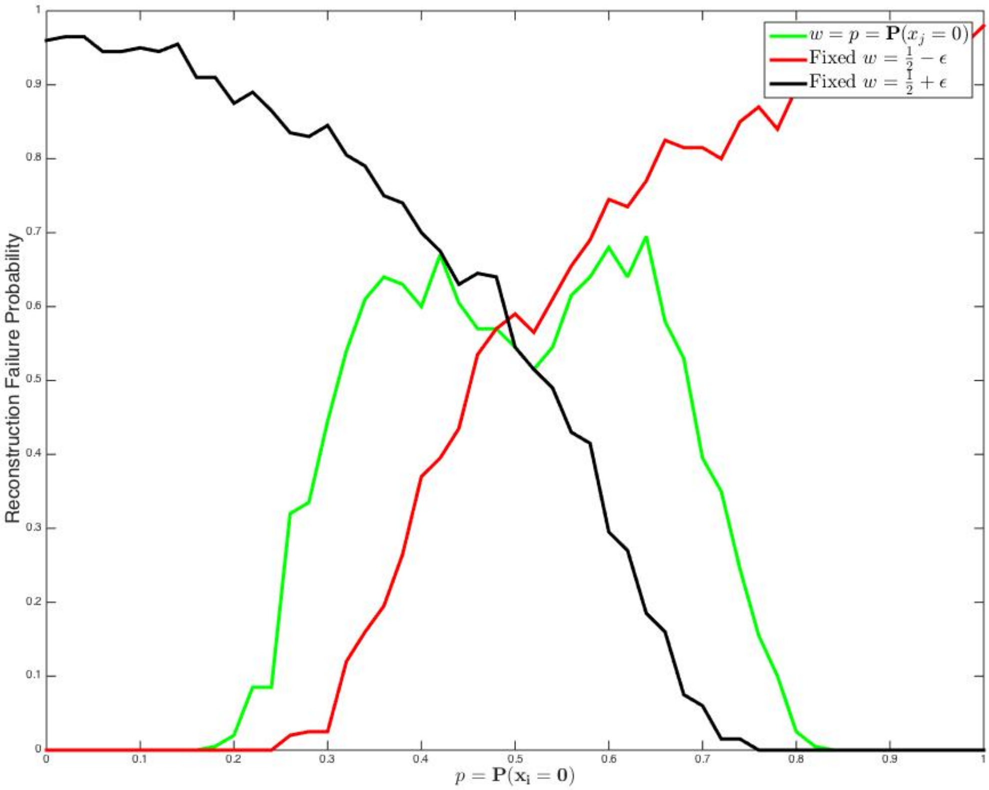

Figure 4.2: Comparing the choice of w

In Figure 4.2, we observe that the red and black curves have the lower RFP on the left and right side of $p = 0.5$ when compared with the green curve. The figure shows that, we can achieve better reconstruction performance, if we fix $w = (\frac{1}{2} - \epsilon)$, $\epsilon > 0$ in the case of dense signal, and $w = (\frac{1}{2} + \epsilon)$, $\epsilon > 0$, in the case of sparse signal. Dr. Dae Gwan Lee formally proved the above result by analyzing the tangent cone associated with the SAV problem. The work is not yet released and we will not include the proof in this thesis.

The observations above motivate us to propose the Binary Fixed-Weight SAV optimization:

Binary Fixed Weights Sum of Absolute Value minimization(BFW-SAV)

$$\text{minimize } \rho||\mathbf{z}||_1 + (1-\rho)||\mathbf{z} - \mathbf{1}_N||_1 \text{ subject to } \Phi\mathbf{z} = \mathbf{y}$$

$$\text{where } \rho = \text{sgn}(p - \frac{1}{2}) \cdot \epsilon + \frac{1}{2}, \epsilon > 0,\ p = \mathbf{P}(z_j = 0), \forall j \in [N]$$

The ideas is that, when $p < 0.5$, we let our RFP curve be the red curve in Figure 4.2; when $p > 0.5$, we let the RFP curve be the blue one. If $p = 0.5$, we let the RFP curve be the green one. To realize this, we choose the weight ρ as

$$\rho = \text{sgn}(p - \frac{1}{2}) \cdot \epsilon + \frac{1}{2}.$$

Or equivalently,

$$\rho = \begin{cases} \frac{1}{2} - \epsilon & p < \frac{1}{2} \\ \frac{1}{2} + \epsilon & p > \frac{1}{2} \\ \frac{1}{2} & p = \frac{1}{2} \end{cases}.$$

4.2.2 On the Reconstruction Performance of SAV

In [21], the author compares the performance of binary signal reconstruction by SAV and BP and shows that SAV outperforms BP. Please refer to the result in Figure 4.3 below, which is copied directly from Fig. 2 subplot 1 of [21].

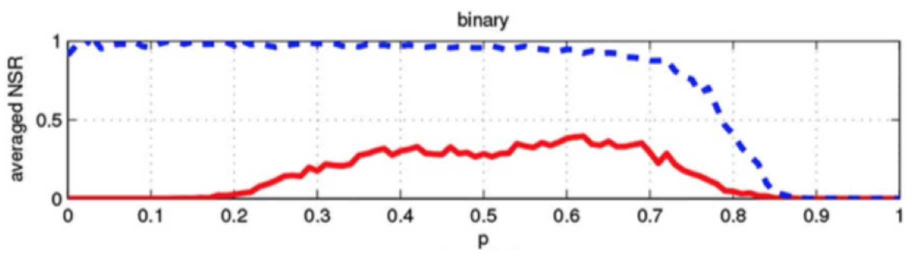

Figure 4.3: (cf. Fig.2 subplot 1 in [21]) Averaged SNR $\frac{||\mathbf{x} - \mathbf{z}||_2}{||\mathbf{x}||_2}$ vs p by SAV (solid) and the basis pursuit (dash)

However, note that the averaged SNR is not really the appropriate reconstruction measurement if we are interested in the exact reconstruction rate. Hence, instead of using averaged SNR, we repeat the experiment using RFP as reconstruction accuracy measurement. We get the following result:

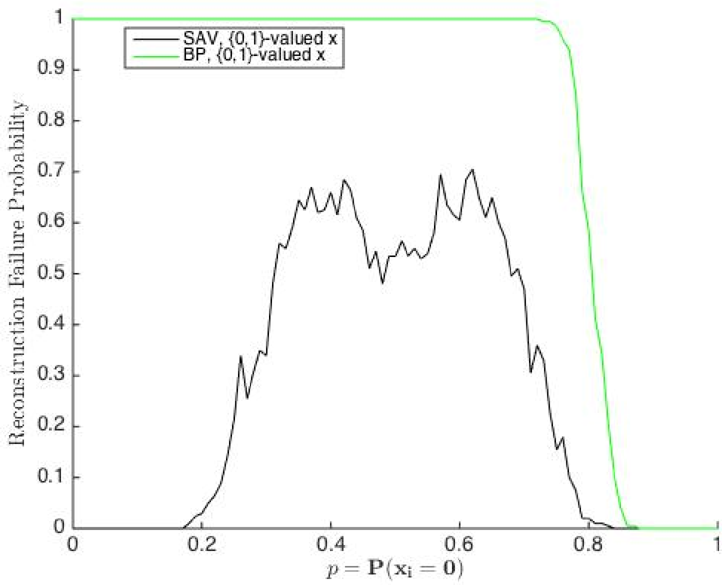

Figure 4.4: Step size of $p = 0.01$, Number of experiments for each p is 200

We shall zoom in and have a closer look at the range of $0.3 \leq p \leq 0.7$. In the meanwhile, we increase the number of trials per each p and make the step-size finer. The result is shown in Figure 4.5.

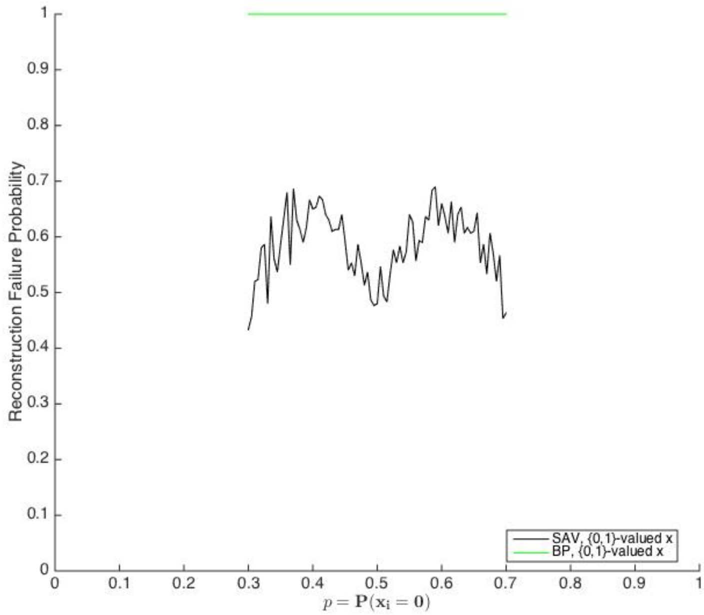

Figure 4.5: Step size of $p = 0.005$, Number of experiments for each p is 300

The experiment results in Figure 4.4 and 4.5 are somewhat surprising. We may ask the following questions:

- Given the step-size fine enough and the number of experiments large enough, why is the SAV curve zigzagging in a large variation?
- Why does the SAV curve has peak near $p = \mathbf{P}(x_i = 1) = 0.35$ and 0.65, and what happens in the "V-shaped" range between 0.35 and 0.65?

So far we do not have solid arguments for the above questions.

4.3 Numerical Experiments on BRSN

Recall the proposed BRSN minimization from the section 3.3.7:

Binary Reweighted Sum of Norm(BRSN)

$$\text{minimize } (1-p)||\mathbf{z}||_1 + p||\mathbf{z} - \mathbf{1}_N||_1 + \lambda\,||\mathbf{z} - \frac{1}{2}\,\mathbf{1}_N||_\infty$$

$$\text{subject to } \Phi\,\mathbf{z} = \mathbf{y},\ \text{ where } p = \mathbf{P}(z_j = 0), j \in [N]$$

4.3.1 On the Parameter Tuning for BRSN

Our first experiment is designed to tune the parameter λ for BRSN. The result is shown in Figure 4.6.

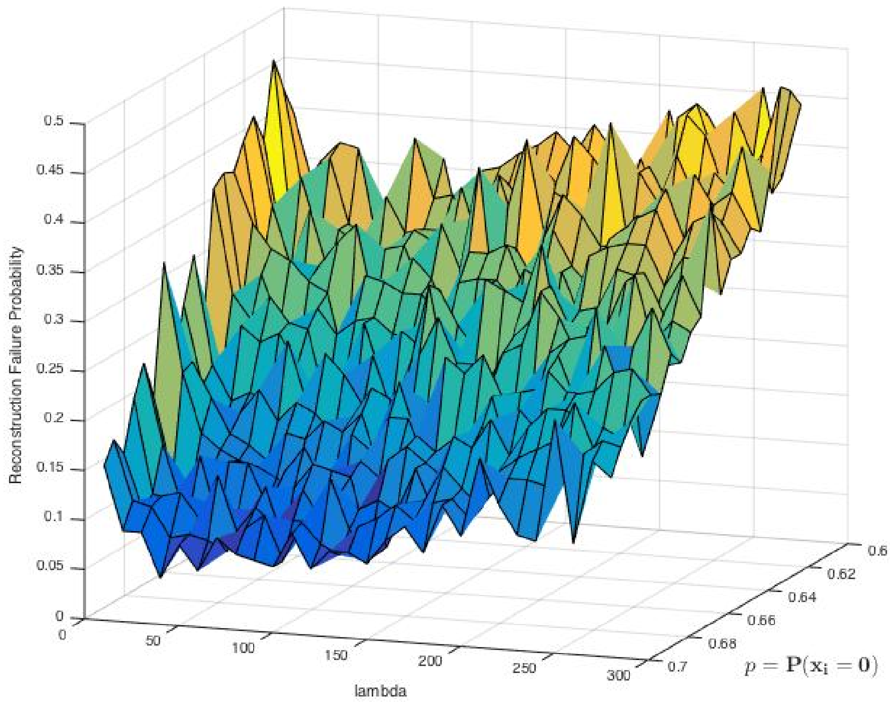

Figure 4.6: Tuning λ for BRSN

Increasing λ from 5 to 275 in the step size of 10, we see that $\lambda = 35$ is a fair choice for BRSN, as it gives one of the lowest RFP curves. In the experiments in the subsequent sections, we fix $\lambda = 35$.

4.3.2 On the Reconstruction Performance of 1-dim Binary Vectors

In the following experiment, increasing p by the step size of 0.01, we compare reconstruction performance of sparse binary vector between BP(Basis Pursuit), Boxed-BP, SAV(Sum of Absolute Values), SN(Sum of Norms), the proposed BFW-SAV(Binary

Fixed Weights Sum of Absolute Values), and the proposed BRSN(Support-based Sum of Norms). The result is shown in Figure 4.7.

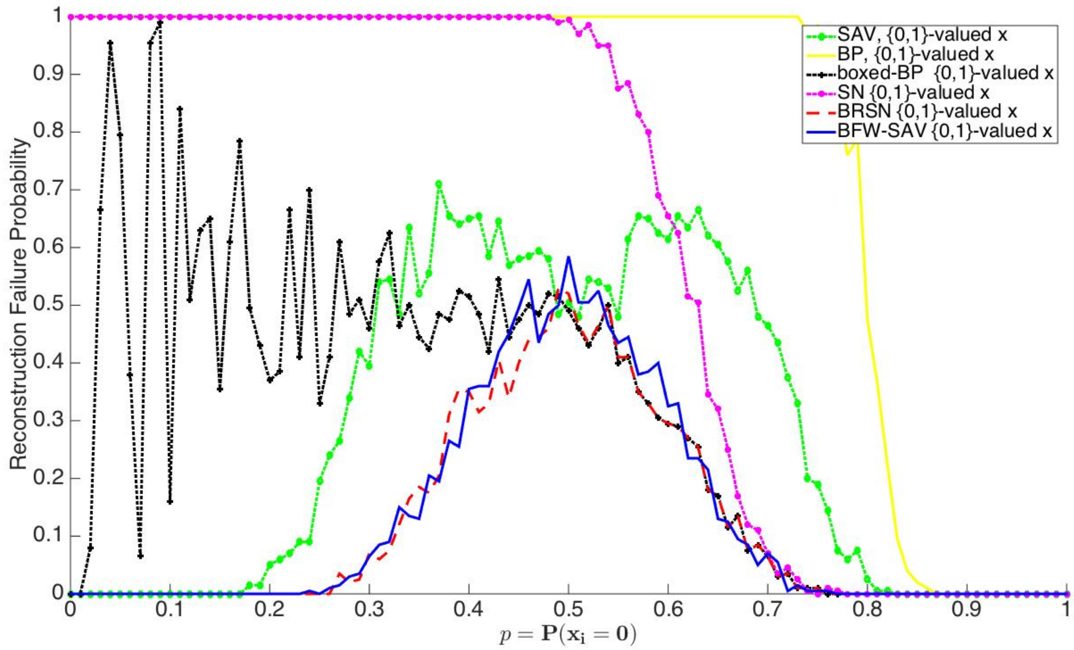

Figure 4.7: RFP of SAV, BP, Boxed-BP, SN and BRSN

We have the following observations:

- Overally, our proposed BFW-SAV(blue) and BRSN(red) are the optimization methods with the two lowest RFP curves.
- Similar as SAV, the BFW-SAV and BRSN have symmetric RFP curves.
- Given the binary signal is dense ($p < 0.5$), BRSN has the better reconstruction accuracy when compared to boxed-BP; on the other hand, if the binary signal is sparse, the former has no advantage compared with the latter.
- The boxed BP has a very unstable reconstruction accuracy, given the signal is dense.
- In the range of $p \in [0.4, 0.6]$, BRSN out-performs BFW-SAV. In other ranges, the comparison is not clear.

4.3.3 On the Reconstruction Performance of 2-dim Bitonal Images

In the following experiment, we compare the reconstruction performance of bitonal (or Black-and-white) image, when using Basis Pursuit, SAV, and proposed BRSN. We essentially mimic the second experiment in [21] and proceed the experiment in the following steps.

1. Consider the 37×37-pixel bitonal image shown on the left of Figure 4.8. We add Gaussian white noise with a mean of 0 and standard deviation of 0.1 to each pixel. The noised image is shown on the right of 4.8. The noised image is another 37×37 real-valued matrix, and we call this matrix X, $X \in \mathbb{R}^{37\times 37}$. We concatenate the matrix column-wisely and get the column vector $\mathrm{vec}(X)$.

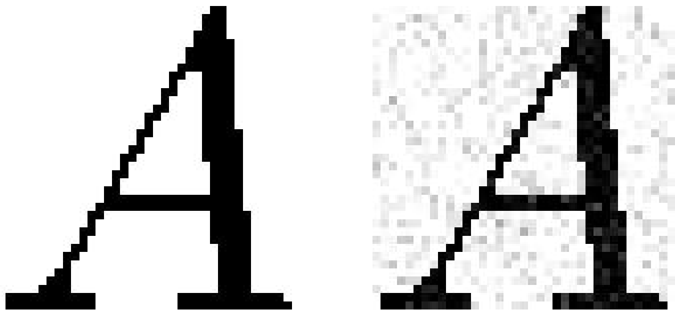

Figure 4.8: The original bitonal image (left) and the noised image (right)

2. Apply the DFT (Discrete Fourier Transformation) to the matrix X and get the transformed matrix $\hat{X} = WXW$, where

$$W := \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{K-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{K-1} & \omega^{2(K-1)} & \cdots & \omega^{(K-1)(K-1)} \end{pmatrix} \tag{4.3.1}$$

where $k = 37$, $\omega := \exp(-2\pi j/K)$. Again,we concatenate $\hat{X}$ column-wisely and get a column vector $\mathrm{vec}(\hat{X})$, which can be written in the following way because of the above relation:

$$\mathrm{vec}(\hat{X}) = (W \otimes W)\mathrm{vec}(X) \in \mathbb{C}^{1369}$$

where $\otimes$ is the Kronecker tensor product.

3. Randomly half-size downsample $\text{vec}(\hat{X})$ to get $\mathbf{y} \in \mathbb{C}^{685}$. Now, the sensing matrix $\Phi \in \mathbb{C}^{685\times 1369}$ is composed of the down-sampling row vectors of $W \otimes W$.

4. Assume $\mathbf{P}(\text{vec}(\hat{X})_i = 1) = \mathbf{P}(\text{vec}(\hat{X})_i = 0) = \frac{1}{2}$. Apply BP, SAV, and BRSN on

$$\Phi \ \text{vec}(X) = \mathbf{y}$$

to reconstruct the $\text{vec}(X)$.

The experiment results are shown in Figure 4.9.

Figure 4.9: Bitonal image reconstruction results using BP(left), SAV(middle), and proposed BRSN(right)

The results show that the performance of the proposed BRSN is somewhat better than that of BP and SAV.

Notice that the BFW-SAV is equivalent to SAV under our assumption of

$$\mathbf{P}(\text{vec}(\hat{X})_i = 1) = \mathbf{P}(\text{vec}(\hat{X})_i = 0) = \frac{1}{2},$$

and therefore we do not have to consider BFW-SAV in this experiment. The reconstruction results from SN and boxed BP are extremely poor and we omit them here.

4.4 Summary

We summarize our conclusion from the previous experiments:

- We can choose the weights as fixed numbers in SAV and get a better reconstruction result. The resulted BFW-SAV out-performs the original SAV in the binary signal reconstruction.

- BRSN outperforms l_1, SAV, SN in binary signal and bitonal image reconstruction.
- In the case that the to-be-recovered-signal is sparse, BRSN has the same reconstruction performance as Boxed BP; on the other hand, in the case that the to-be-recovered signal is dense, BRSN has the more accurate and stable reconstruction performance.

Chapter 5

Conclusion and Open questions

In this thesis, we study the Discrete CS in two perspectives: first, we develop the unique reconstruction guarantee of discrete signals from random matrices; second, we provide a survey of convex optimization methods for discrete signals, and particularly we propose the BFW-SAV(Binary Fixed Weights Sum of Absolute Values) and BRSN for binary signal reconstruction.

Note that, however, there are various of limitations of our approaches in both perspectives:

- Although the Discrete NSP is weaker than the General NSP, the gap between Discrete NSP and General NSP is in too thin to yield any useful property. One might notice that in the experiments, we did not use the Discrete NSP to guarantee the unique reconstruction. In fact, Gaussian random matrices satisfy the General NSP, and therefore automatically satisfy Discrete NSP. Discrete NSP does not give us more choices for random matrices.

  Dr. Dae Gwan Lee and coworkers developed the "Weak Null Space Property," which is weaker than the Discrete Null Space property. We will not include the Weak Null Space Property in this thesis since the work is not yet released.

- From our experiment, our proposed BRSN has the same reconstruction performance as boxed l_1 in the case of p close to 1, i.e., when the signal is sparse enough. Although it is true that BRSN has the better performance in other cases, it somehow deviate from our original interests in the K-sparse model, where $K \ll N$.

- The proposed BFW-SAV and BRSN together with SAV, SN, l_1, boxed l_1, have the drawbacks of high computational complexity. We perform these optimization methods by using the **CVX** toolbox, which relies on the interior point method of linear programming.

In below we list some topics are still open for discussion and should be investigated in the future. Particularly, we classify them into three categories.

- **Category I**: in this category, we list those questions we observe in the experiments but fail to explain in the theoretical level.
  - The non-symmetric, bimodal, zigzagging behaviour of SAV described in the end of section 4.2.2.
  - The unstable behaviour of Boxed-BP in the range $p < 0.5$ described in the end of section 4.3.2.
  - The inferior reconstruction performance of Boxed-BP on bitonal image, given its superior reconstruction performance on binary signal reconstruction. It was mentioned in the end of section 4.3.3.
- **Category II**: In this category, we present two concrete future research orientations. The difference between Category II and I is that, for the questions in Category II, we do not have evidence from the experiments.
  - Study the following optimization method under the Binary CS scheme:

$$\text{minimize } (1-p)||\,\mathbf{z}\,||_1 + p||\,\mathbf{z} - \mathbf{1}_N^T||_1 + \lambda \sum_{i<j} \max\,\{|x_i - \frac{1}{2}|, |x_j - \frac{1}{2}|\}$$

$$\text{subject to } \Phi\,\mathbf{z} = \mathbf{y},\ \text{ where } p = \mathbf{P}(z_j = 0), \forall j \in [N]$$

    In the above minimization, we replace the uniform infinity norm in BRSN by the pairwise infinity norms. This replacement aims to punish the entries that have distinct magnitude harder. However, this optimization, though convex, can not be realized in **CVX** toolbox due to its combinatorial nature. But we believe that it could be solved by using the "proximal splitting algorithms" proposed in [28].
  - Follow the same routine as in section 4.2.1 to check that if we can select better weights for BRSN. That is to say, we study the following optimization,

$$\text{minimize } (1-\frac{t}{N})||\,\mathbf{z}\,||_1 + \frac{t}{N}||\,\mathbf{z} - \mathbf{1}_N||_1 + \lambda\,||\,\mathbf{z} - \frac{1}{2}\,\mathbf{1}_N||_\infty$$

$$\text{subject to } \Phi\,\mathbf{z} = \mathbf{y},\; N = \dim(\mathbf{z}).$$

where t is just some number that is no longer related to the sparsity. Particularly, it is important to understand the tangent cone associated with the BRSN problem.

- **Category III**: In this category, we present some open questions in the general sense. The difference between Category III and II is that, for the open questions in Category III, we do not have a concrete approach.
  - Develop the Discrete Restricted Isometry Property, which should be strong enough to imply Discrete Null Space Property, but still weak enough to relax the condition of General Restricted Isometry Property.
  - Study the *Phase Transition Diagram* [11] of SAV, BP, boxed-BP, SN, BFW-SAV, and BRSN.
  - Study the reconstruction performance of SAV, BP, boxed-BP, SN, BFW-SAV, and BRSN under noise.
  - Develop methods that systematically tunes the λ scaling scalar in SN and BRSN.
  - Consider the Discrete CS problem in the non-convex-optimization scenario.
  - Develop greedy methods tailored for discrete signal reconstruction which have low computational complexity.

References

[1] E. Candes. *Compressive sampling.* In Proc. Int. Congress of Math., Madrid, Spain, Aug. 2006.

[2] E. Candes and J. Romberg. *Quantitative robust uncertainty principles and optimally sparse decompositions.* Found. Comput. Math., 6(2):227-254, 2006

[3] E. Candes, J. Romberg, and T. Tao. *Stable signal recovery from incomplete and inaccurate measurements.* Comm. Pure Appl. Math., 59(8):1207-1223, 2006.

[4] E. Candes and T. Tao. *Near optimal signal recovery from random projections: Universal encoding strategies?* IEEE Trans. Inform. Theory, 52(12):5406-5425, 2006

[5] D. L. Donoho. *Compressed sensing.* IEEE Trans. Inform. Theory, 52(4):1289-1306, 2006.

[6] Bioglio, Valerio, Giulio Coluccia, and Enrico Magli. *Sparse image recovery using compressed sensing over finite alphabets.* Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014.

[7] Sasahara, Hampei, Kazunori Hayashi, and Masaaki Nagahara. *Symbol Detection for Frame-Based Faster-than-Nyquist Signaling via Sum-of-Absolute-Values Optimization.* arXiv preprint arXiv:1508.06021 (2015).

[8] Bemporad, Alberto, and Manfred Morari. *Control of systems integrating logic, dynamics, and constraints.* Automatica 35.3 (1999): 407-427.

[9] Sasahara, Hampei, Kazunori Hayashi, and Masaaki Nagahara. *Multiuser Detection by MAP Estimation with Sum-of-Absolute-Values Relaxation.* arXiv preprint arXiv:1510.07273 (2015).

[10] Knoop, Benjamin, et al. *Compressed sensing K-best detection for sparse multi-user communications.* Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European. IEEE, 2014.

[11] Donoho, David L., and Jared Tanner. *Precise undersampling theorems.* Proceedings of the IEEE 98.6 (2010): 913-924.

[12] H. Boche, Robert Calderbank, Gitta Kutyniok, and Jan Vybíral. *Compressed Sensing and Its Applications: MATHEON Workshop 2013*(pp. 6-7). Cham: Birkhäuser, 2015. Print.

[13] S. S. Chen, D.L. Donoho, M.A. Saunders, *Atomic decomposition by Basis Pursuit.* SIAM J. Sci. Comput. 20(1), 33-61 (1999) (Cited on pp. 34, 36, 71.)

[14] S. Mallat, Z. Zhang, *Matching pursuits with time-frequency dictionaries.* IEEE Trans. Signal Process. 41(12), 3397-3415 (1993) (Cited on pp. 34, 36, 71.)

[15] Pati, Yagyensh Chandra, Ramin Rezaiifar, and P. S. Krishnaprasad. *Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition.* Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on. IEEE, 1993.

[16] Tropp, Joel A., and Anna C. Gilbert. *Signal recovery from random measurements via orthogonal matching pursuit.* Information Theory, IEEE Transactions on 53.12 (2007): 4655-4666.

[17] Maleki, Arian, and David L. Donoho. *Optimally tuned iterative reconstruction algorithms for compressed sensing.* Selected Topics in Signal Processing, IEEE Journal of 4.2 (2010): 330-341.

[18] S. Foucart, and H. Rauhut. *A Mathematical Introduction to Compressive Sensing.* New York: Springer, 2013. Print.

[19] Mangasarian, Olvi L., and Benjamin Recht. *Probability of unique integer solution to a system of linear equations.* European Journal of Operational Research 214.1 (2011): 27-30.

[20] Aissa-El-Bey, Abdeldjalil, et al. *Sparsity-Based Recovery of Finite Alphabet Solutions to Underdetermined Linear Systems.* Information Theory, IEEE Transactions on 61.4 (2015): 2008-2018.

[21] Nagahara, Masaaki. *Discrete Signal Reconstruction by Sum of Absolute Values*. Signal Processing Letters, IEEE 22.10 (2015): 1575 -1579.

[22] Wang, Sheng, and Nazanin Rahnavard. *Binary Compressive Sensing via Sum of l1-Norm and l (infinity)-Norm Regularization*. Military Communications Conference, MILCOM 2013-2013 IEEE. IEEE, 2013.

[23] Michael Grant and Stephen Boyd. *CVX: Matlab software for disciplined convex programming, version 2.0 beta*. http://cvxr.com/cvx, September 2013.

[24] Michael Grant and Stephen Boyd. *Graph implementations for nonsmooth convex programs*, Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, pages 95-110, Lecture Notes in Control and Information Sciences, Springer, 2008. http://stanford.edu/~boyd/graph_dcp.html.

[25] Baraniuk, Richard G. *Compressive sensing.* IEEE signal processing magazine 24.4 (2007).

[26] Studer, Christoph, Wotao Yin, and Richard G. Baraniuk. *Signal representations with minimum* l_∞*-norm.* Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on. IEEE, 2012.

[27] Fuchs, Jean-Jacques. *Spread representations.* Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on. IEEE, 2011.

[28] Zeng, Xiangrong, and Mario AT Figueiredo. *Solving OSCAR regularization problems by fast approximate proximal splitting algorithms.* Digital Signal Processing 31 (2014): 124-135.

List of Figures

Appendix

Acronyms

BP .... **B**asis **Pursuit**
BRSN .... **B**inary **R**eweighted **S**um of **N**orm
CS .... **C**ompressive **S**ensing
CVX .... Matlab Software for Disciplined **Con**v**ex** Programming
DFT .... **D**isecrete **F**ourier **T**ransformation
IT .... **I**terative **T**hreshold
MATLAB .... **MAT**rix **LAB**oratory, Numeric Programming Software
NSP .... **N**ull **S**pace Property
MP .... **M**atching **P**ursuit
NSP .... **N**ull **S**pace **P**roperty
OMP .... **O**rthogonal **M**atching **P**ursuit
RF .... **R**econstruction **F**ailure **P**robability
SAV .... **S**um of **A**bsolute **V**alue minimization
SN .... **S**um of **N**orm minimization
SNR .... **S**ignal to **N**oise **R**atio

Symbols

$\mathbb{C}$ .... Field of Complex Numbers
Card .... Cardinality of a set
K .... Set that contains locations of non-zero entries of a vector.
$\overline{K}$ .... Complement of the set K.
ker .... Kernel of a matrix
$[N]$ .... $\{1, 2, \ldots, N\}$
Φ .... Sensing Matrix

$\mathbb{R}$ .... Field of Real Numbers
sgn .... the signum function that extracts the sign of a real number or vector
supp .... Support of a vector