WORCESTER POLYTECHNIC INSTITUTE

Compressed Sensing via Partial ℓ_1 Minimization

by

Lu Zhong

A thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science

in

Data Science

April 2017

APPROVED:

Professor Randy C. Paffenroth, Advisor:

_____

Professor Andrew C. Trapp, Reader:

_____

Contents

WORCESTER POLYTECHNIC INSTITUTE

Abstract

Data Science

Master of Science

by Lu Zhong

Reconstructing sparse signals from undersampled measurements is a challenging problem that arises in many areas of data science, such as signal processing, circuit design, optical engineering and image processing. The most natural way to formulate such problems is by searching for sparse, or parsimonious, solutions in which the underlying phenomena can be represented using just a few parameters. Accordingly, a natural way to phrase such problems revolves around ℓ_0 minimization in which the sparsity of the desired solution is controlled by directly counting the number of non-zero parameters. However, due to the nonconvexity and discontinuity of the ℓ_0 norm such optimization problems can be quite difficult. One modern tactic to treat such problems is to leverage *convex relaxations*, such as exchanging the ℓ_0 norm for its convex analog, the ℓ_1 norm. However, to guarantee accurate reconstructions for ℓ_1 minimization, additional conditions must be imposed, such as the *restricted isometry property*. Accordingly, in this thesis, we propose a novel extension to current approaches revolving around *truncated* ℓ_1 *minimization* and demonstrate that such approach can, in important cases, provide a better approximation of ℓ_0 minimization. Considering that the nonconvexity of the truncated ℓ_1 norm makes truncated ℓ_1 minimization unreliable in practice, we further generalize our method to *partial* ℓ_1 *minimization* to combine the convexity of ℓ_1 minimization and the robustness of ℓ_0 minimization. In addition, we provide a tractable iterative scheme via the augmented Lagrangian method to solve both optimization problems. Our empirical study on synthetic data and image data shows encouraging results of the proposed partial ℓ_1 minimization in comparison to ℓ_1 minimization.

Keywords. Compressed sensing, ℓ_0 minimization, ℓ_1 minimization, truncated ℓ_1 minimization, partial ℓ_1 minimization, the augmented Lagrangian method, the restricted isometry property

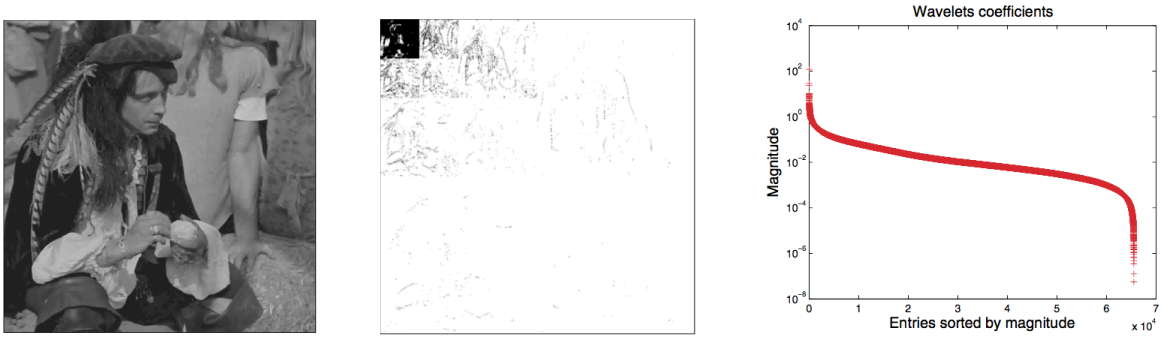
Chapter 1

Introduction

Compressed sensing (CS) is a technique for reconstructing sparse or compressible signals accurately from very limited measurements [1] and has attracted considerable recent attention. It provides a protocol for sensing and compressing information simultaneously [2]. Originally thought of as a technique for signal processing and first proposed by Candès, Romberg and Tao in 2006, CS has found connections and applications in many domains such as applied mathematics, computer science, and electrical engineering [28]. CS has been widely applied in magnetic resonance imaging (MRI) [6], image processing [7], imaging techniques ([9], [31]), holography [29], and so forth. Such broad applicability is because CS is quite effective when taking measurements is costly, in terms of either the cost of a sensor, i.e. sensing outside of the visible frequency band, or the time it takes to acquire a piece of information. For example, MRI is an essential medical imaging tool with an inherently slow data acquisition process. Applying CS to MRI offers potentially significant scan time reductions, with benefits for patients and health care.

However, for successful application of CS, there are two key requirements. The first one is that the signal of interest, $\mathbf{u}$, should be *sparse*. By sparse, we mean that only a small fraction of the entries are nonzero, or compressible. Many real-world signals are naturally compressible by sparse coding in appropriate transform domains [3], i.e. the images we take in our daily life are sparse in the wavelet domain. The second requirement is that the measurement matrix $\mathbf{A}$ should satisfy the *restricted isometry property(RIP)*, and this requirement limits the application of CS. Roughly speaking, RIP indicates that $\mathbf{A}$ behaves like an almost orthonormal system but only for *sparse* linear combinations (the formal definition of RIP is given in Section 2.1.2). RIP limits the application of CS because some real-world sensing systems do not satisfy this

property, and what makes this issue even more problematic is that checking to what degree the system satisfy RIP is not computationally tractable. In practice, it is well known that almost all random matrices, such as Gaussian, Bernoulli, and Fourier ensemble, satisfy RIP with high probability. But again, there is no practical algorithm for verifying whether a given matrix has this property, and to use random matrices, in some cases, is very expensive. The study of RIP will be the key focus of this thesis.



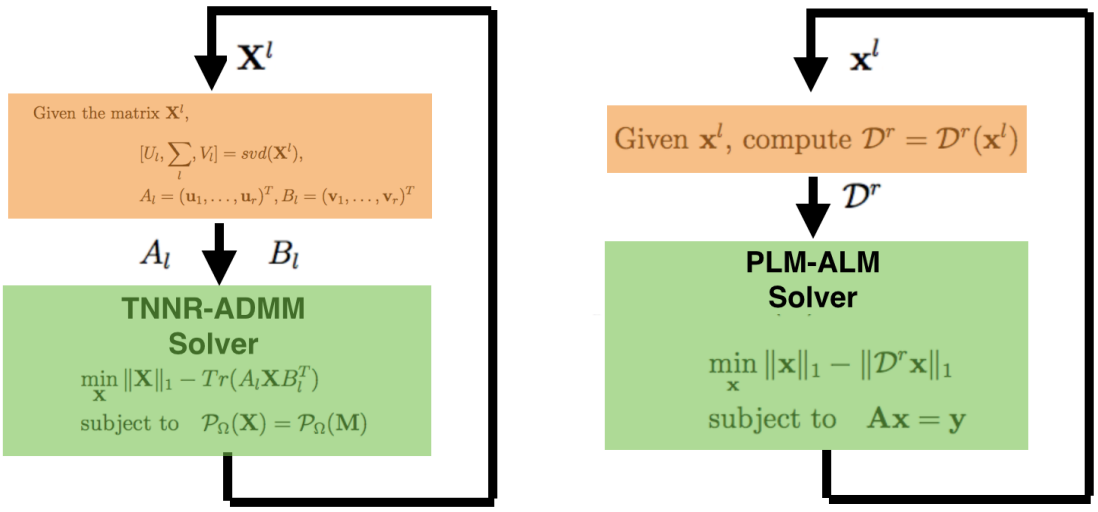
(a) 1024×1024-pixel image (b) Locations of significant wavelet coefficients (c) compressibility

Figure 1.1: In general, significant wavelet coefficients of natural images concentrate at the upper left corner, as shown in the middle plot. The plot on the right displays that significant coefficients are only a small fraction of all the coefficients, while the rest decay quickly.

This thesis considers the model problem of reconstructing a sparse signal $\mathbf{u} \in \mathbb{R}^N$ from under-sampled measurements $\mathbf{y} = \mathbf{A}\mathbf{u}$, where $\mathbf{y} \in \mathbb{R}^m$, with $m \ll N$, and $\mathbf{A} \in \mathbb{R}^{m \times N}$ is the known, nonadaptive measurement matrix. The problem we are interested in is whether it is possible to recover $\mathbf{u}$ exactly from $\mathbf{y}$ when $\mathbf{A}$ is less/non-RIP, and if so, how? By less/non-RIP, we mean that a given $\mathbf{A}$ does not satisfy RIP with high probability, or $\mathbf{A}$ does not behave like an orthonormal basis, for example including a pair of columns that are almost identical.

One of the main contributions of this thesis is a careful treatment of RIP. We propose a novel optimization problem, called *partial* ℓ_1 *minimization*(PLM), aiming to be more tolerant to less/non-RIP $\mathbf{A}$ compared to the classical ℓ_1 minimization. ℓ_1 minimization achieves the sparsest solution by minimizing the sum of all absolute values of entries in $\mathbf{u} \in \mathbb{R}^N$. On the other hand, PLM is only concerned with the sum of the absolute values of the entries in the set $\{1, \ldots, N\} \setminus \Omega_{\mathcal{P}}$, with the assumption that $\Omega_{\mathcal{P}}$ is a known set that overlaps with the support of $\mathbf{u}$ (the support of a vector $\mathbf{u} \in \mathbb{R}^N$, $supp(\mathbf{u})$, is defined as: $supp(\mathbf{u}) := \{i | u_i \neq 0, \forall\, i \in \{1, \ldots, N\}\}$). To show how the requirements of RIP can be relaxed given $\Omega_{\mathcal{P}}$, we perform experiments covering a wide range of cases, including where $\Omega_{\mathcal{P}}$ belongs to $supp(\mathbf{u})$, as well

as where $\Omega_{\mathcal{P}}$ has some spurious indices outside of $supp(\mathbf{u})$. With the concern that evaluating a reconstruction from less/non-RIP measurement matrix by standard error representations is not reliable, we present a thoughtful evaluation metric to identify acceptable reconstructions with the consideration of significant entries, as well as the support. Even more, how to derive less/non-RIP measurement matrices is also important in our work. So, we give details for generating less/non-RIP matrices using spherical coordinate systems that correspond to the geometrical interpretation of RIP. Finally, a generalization of the PLM solver, called PLM-ALM, is provided via solving a convex, unconstrained problem formed by using the augmented Lagrangian method. The experiments on synthetic data demonstrate that PLM is more robust to less/non-RIP $\mathbf{A}$ than ℓ_1 minimization, and performs quite well even if some spurious indices are included in $\Omega_{\mathcal{P}}$. Another good property of PLM is that it is not concerned with how large the entries in the given $\Omega_{\mathcal{P}}$ are compared with other significant ones. In addition, we provide details of the *truncated* ℓ_1 *norm*(TLM), which is another optimization problem we proposed in this thesis. It plays an important role in incubating the idea of PLM. The key idea underlying PLM and TLM is to subtract part of entries from the ℓ_1 norm so that the directions corresponding to the truncated coefficients would be free. Denote the subtracted index set as Ω. The most significant difference between PLM and TLM is the selection rule of Ω. Ω in PLM is fixed and given by, say, an Oracle who is able to reveal hidden knowledge or information, and it does not attempt to verify this given information. Ω in TLM is the set of the largest r entries of the unknown, to-be-reconstructed signal, and accordingly is adaptive to $\mathbf{x}$ in the sparest solution searching process. While PLM is convex, TLM is nonconvex. In [4], the two-step iterative framework provides an efficient way to solve the nonconvex truncated nuclear norm regularization in matrix completion. Building on this idea, we derive a two-step iterative scheme to solve TLM, by proving a theorem. The solver for TLM is called TLM-2SI, where convergence to a local minimum is guaranteed.



(a) two-step Iterative Algorithm Framework in TNNR

(b) two-step Iterative Algorithm Framework in TLM

Figure 1.2: two-step iterative TNNR framework (shown on the left) proposed in [4] provides an efficient procedure to deal with the nonconvex truncated nuclear norm regularization arising in matrix completion. We borrow this framework to formulate our two-step iterative TLM framework (shown on the right) to solve the nonconvex truncated ℓ_1 minimization problem. Details about PLM-ALM solver are in Section 4.2. The generalization of (b) can be thought of similar to (a), though the details of the method and the solver are, of course, quite different because we are using this methodology on quite different problems.

The remainder of the thesis is organized as follows: In the next chapter, we recall a brief description of the necessary background and the related works that form our ideas. Chapter 3 presents truncated ℓ_1 minimization (TLM) and the corresponding solver TLM-2SI, followed by the observations that inspire our idea about PLM. In Chapter 4, we propose partial ℓ_1 minimization (PLM), and the corresponding solver, PLM-ALM. Experimental results are displayed in Chapter 5. Chapter 6 is devoted to concluding remarks and future works.

Chapter 2

Background

This chapter presents the necessary background and the related works that inspire our ideas and methods. The basic knowledge and core conclusions in CS are detailed first. Then we introduce the augmented Lagrangian method, which helps us to replace a constrained optimization problem with a convex, unconstrained problem. Moreover, we give a brief introduction on the work of truncated nuclear norm regularization. As last, we introduce a Bayesian variable selection, called spike-and-slab, to exhibit that how the sparsity models are solved in statistical learning theory.

2.1 Two Requirements in CS

The reconstruction of an arbitrary vectors from vastly underdetermined linear systems is an ill-posed problem and requires additional information to make the reconstruction possible. Accordingly, the two fundamental requirements in CS are that [1]

- $\mathbf{u} \in \mathbb{R}^N$ is sufficiently sparse, or compressible, in some known basis,
- $\mathbf{A} \in \mathbf{R}^{m \times N}$ obeys the restricted isometry property (RIP) which, roughly speaking, means that $\mathbf{A}$ behaves like an almost orthonormal system but only for *sparse* linear combinations.

2.1.1 Sparsity

Before explaining why the requirement of sparsity is necessary for signal reconstruction from undersampled measurements, we give the definition of sparsity as:

Definition 2.1.1. A vector $\mathbf{x} \in \mathbb{R}^N$ is called *s - sparse* if at most s of its entries are nonzero.

Given $\mathbf{A} \in \mathbb{R}^{m \times N}$, $m \ll N$, suppose that $\mathbf{u}$ is not sparse at all, meaning all of the entries of $\mathbf{u}$ are nonzero. As $m < N$, $\mathcal{N}(A) \setminus \mathbf{0} \neq \varnothing$, where $\mathcal{N}(A) := \{\mathbf{b} : \mathbf{Ab} = \mathbf{0}\}$ is the *null-space* of A. Correspondingly,

$$\mathbf{Au} = \mathbf{A}(\mathbf{u} + \beta) = \mathbf{y}, \qquad \forall \beta \in \mathcal{N}(A).$$

Obviously, reconstruction such a $\mathbf{u}$ from $\mathbf{y} = \mathbf{Au}$ is impossible. Therefore, the requirement that $\mathbf{u}$ has to be sparse is necessary. In practice, compressibility, which is to say that the signal of interested can be well-approximated by a sparse solution, is preferable as it is generally a weaker assumption than sparsity. This is because that most of real world signals are naturally compressible (while perhaps contaminated with noise). In this thesis, we focus our interest on the reconstruction of sparse $\mathbf{u}$.

2.1.2 RIP and Random Matrix

Once $\mathbf{u}$ is known as an s-sparse vector, we are able further to constrain the property of the measurement matrix $\mathbf{A}$ to reconstruct $\mathbf{u}$. This is where the second requirement comes from. Suppose that the only s-sparse vector $\beta \in \mathcal{N}(A)$ is $\mathbf{0}$, then any s-sparse $\mathbf{u}$ can be reconstructed from $\mathbf{y}$. This property is sometimes called the *null space property (NSP)* [3]. Even though the NSP is a straightforward description of the reconstruction condition, it is not sufficient to guarantee the reconstruction when $\mathbf{y}$ is contaminated with noise or has been corrupted by some error. Therefore, Candès and Tao proposed to use the RIP [5] in their recovery proofs as it is applicable in sparse signal reconstruction from noisy or noiseless measurements.

Definition 2.1.2. A matrix $\mathbf{A}$ satisfies the *restricted isometry property* with δ_s, if, for all s-sparse vector $\mathbf{x} \in \mathbb{R}^N$,

$$(1 - \delta_s)\|\mathbf{x}\|_2^2 \leq \|\mathbf{Ax}\|_2^2 \leq (1 + \delta_s)\|\mathbf{x}\|_2^2, \tag{2.1}$$

where $\delta_s \in (0, 1)$ is the smallest quantity to satisfy (2.1).

RIP requires all the possible combination of column vectors with the cardinality at most s to behave like an almost orthonormal system. If $\mathbf{A}$ satisfies RIP of the order $2s$, then $\mathbf{A}$ approximately preserves the distance between any pair of s-sparse vectors. we can say that if $\mathbf{A}$ satisfies RIP, it also satisfies NSP, but not vice versa [2]. Based on the equation (2.1), the smaller δ_{2s} is, the better such distance is preserved. The optimal case is that δ_{2s} is equal

to 0, then there is no distance information missing or distorted via the transformation. This is exactly what happens in the orthonormal basis. However, in our case, $\mathbf{A}$ is a rectangular matrix. In addition, to compute δ_s is NP-hard [37]. For example, computing δ_s in a brute force fashion would require testing all entries of x, followed by all pairs of entries in x, followed by all triples, and so forth. Such a calculation is not practical. In practice, random matrices are used. They satisfy RIP with overwhelming probability [5]. For example, suppose that m, N, and δ_{2s} are given and fixed. For a Gaussian or Bernoulli measurement matrix, the RIP holds for such matrix with the prescribed δ_{2s} and any $s \leq c_1 m / \log \frac{N}{s}$ with probability $\geq 1 - 2e^{-c_2 m}$. The constants c_1 and c_2 depend only on δ_{2s} [38]. In this thesis, Gaussian matrices are used as the measurement matrix for conducting our experiments.

2.2 Theoretical δ_{2s} Bound for Reconstruction

For sparse signal reconstruction, ℓ_0 minimization and ℓ_1 minimization are two well-understood optimization problems. The sparsity requirement directly leads us to reconstruct a signal by searching for the sparsest solution $\mathbf{x}$ which satisfies the constraint $\mathbf{Ax} = \mathbf{y}$, that is ℓ_0 minimization and formulated as follows:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^N} \quad & \|\mathbf{x}\|_0 \\ \text{subject to} \quad & \mathbf{Ax} = \mathbf{y}, \end{aligned} \tag{2.2}$$

where $\|.\|_0$ is the ℓ_0 *norm*, equal to the number of nonzero entries in a vector, that is

$$\|\mathbf{x}\|_0 := \text{card}\,(i \,|\, x_i \neq 0, i \in [N]),\ \mathbf{x} \in \mathbb{R}^N.$$

However, $\|\mathbf{x}\|_0$ is not convex, and solving (2.2) is NP-hard [3]. One corresponding convex relaxation of ℓ_0 minimization is ℓ_1 *minimization*, and is formulated as follows:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^N} \quad & \|\mathbf{x}\|_1 \\ \text{subject to} \quad & \mathbf{Ax} = \mathbf{y}, \end{aligned} \tag{2.3}$$

where $\|.\|_1$ is the ℓ_1 *norm*, defined as $\|\mathbf{x}\|_1 := \sum_{i=1}^{N} |x_i|$.

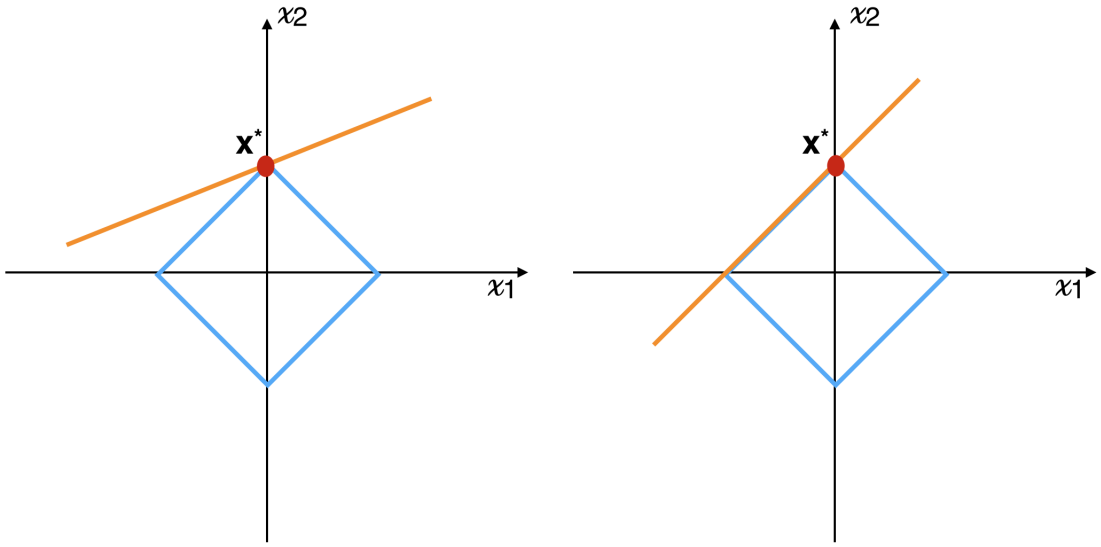


Figure 2.1: In the 2-dimensional space, the set of vectors such that $\|\mathbf{x}\| = 1$ is a diamond. The sparse vector $\mathbf{x}^*$ locates at a coordinate axis. The set of vectors $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{y}$ is a plane passing through the point $\mathbf{x}^*$. As shown in the figure on the left, most planes that pass through x^*, intersect the diamond only at the point x^*. This is the situation where the ℓ_0 minimization is equivalent to the ℓ_1 minimization. However, when the plane unfortunately overlaps with part of the diamond, as shown in the figure on the right, the ℓ_1 minimization is not the same as the ℓ_0 minimization. Of course, in higher dimensions one must consider the intersection of the constraint with various higher dimensional faces/edges.

The exact reconstruction condition is the one under which, given $\mathbf{A} \in \mathbb{R}^{m \times N}$ and $\mathbf{y} = \mathbf{A}\mathbf{u}$, $\mathbf{u}$ is the unique minimizer to the corresponding optimization problem, $\forall\, s$-sparse $\mathbf{u}$. The theoretical upper bound on δ_{2s} for guaranteeing exact reconstruction for ℓ_0 minimization is

$$m \sim \mathcal{O}(s \ln \tfrac{N}{s}), \text{ and } \delta_{2s} < 1,$$

while the one for ℓ_1 minimization is

$$m \sim \mathcal{O}(s \ln N), \text{ and } \delta_{2s} < \tfrac{1}{3}.$$

Obviously, the requirement on δ_{2s} for ℓ_1 minimization is roughly "three times as strict" as the one for ℓ_0 optimization [5]. However, δ_{2s} cannot be accurately computed as previously mentioned. And random matrices are only said to satisfy RIP with high probability when the number of measurements, m, is sufficient, for given δ_{2s} and s.

2.3 The General Sampling Process in CS

As shown in Figure 1.1, the image is not sparse when thought of as pixels, but it can be well compressible in the wavelet domain. In general, to get measurements $\mathbf{y}$ via sampling $\mathbf{z}$, $\mathbf{A}$, actually, is the combination of two operations or steps [11].

$$\mathbf{A} = \mathbf{\Phi}^* \mathbf{\Psi}.$$

For example, we first get a pixel image flattened as a vector $\mathbf{z} \in \mathbb{R}^N$, then use $\mathbf{\Psi}$ to transform $\mathbf{z}$ into a sparse representation $\mathbf{u}$. The rest of things is to pick up m rows from an appropriate domain $\mathbf{\Phi}$ so that to get $\mathbf{\Phi}^* \in \mathbb{R}^{m \times N}$, then use $\mathbf{\Phi}^*$ to encode the sparse $\mathbf{u}$, and finally get the measurements $\mathbf{y}$.

Formally speaking, the sparsity basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ is an orthornomal operator to project the original signal $\mathbf{z} \in \mathbb{R}^N$ into a space with the sparse representation $\mathbf{u} \in \mathbb{R}^N$.

$$\mathbf{u} = \mathbf{\Psi z} \text{ , } \mathbf{u} \text{ is sparse.}$$

If $\mathbf{z}$ is sparse already, $\mathbf{\Psi}$ simply is an identical matrix $\mathbf{I}_N$.

The usually choice of the sensing basis, $\mathbf{\Phi} \in \mathbb{R}^{N \times N}$, is an orthonormal basis as well. Because this thesis involves the use of less/non-RIP matrix, our choice will go beyond orthonormal basis (see details in Section 5.1.1). Next, m rows in $\mathbf{\Phi}$ are uniformly, and randomly picked to form the sensing matrix $\mathbf{\Phi}^*$. Then, the measurements $\mathbf{y} \in \mathbb{R}^m$ is generated by applying $\mathbf{\Phi}^*$ on the sparse representation $\mathbf{u}$.

Correspondingly, to reconstruct $\mathbf{z}$, we first reconstruct $\mathbf{u}$ from $\mathbf{y}$, and then project $\mathbf{u}$ back to the original basis through

$$\mathbf{z} = \mathbf{\Psi}^{-1}\mathbf{u}.$$

However, there are some applications for which the signal $\mathbf{z}$ is sparse not in an orthonormal basis, but in some overcomplete dictionary [35]. This means that $\mathbf{u} = \mathbf{\Psi z}$, where $\mathbf{\Psi}$ is a redundant dictionary. When $\mathbf{\Psi}$ is not orthonormal, $\mathbf{A} = \mathbf{\Phi}^* \mathbf{\Psi}$ is not at all likely to satisfy RIP. From such reality, the interest of relaxing RIP occurs to us again.

For simplification, in this thesis, we use $\mathbf{A}$ to denote the sampling process, and consider the

input vector $\mathbf{u}$ is sparse already.

2.4 The Augmented Lagrangian Method

The augmented Lagrangian method is an efficient way to transform a constrained optimization problem, i.e. ℓ_1 minimization, into an unconstrained, and strongly convex function [36].

Definition 2.4.1. Let $\mathbf{X}$ be a convex set in a real vector space, and $f : X \mapsto \mathbb{R}$ be a function, then with the parameter $m \in (0, 1)$, f is called *strongly convex* if and only if, $\forall\, \mathbf{x_1}, \mathbf{x_2} \in \mathbf{X}$ and $\forall\, t \in [0, 1]$, it satisfies

$$f\left(t\mathbf{x_1} + (1-t)\mathbf{x_2}\right) \leq t\, f(\mathbf{x_1}) + (1-t)\, f(\mathbf{x_2}) - \frac{1}{2} mt(1-t)\|\mathbf{x_1} - \mathbf{x_2}\|_2^2. \tag{2.4}$$

Correspondingly, when $m = 0$, f is convex such that

$$f\left(t\mathbf{x_1} + (1-t)\mathbf{x_2}\right) \leq t\, f(\mathbf{x_1}) + (1-t)\, f(\mathbf{x_2}). \tag{2.5}$$

A strongly convex function is also strictly convex, but not vice versa.

Suppose that a constrained convex optimization problem is given as,

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^N} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{y}. \end{aligned}$$

The augmented Lagrangian method uses the Lagrangian of the constrained problem and an additional penalty term (the augmentation) to give the dual function:

$$\begin{aligned} g_\rho(\lambda) &= \inf_{\mathbf{x}\in\mathbb{R}^N} L_\rho(\mathbf{x}, \lambda) \\ &= \inf_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + \lambda^T(\mathbf{A}\mathbf{x} - \mathbf{y}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2, \end{aligned} \tag{2.6}$$

where $\mathbf{A} \in \mathbb{R}^{m\times N}$, $\lambda \in \mathbb{R}^m$ is the Lagrangian multiplier, and $\rho > 0$ is the penalty parameter.

With the quadratic penalty term, there is only one minimizer of $L_\rho(\mathbf{x}, \lambda)$ given f is convex. Therefore, we can recover a primal optimal point $\mathbf{x}^*$ from a dual optimal point λ^* as

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathbb{R}^N} L_\rho(\mathbf{x}, \lambda^*).$$

We solve the dual problem using gradient ascent [16]. Assuming that g is differentiable, the gradient $\nabla g(\lambda)$ can be evaluated as follows. First, we find $\dot{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^N} L_\rho(\mathbf{x}, \lambda)$. Then, we have $\nabla g(\lambda) = \mathbf{A}\dot{\mathbf{x}} - \mathbf{y}$, which is the residual for the equality constraint. The corresponding iterative steps are as:

$$\begin{aligned} \mathbf{x}^{k+1} :&= \arg\min_{\mathbf{x}\in\mathbb{R}^N} L_\rho(\mathbf{x}, \lambda^k) \\ &= \arg\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + (\lambda^k)^T(\mathbf{A}\mathbf{x} - \mathbf{y}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2, \end{aligned} \tag{2.7}$$

$$\lambda^{k+1} := \lambda^k + \rho\,(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y}). \tag{2.8}$$

The first step (2.7) is an $\mathbf{x}$-minimization step, and the second step (2.8) is a dual variable update. This method can be used when g is not differentiable [40].

To be more convenient, we reformulate (2.7) as the scaled form [16], f plus a quadratic term.

$$\mathbf{x}^{k+1} := \arg\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + \frac{\rho}{2}\,\|\mathbf{A}\mathbf{x} - \mathbf{y} + \frac{1}{\rho}\lambda^k\|_2^2. \tag{2.9}$$

(2.7) and (2.9) are equivalent, because

$$\begin{aligned} \mathbf{x}^{k+1} :&= \arg\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + (\lambda^k)^T(\mathbf{A}\mathbf{x} - \mathbf{y}) + \frac{\rho}{2}\,\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \\ &= \arg\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + \frac{\rho}{2}\,\|\mathbf{A}\mathbf{x} - \mathbf{y} + \frac{1}{\rho}\,\lambda^k\|_2^2 - \frac{1}{2\rho}\,\|\lambda^k\|_2^2. \end{aligned}$$

Discard the constant term $-\frac{1}{2\rho}\,\|\lambda^k\|_2^2$, then we have

$$\mathbf{x}^{k+1} := \arg\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + \frac{\rho}{2}\,\|\mathbf{A}\mathbf{x} - \mathbf{y} + \frac{1}{\rho}\lambda^k\|_2^2.$$

2.5 Truncated Nuclear Norm Regularization in Matrix Completion

Matrix completion is to estimate the missing values in a matrix from a small subset of its entries with the assumption that the matrix is low-rank. Low-rank matrix completion constrains the number of nonzero singular values. It is very relevant to the sparsity in CS. Similar to using the ℓ_1 minimization as the convex relaxation of the original ℓ_0 minimization in CS, low-rank matrix completion uses the nuclear norm regularization, which sums all the singular values together, as

the convex relaxation of the rank regularization. The work in [4] proposed a novel method called truncated nuclear norm regularization (TNNR) to solve the matrix completion problem. Unlike the traditional nuclear norm approaches, which consider all the singular values, TNNR sets the r largest singular values free and only sums up the rest, where r is the matrix rank. This idea and our proposed *truncated ℓ_1 minimization* (see details in Chapter 3) are very similar. As TNNR is no longer convex, by introducing an interim variable, they designed a two-step iterative scheme to provide a tractable process which guarantees the convergence to a local minimum. The experimental results on both synthetic and real visual datasets demonstrate the advantages of the TNNR-based algorithms in comparison with other state-of-art matrix completion methods based on the nuclear norm. *Our framework for solving truncated ℓ_1 minimization problems was inspired by this work.*

2.6 The Sparsity Model in Statistical Learning Theory

In statistical learning theory, the sparsity models are used as a method for avoiding overfitting, e.g., the sparsity assumption embedded in Bayesian variable selections. This view is especially well illustrated by *spike-and-slab variable selection*. Although CS and the spike-and-slab variable selection share a similar model and assumptions, they respectively use different approaches to search for sparse solutions. The sparse solution in CS is achieved by solving a constrained optimization problem, while the one in the spike-and-slab variable selection is achieved through a stochastic process, such as Markov chain Monte Carlo (MCMC), to draw samples from the model space in which models with high posterior probability will be visited more often than low probability models [17].

For interested readers, we give some details of the spike-and-slab framework. Consider a regression model given as

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon,$$

where $\mathbf{y} \in \mathbb{R}^m$ are independent responses, with each corresponding to a N-dimensional variable $X_i\,(\,i \in 1, \ldots, m\,)$, and the error term $\varepsilon \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$, $\sigma \in \mathbb{R}^m$. Regression through variable selection is firstly to identify a subset Ω from the N predictors that fits the data well, then secondly to get the final regression coefficients for predictors in Ω while the others are not used in the regression model. To identify the subset Ω, one way is to use a frequentist approach, the other is to use the posterior mean. The *spike-and-slab* framework can be viewed as a mask

Factor	CS	Spike-and-Slab	Similar
Model	$\mathbf{y} = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \mathbb{R}^{m \times N}$, $\mathbf{y} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^N$	$E(\mathbf{y}) = \mathbf{X}\beta$, $\mathbf{X} \in \mathbb{R}^{m \times N}$, $\mathbf{y} \in \mathbb{R}^m, \beta \in \mathbb{R}^N$	yes
Matrix Generation	$\mathbf{A}$ is randomly generated	$\mathbf{X}$ is randomly sampled	yes
Sparsity	nonzero values of $\mathbf{x}$ is sparse	nonzero values/coefficients of β is sparse	yes
Spike	the value of nonzero x_i	the distribution of zero β_k	no
Matrix Shape	$m \ll n$	$m < n$ is not necessary	no
Vector Reconstruction	Convex Optimization	Stochastic Process	no

Table 2.1: The comparison between CS and the Spike-and-Slab variable selection. By involving mechanisms of sparsity and randomness, both model problems use different strategies to acquire the sparse solution.

on β, in which spikes correspond to predictors with zero coefficients and slabs correspond to predictors with nonzero coefficients. We use the simple two-component prior [18] to explain how it works. The prior hierarchical distribution is given as follows:

$$(\beta_k|\zeta_k) \overset{ind}{\sim} (1-\zeta_k)\mathcal{N}(0, \tau_k^2) + \zeta_k \mathcal{N}(0, c_k\tau_k^2),\ \forall k \in \{1, N\},$$

$$\zeta_k \overset{ind}{\sim} \text{Bern}(\omega_k),\ 0 < \omega_k < 1,$$

where ζ_k is a posterior latent variable defined as

$$\zeta_k = \begin{cases} 1, & \text{if the } k^{th} \text{ predictor is used in the model} \\ 0, & \text{otherwise .} \end{cases}$$

The value of τ_k^2 is chosen to be small and consequently, the shape of $\mathcal{N}(0, \tau_k^2)$ corresponding to a predictor with $\zeta_k = 0$ exhibits as a spike. On the other hand, $c_k > 0$ is chosen to be large, thus the shape of $\mathcal{N}(0, c_k\tau_k^2)$ corresponding to a predictor with $\zeta_k = 1$ shows as a slab. Based on such priors, predictors with promising coefficients will have large posterior β_k values. In practice, it is difficult to manually set values for ω_k, τ_k^2, and c_k. Improperly chosen values lead to models that concentrate on either too few or too many coefficients. To overcome this problem, Ishwaran and Rao [19] proposed a continuous bimodal prior in place of the two-component mixture distribution. It is a more sophisticated hierarchy distribution, detailed in [18].

Chapter 3

Truncated ℓ_1 Minimization

3.1 The optimization problem and Reformulation

Compared to $\|\mathbf{x}\|_0$, in which all the nonzero entries have equal contributions, $\|\mathbf{x}\|_1$ treats the entries differently by summing their absolute values. We propose the truncated ℓ_1 norm, $\|\mathrm{Trun}_r(\mathbf{x})\|_1$, as a better approximation of $\|\mathbf{x}\|_0$ by subtracting the r largest absolute values from $\|\mathbf{x}\|_1$.

Definition 3.1.1. Given a vector $\mathbf{x} \in \mathbb{R}^N$, the *truncated ℓ_1-norm* with order r, $\|\mathrm{Trun}_r(\mathbf{x})\|_1$, is defined as the sum of $N - r$ minimum absolute entries in $\mathbf{x}$, that is

$$\|\mathrm{Trun}_r(\mathbf{x})\|_1 := \sum_{i=r+1}^{N} |\mathbf{x}|^*,$$

where the vector $|\mathbf{x}|^*$ is the nonincreasing rearrangment of $\mathbf{x}$, sorted by the absolute value of each entry.

The corresponding optimization problem, called *truncated ℓ_1 minimization*(TLM), is formulated as follows:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^N} \quad & \|\mathrm{Trun}_r(\mathbf{x})\|_1 \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{y}. \end{aligned} \tag{3.1}$$

Following this optimization problem, the r largest entries do not need to compete with entries in other directions to fulfill the objective because their values do not appear in the objective. Particularly, when $r = s$, the implication behind TLM is first to detect the support and then to minimize all the rest to be 0. The support detection part is similar to counting the number

of nonzeros in calculating $\|\mathbf{x}\|_0$.

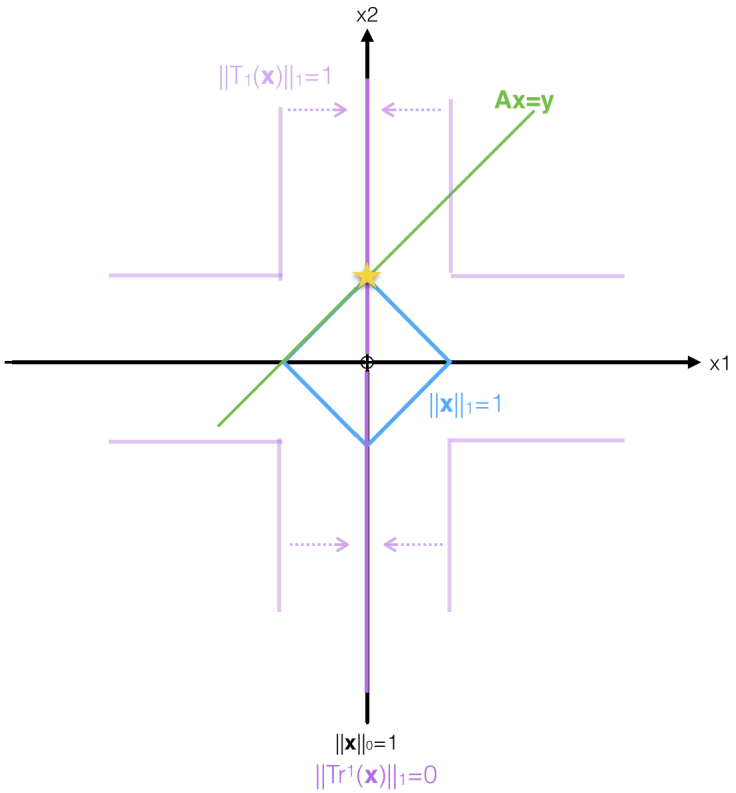


Figure 3.1: In the geometric view, the truncated ℓ_1 norm is a better approximation of the ℓ_0 norm than the ℓ_1 norm. As shown in the plot, the vector set corresponding to $\|\text{Trun}_1(\mathbf{x})\|_1 = 0$ is the union of the zero vector and the vector set corresponding to $\|\mathbf{x}\|_0 = 1$.

Technically speaking, the truncated ℓ_1 function is not a norm or seminorm [20], because $\|\text{Trun}_r(.)\|$ fails to satisfy the triangle inequality.

Definition 3.1.2. Given a vector space $\mathbf{V}$, then a function $f : V \mapsto \mathbb{R}$ satisfies the *triangle inequality* if and only if

$$f(\mathbf{x}_1 + \mathbf{x}_2) \leq f(\mathbf{x}_1) + f(\mathbf{x}_2), \forall\, \mathbf{x}_1, \mathbf{x}_2 \in \mathbf{V}.$$

Here is a counterexample. Given $\mathbf{x}_1 = (1, 2, 3)^T$ and $\mathbf{x}_2 = (2, 5, 2)^T$, then $\|\text{Trun}_1(\mathbf{x}_1)\| = 3$, $\|\text{Trun}_1(\mathbf{x}_2)\| = 4$, and $\|\text{Trun}_1(\mathbf{x}_1 + \mathbf{x}_2)\| = 8$. Obviously, $\|\text{Trun}_1(\mathbf{x}_1 + \mathbf{x}_2)\| \nleq \|\text{Trun}_1(\mathbf{x}_1)\| + \|\text{Trun}_1(\mathbf{x}_2)\|$.

Note, $\|\text{Trun}_r(.)\|$ is not convex (see Definition 2.4.1), because the indices of the locations of the r largest entries is not fixed. Recalling that given $\mathbf{x}_1$ and $\mathbf{x}_2$ in the above example, we set $t = 0.5$, obviously, $\|\text{Trun}_1(0.5\,\mathbf{x}_1 + 0.5\,\mathbf{x}_2)\| \nleq 0.5\,\|\text{Trun}_1(\mathbf{x}_1)\| + 0.5\,\|\text{Trun}_1(\mathbf{x}_2)\|$.

As the truncated ℓ_1 norm is nonconvex, we cannot use the augmented Lagrangian method directly to solve (3.1). Before we introduce an interim variable to reformulate (3.1) as a tractable procedure, we provide the following theorem.

Theorem 3.1.1. *Let $\Omega_{\mathbf{v}}$ denote the index set corresponding to the r largest absolute entries of $\mathbf{v}$, and $\mathcal{D}^r(\mathbf{v})$ be an diagonal matrix defined as, $\forall i \in \{1, \ldots, N\}$,*

$$\mathcal{D}^r_{ii} = \begin{cases} 1, & \text{if } i \in \Omega_{\mathbf{v}} \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

Given two vectors $\mathbf{x}, \mathbf{z} \in \mathbb{R}^N$ and $r \leq N$, then

$$\|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1 \leq \sum_{i \in \Omega_{\mathbf{x}}} |x_i|. \tag{3.3}$$

Proof. According to (3.2), we have

$$\|\mathcal{D}^r(\mathbf{x})\,\mathbf{x}\|_1 = \sum_{i \in \Omega_{\mathbf{x}}} |x_i|, \qquad \|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1 = \sum_{i \in \Omega_{\mathbf{z}}} |x_i|. \tag{3.4}$$

Let $\Omega^* = \Omega_{\mathbf{z}} \cap \Omega_{\mathbf{x}}$, then $|\Omega_{\mathbf{z}} \setminus \Omega^*| = |\Omega_{\mathbf{z}} \setminus \Omega^*|$, and

$$\min_{i \in \Omega_{\mathbf{x}} \setminus \Omega^*} |x_i| \geq \max_{i \in \Omega_{\mathbf{z}} \setminus \Omega^*} |z_i|. \tag{3.5}$$

Rewrite (3.4) as follows:

$$\begin{aligned} \|\mathcal{D}^r(\mathbf{x})\,\mathbf{x}\|_1 &= \sum_{i \in \Omega_{\mathbf{x}}} |x_i| = \sum_{i \in \Omega^*} |x_i| \; + \sum_{i \in \Omega_{\mathbf{x}} \setminus \Omega^*} |x_i|, \\ \|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1 &= \sum_{i \in \Omega_{\mathbf{z}}} |x_i| = \sum_{i \in \Omega^*} |x_i| \; + \sum_{i \in \Omega_{\mathbf{z}} \setminus \Omega^*} |x_i|. \end{aligned}$$

To compare $\|\mathcal{D}^r(\mathbf{x})\,\mathbf{x}\|_1$ with $\|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1$ is equivalent to the comparison between $\sum_{i \in \Omega_{\mathbf{x}} \setminus \Omega^*} |x_i|$ and $\sum_{i \in \Omega_{\mathbf{z}} \setminus \Omega^*} |x_i|$.

If $\Omega_{\mathbf{z}} \setminus \Omega^* = \Omega_{\mathbf{x}} \setminus \Omega^* = \varnothing$, then

$$\sum_{i \in \Omega_{\mathbf{x}} \setminus \Omega^*} |x_i| \; = \sum_{i \in \Omega_{\mathbf{z}} \setminus \Omega^*} |x_i| = 0.$$

Otherwise,

$$\sum_{i\in\Omega_{\mathbf{z}}\setminus\Omega^*}|x_i|\leq\sum_{i\in\Omega_{\mathbf{x}}\setminus\Omega^*}|x_i|,$$

verified directly by (3.5).

The claims (3.3) now follows. □

Based on Definition 3.1.1,

$$\begin{aligned}\|\mathrm{Trun}_r(\mathbf{x})\|_1 :&= \sum_{i=r+1}^{N}|\mathbf{x}|^*\\ &= \|\mathbf{x}\|_1-\sum_{i\in\Omega_{\mathbf{x}}}|x_i|\\ &= \|\mathbf{x}\|_1-\|\mathcal{D}^r(\mathbf{x})\,\mathbf{x}\|_1,\end{aligned}$$

according to Theorem 3.1.1, $\|\mathcal{D}^r(\mathbf{x})\,\mathbf{x}\|_1=\sup_{\mathbf{z}\in\mathbb{R}^N}\|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1$, then we have

$$\|\mathrm{Trun}_r(\mathbf{x})\|_1:=\|\mathbf{x}\|_1-\sup_{\mathbf{z}\in\mathbb{R}^N}\|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1.$$

Accordingly, (3.1) can be rewritten as follows:

$$\begin{aligned}&\min_{\mathbf{x}\in\mathbb{R}^N}\|\mathbf{x}\|_1-\max_{\mathbf{z}\in\mathbb{R}^N}\|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1\\ &\text{subject to}\quad \mathbf{A}\mathbf{x}=\mathbf{y}.\end{aligned}\tag{3.6}$$

With fixed $\hat{\mathbf{x}}$, the optimal solution of (3.6) is when $\mathbf{z}=\hat{\mathbf{x}}$. Instead of considering (3.6) as a constrained optimization problem with two separable variables, we treat $\mathbf{z}$ as an interim variable and enforce it to be an approximation of $\mathbf{x}$. Based on that, we design a two-step iterative scheme, inspired by the algorithmic framework in [4], to solve (3.6). We call the solver for (3.6) TLM-2SI (see details in Algorithm 1). In the l-th iteration, we firstly fix $\mathbf{x}^l$ and compute $\mathcal{D}^r=\mathcal{D}^r(\mathbf{x}^l)$. Then we fix $\mathcal{D}^r$ to update $\mathbf{x}^{l+1}$ by solving the following problem,

$$\begin{aligned}&\min_{\mathbf{x}\in\mathbb{R}^N}\|\mathbf{x}\|_1-\|\mathcal{D}^r\,\mathbf{x}\|_1\\ &\text{subject to}\quad \mathbf{A}\mathbf{x}=\mathbf{y}.\end{aligned}\tag{3.7}$$

The objective function in (3.7) is convex, because

$$F(\mathbf{x}) = \|\mathbf{x}\|_1 - \|\mathcal{D}^r \mathbf{x}\|_1 = \sum_{i \in \Omega_{\mathbf{x}}^c} |x_i|.$$

$F(\mathbf{x})$ is a linear combination of $|x_i|$ with nonnegative coefficients, and $|x_i|$ is a convex function. Moreover, a sum of convex functions is convex. Thus, $F(\mathbf{x})$ is convex given by straightforward verification of Definition 2.4.1. Correspondingly, we can use the augmented Lagrangian method to transform (3.7), a convex constrained optimization problem, into an unconstrained strongly convex function which guarantees convergence. The solver for (3.7) (the step 2 in TLM-2SI) is called PLM-ALM. We will detail this solver in the next chapter.

Algorithm 1 Truncated ℓ_1 Minimization via two-step Iterative (TLM-2SI)

Input: the measurement matrix $\mathbf{A} \in \mathbb{R}^{m \times N}$, the measurements $\mathbf{y} \in \mathbb{R}^m$, the order of the truncated norm $r \in \mathbb{Z}^+$, and the tolerance ϵ
Initialize: randomly pick up a vector $\mathbf{x}_0 \in \mathbb{R}^N$
repeat
STEP 1. Given $\mathbf{x}^l$, compute $\mathcal{D}^r = \mathcal{D}^r(\mathbf{x}^l)$
STEP 2. Solve

$$\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_1 - \|\mathcal{D}^r \mathbf{x}\|_1$$
$$\text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}$$

to update $\mathbf{x}^{l+1}$
until $\|\mathbf{x}^{l+1} - \mathbf{x}^l\|_\infty \leq \epsilon$
return the reconstructed vector $\mathbf{x}$

Here, $\|\mathbf{x}\|_\infty$ is the infinite norm of $\mathbf{x}$ and defined as $\|\mathbf{x}\|_\infty := \max_i |x_i|$.

3.2 The Order of Truncated ℓ_1 Norm in TLM

In TLM-2SI, we solve (3.6) iteratively given the order r of the truncated ℓ_1 norm. When $r \geq s$, obviously $\|\text{Trun}_r(\mathbf{x})\| = 0$. The implicit assumption in TLM is that we know s. However, in practice, we do not exactly know the sparsity of the signal, therefore we use a brute-force way to get a heuristic about the choice of r in TLM. It is simply an iterative process with built-in TLM-2SI. However, the stopping criterion for this naive searching strategy is ambiguous. First, the convergence threshold is not suitable as each TLM-2SI with fixed r only guarantees to return a local minimum. Second, achieving a *zero* value for the objective is too stringent to identify the exact reconstruction, because it may happen when $r < s$. Even worse is to set the maximum for r to be N. There would be too much wasted computation, because usually N is

Sparsity	**r=0**	**r=1**	**r=2**	**r=3**	**r=4**	**r=5**
3	100%	100%	90%	80%	50%	20%
4	100%	0	100%	80%	60%	50%
5	0	20%	30%	30%	20%	10%

Table 3.1: The rate of exact reconstruction with 10 different initialization ($\mathbf{x}^0$) by using TLM-2SI. In this toy example, $\mathbf{A} \in \mathbb{R}^{10\times 30}$ is a Gaussian matrix. The interested signals with sparsity equal to 3, 4, and 5, respectively. They are randomly generated under an uniform distribution. TLM-2SI with $r = 0$ is equivalent to solving the convex ℓ_1 norm minimization. The problem (3.6) is nonconvex and the TLM-2SI only guarantees the local minimum, the reconstruction from TLM is sensitive to initializations, while the reconstruction through ℓ_1 norm minimization only depends on the structure of $\mathbf{A}$ and the interested signal. However, TLM shows the tolerance to tight conditions when ℓ_1 norm minimization fails. And TLM-2SI with $r < s$ performs not worse than the one with $r = s$.

very large and $r \ll N$. However, we've observed that there is improvement by removing some of entries in the support from the objective.

Moreover, as TLM is not convex and thus only the local minimum can be guaranteed, to identify if TLM-2SI can reconstruct a sparse signal is time-consuming. We define *success* as being able to reconstruct the signal of interest through multiple trials with different initializations. As the sparsity s increases, the number of trials becomes larger to identify whether it is a success. Compared with ℓ_1 minimization, truncated ℓ_1 minimization is unreliable because

$$p_{\ell_1}(\text{success}|\mathbf{A},\mathbf{x}) \in \{0,1\}, \quad \text{while } p_{T\ell_1}(\text{success}|\mathbf{A},\mathbf{x}) \in [0,1],$$

where given $\mathbf{A}$ and $\mathbf{x}$, $p_{\ell_1}(.)$ denote the probability of successful reconstruction through ℓ_1 minimization, and $p_{T\ell_1}(.)$ denote the probability of successful reconstruction through truncated ℓ_1 minimization.

Based on this observation, we propose another optimization problem, *partial* ℓ_1 *minimization (PLM)* , which leverages prior information about *supp*($\mathbf{u}$) to obtain a convex optimization problem.

Chapter 4

Partial ℓ_1 Minimization

4.1 The optimization problem

Suppose that an oracle provides information on which d entries are nonzero, $d \leq s$. Instead of trying to figure out if such entries are the largest, we just effortlessly give priority to them by removing them from the objective and minimizing the sum of the remaining absolute values of the entries, that is the *partial ℓ_1 norm*.

Definition 4.1.1. Given a vector $\mathbf{x} \in \mathbb{R}^N$, and $\mathbb{J}$ is the support of $\mathbf{x}$, let $\Omega_{\mathcal{P}}$ be an arbitrary subset of $\mathbb{J}$, and $\overline{\Omega_{\mathcal{P}}}$ be the complement of $\Omega_{\mathcal{P}}$ in $\{1,\ldots,\mathrm{N}\}$. The *partial ℓ_1 norm* on $\overline{\Omega_{\mathcal{P}}}$, $\|\mathbf{x}_{\overline{\Omega_{\mathcal{P}}}}\|_1$, is defined as the sum of absolute value of each entry in $\overline{\Omega_{\mathcal{P}}}$.

$$\begin{aligned}\|\mathbf{x}_{\overline{\Omega_{\mathcal{P}}}}\|_1 &:= \sum_{i\in\overline{\Omega_{\mathcal{P}}}} |x_i| \\ &= \|\mathbf{x}\|_1 - \|\mathbf{x}_{\Omega_{\mathcal{P}}}\|_1.\end{aligned}$$

The corresponding optimization problem, called *partial ℓ_1 minimization*, is a derivative of (3.7), and thus is formulated as follows:

$$\begin{aligned}\min_{\mathbf{x}\in\mathbb{R}^N} \quad & \|\mathbf{x}\|_1 - \|\mathcal{P}\mathbf{x}\|_1 \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{y},\end{aligned} \tag{4.1}$$

where $\mathcal{P} \in \mathbb{R}^{N\times N}$ is a diagonal matrix with

$$\mathcal{P}_{ii} = \begin{cases} 1, & \text{if } i \in \Omega_{\mathcal{P}} \\ 0, & otherwise \end{cases}, \quad \forall\, i \in \{1, \ldots, N\}. \tag{4.2}$$

PLM differs from TLM in several important ways. First, PLM makes use of appropriate information about $\mathbb{J}$ and thus is formulated as a convex problem. Second, TLM needs to know the cardinality of $\mathbb{J}$. This is a strong assumption. The prior knowledge assumption in PLM is less strong, as explained in Figure 1.1. In particular, we do not require that the oracle is perfectly accurate! Actually, the oracle is allowed to provide indices of $\mathbf{u}$ which have *no role* in the support. And important indices of $\mathbf{u}$, which correspond to large absolute value, are also allowed *not to appear in the subset* provided by the oracle. All that we require is that the subset provided by the oracle overlaps in some way with $\mathbb{J}$.

4.2 The Optimization Solver Using The Augmented Lagrangian Method

We use the augmented Lagrangian method to solve (4.1) as a strongly convex function with the global minimum guaranteed. The augmented Lagrange function of (4.1) is

$$L_\rho(\mathbf{x}, \lambda) = \|\mathbf{x}\|_1 - \|\mathcal{P}\mathbf{x}\|_1 + \lambda^T(\mathbf{A}\mathbf{x} - \mathbf{y}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2. \tag{4.3}$$

Finding the optimal $\mathbf{x}$ in (4.1) is equivalent to minimizing the augmented Lagrangian function $L_\rho(\mathbf{x}, \lambda)$, by computing $\mathbf{x}^{k+1}$ and the Lagrangian multiplier λ^{k+1} alternatively.

Computing $\mathbf{x}^{k+1}$.

Fix λ^k and minimize $L_\rho(\mathbf{x}, \lambda^k)$ as follows:

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}\in\mathbb{R}^N} [\, \|\mathbf{x}\|_1 - \|\mathcal{P}\mathbf{x}\|_1 + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y} + \frac{1}{\rho}\lambda^k\|_2^2 \,]. \tag{4.4}$$

Because (4.4) is strictly convex, there is a unique minimizer. However, the nonsmooth ℓ_1 norm function is nondifferentiable with respect to a coordinate which is zero. Therefore, it is not possible to obtain a closed-form solution for (4.4) [22]. As the quadratic term is smooth and has a Lipschitz continuous gradient, there are still multiple approaches, such

as *iterative shrinkage-threshold (IST)* method [21], *fast iterative shrinkage-threshold (FIST)* method [21], and *proximal gradient (PG)* method ([23], [24]), to solve (4.4) efficiently with theoretical guarantees. Here, we use an approximate and straightforward approach to formulate a sequential, point-wise solution for (4.4).

A point $\mathbf{x}^*$ is the minimizer of a convex function f if and only if f is subdifferentiable at $\mathbf{x}^*$ and

$$\mathbf{0} \in \partial f(\mathbf{x}^*).$$

This condition reduces to $\nabla f(\mathbf{x}^*) = \mathbf{0}$, when f is differentiable at $\mathbf{x}^*$ [27].

We define the subgradient of the ℓ_1 norm defined as

$$\frac{\partial \|\mathbf{x}\|_1}{\partial x_i} = \begin{cases} 1, & x_i = 0 \\ u_i \in [-1, 1], & x_i = 0 \\ -1, & x_i < 0. \end{cases} \tag{4.5}$$

Because $L_\rho(\mathbf{x}, \lambda^k)$ is subdifferentiable, the condition for the unique minimizer is

$$\mathbf{0} \in \partial L_\rho(\mathbf{x}^{k+1}, \lambda^k).$$

Based on matrix differentiation,

$$\begin{aligned} \partial\, L_\rho(\mathbf{x}^{k+1}, \lambda^k) &= \partial\, (\|\mathbf{x}^{k+1}\|_1 - \|\mathcal{P}\mathbf{x}^{k+1}\|_1) + \partial\, \frac{\rho}{2}\, \|\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y} + \frac{1}{\rho}\lambda^k\|_2^2 \\ &= \partial\, (\|\mathbf{x}^{k+1}\|_1 - \|\mathcal{P}\mathbf{x}^{k+1}\|_1) + \rho\, \mathbf{A}^T(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y} + \frac{1}{\rho}\lambda^k). \end{aligned}$$

We use $\mathbf{V}$ to denote $\partial\, L_\rho(\mathbf{x}^{k+1}, \lambda^k)$, which is a vector in $\mathbb{R}^N$. Because $\|.\|_1$ is a separable function, each entry v_i of $\mathbf{V}$ can be expressed as, $\forall\, i \in \{1, \ldots, N\}$

$$\begin{aligned} v_i &= \partial\, (|x_i^{k+1}| - |\mathcal{P}_{ii} x_i^{k+1}|) + [\rho\, \mathbf{A}^T(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y} + \frac{1}{\rho}\lambda^k)]_i \\ &= \partial\, |x_i^{k+1}| - \mathcal{P}_{ii}\, \partial\, |x_i^{k+1}| - \rho\, [\mathbf{A}^T\mathbf{y}]_i + [\mathbf{A}^T\lambda^k]_i + \rho \sum_{j=1}^{N} [\mathbf{A}^T\mathbf{A}]_{ij}\, x_j^{k+1}, \end{aligned} \tag{4.6}$$

and $0 \in v_i$.

If $x_i^{k+1} \neq 0$, then $|x_i|$ is differentiable at the point x_i^{k+1}, and consequently, $v_i = 0$.

From $v_i = 0$ with the assumption that $x_i^{k+1} \neq 0$, we have

$$\partial\,|x_i^{k+1}| - \mathcal{P}_{ii}\,\partial\,|x_i^{k+1}| - \rho\,[\mathbf{A}^T\mathbf{y}]_i + [\mathbf{A}^T\lambda^k]_i + \rho\sum_{j=1}^{N}[\mathbf{A}^T\mathbf{A}]_{ij}\,x_j^{k+1} = 0,$$

therefore,

$$\begin{aligned}\rho\sum_{j=1}^{N}[\mathbf{A}^T\mathbf{A}]_{ij}\,x_j^{k+1} &= -\partial\,|x_i^{k+1}| + \mathcal{P}_{ii}\,\partial\,|x_i^{k+1}| + \rho\,[\mathbf{A}^T\mathbf{y}]_i - [\mathbf{A}^T\lambda^k]_i \\ &= -\mathbf{sgn}(|x_i^{k+1}|) + \mathcal{P}_{ii}\,\mathbf{sgn}(|x_i^{k+1}|) + \rho\,[\mathbf{A}^T\mathbf{y}]_i - [\mathbf{A}^T\lambda^k]_i.\end{aligned}$$

This equation yields

$$x_i^{k+1} = \frac{[\mathbf{A}^T(\rho\,\mathbf{y} - \lambda^k)]_i - \rho\,\sum_{j=1,\,j\neq i}^{N}[\mathbf{A}^T\mathbf{A}]_{ij}\,x_j^{k+1} - \mathbf{sgn}(|x_i^{k+1}|) + \mathcal{P}_{ii}\,\mathbf{sgn}(|x_i^{k+1}|)}{\rho\,[\mathbf{A}^T\mathbf{A}]_{ii}}, \tag{4.7}$$

and $x_i^{k+1} \neq 0$.

For unknown x_j^{k+1} on the right side of the equation (4.7), we use x_j^k as its approximation. For simplification, we take off the superscript of x_j, and rewrite (4.7) as

$$\begin{aligned}x_i^{k+1} &= \frac{[\mathbf{A}^T(\rho\,\mathbf{y} - \lambda^k)]_i - \rho\,\sum_{j=1,\,j\neq i}^{N}[\mathbf{A}^T\mathbf{A}]_{ij}\,x_j - \mathbf{sgn}(|x_i^{k+1}|) + \mathcal{P}_{ii}\,\mathbf{sgn}(|x_i^{k+1}|)}{\rho\,[\mathbf{A}^T\mathbf{A}]_{ii}} \\ &= \mathbf{G}_i\,(x_1^{k+1},\ldots,x_{i-1}^{k+1},x_{i+1}^{k},\ldots,x_N^{k}) + \mathbf{F}_i(x_i^{k+1}),\end{aligned} \tag{4.8}$$

where

$$\mathbf{G}_i\,(x_1^{k+1},\ldots,x_{i-1}^{k+1},x_{i+1}^{k},\ldots,x_N^{k}) = \frac{[\mathbf{A}^T(\rho\,\mathbf{y} - \lambda^k)]_i - \rho\,\sum_{j=1,\,j\neq i}^{N}[\mathbf{A}^T\mathbf{A}]_{ij}\,x_j}{\rho\,[\mathbf{A}^T\mathbf{A}]_{ii}} = \mathbf{g}_i,$$

and

$$\mathbf{F}_i(x_i^{k+1}) = \frac{-\mathbf{sgn}(|x_i^{k+1}|) + \mathcal{P}_{ii}\,\mathbf{sgn}(|x_i^{k+1}|)}{\rho\,[\mathbf{A}^T\mathbf{A}]_{ii}} = \begin{cases}\frac{-1+\mathcal{P}_{ii}}{\rho\,[\mathbf{A}^T\mathbf{A}]_{ii}} = \mathbf{f}_i^+, & x_i^{k+1} > 0 \\ \frac{1-\mathcal{P}_{ii}}{\rho\,[\mathbf{A}^T\mathbf{A}]_{ii}} = \mathbf{f}_i^-, & x_i^{k+1} < 0.\end{cases}$$

$\mathbf{F}_i(x_i^{k+1})$ indicates that (4.8) has self-constraints to keep consistent with the sign of x_i^{k+1}.

If $|x_i| = 0$, then $|x_i|$ is not differentiable at the point x_i^{k+1}, and consequently, (4.8) would be self-contradictory.

We formulate the sequential, point-wise solution for (4.4) as follows:

$$x_i^{k+1} = \begin{cases} \mathbf{g}_i + \mathbf{f}_i^+, & \mathbf{g}_i + \mathbf{f}_i^+ > 0 \\ \mathbf{g}_i + \mathbf{f}_i^-, & \mathbf{g}_i + \mathbf{f}_i^- < 0 \\ 0, & otherwise \end{cases}, \quad \forall\, i \in \{1, \ldots, N\}. \tag{4.9}$$

Computing λ^{k+1}.

Fix $\mathbf{x}^{k+1}$, and calculate λ^{k+1} as follows:

$$\lambda^{k+1} = \lambda^k + \rho(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y}). \tag{4.10}$$

The whole procedure of solving (4.1) is summarized in Algorithm 2, called PLM-ALM. The main computational cost of PLM-ALM in each iteration is the computation of point-wise solution for $\mathbf{x}^{k+1}$ in step 1. The linear convergence of PLM-ALM is guaranteed by the augmented Lagrangian method [26].

Algorithm 2 Partial ℓ_1 Minimization using The Augmented Lagrangian Method (PLM-ALM)

Input: the measurement matrix $\mathbf{A} \in \mathbb{R}^{m \times N}$, the measurements $\mathbf{y} \in \mathbb{R}^m$, the subset $\Omega_{\mathcal{P}}$, the penalty parameter ρ, the tolerance ϵ

Initialize: randomly pick up a vector $\mathbf{x}_0 \in \mathbb{R}^N$, and set $\lambda_0 = \mathbf{0}$

repeat

step 1, update $\mathbf{x}^{k+1}$

$$x_i^{k+1} = \begin{cases} \mathbf{g}_i + \mathbf{f}_i^+ & , \text{if } \mathbf{g}_i + \mathbf{f}_i^+ > 0 \\ \mathbf{g}_i + \mathbf{f}_I^- & , \text{if } \mathbf{g}_i + \mathbf{f}_I^- < 0 \\ 0 & , otherwise \end{cases}, \quad \forall\, i \in \{1, \ldots, N\}$$

step 2, update λ^{k+1}

$$\lambda^{k+1} = \lambda^k + \rho\,(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y})$$

until $\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_\infty \leq \epsilon$

return the recovered vector $\mathbf{x}$

Chapter 5

Experimental Results

In this section, we investigate the practical ability of PLM to reconstruct sparse vectors from less/non-RIP measurement matrices. The goal here is to evaluate empirically the breakpoint of PLM so as to get a sense of the performance one might expect in practice. To do this, we perform a series of experiments on synthetic data to: 1) check whether PLM is able to reconstruct $\mathbf{u}$ from $\mathbf{y}$ when $\mathbf{A}$ is less/non-RIP; 2) compare the performance of PLM with ℓ_1 minimization; 3) check whether PLM prefers an $\Omega_{\mathcal{P}}$ including indices corresponding to larger entries. Furthermore, we test PLM on real visual data.

The main findings are that: 1) PLM is better than ℓ_1 minimization to reconstruct sparse vectors from fewer measurements; 2) when a subset $\mathcal{C}$ of column vectors in the measurement matrix is enforced to get close to each other, even sometimes to be identical, PLM has much higher probability than ℓ_1 minimization to reconstruct sparse vectors from such corrupted systems, as long as $\Omega_{\mathcal{P}}$ overlaps in some way with $\mathcal{C}$; 3) to make improvements on ℓ_1 minimization, all that PLM requires is $\Omega_{\mathcal{P}}$ overlaps with the support of the signal. It is not sensitive to how large the entries with indices in $\Omega_{\mathcal{P}}$, and is tolerant with false alarms, which means spurious indices are included in $\Omega_{\mathcal{P}}$. The test on the reconstruction of real visual data shows that PLM is able to reconstruct an image from the corrupted measurement matrix, when ℓ_1 minimization fails.

For ℓ_1 minimization, CVXPY [27], a Python embedded package, is applied. For PLM, our proposed PLM-ALM is applied. The experiments are conducted using Python.

5.1 Synthetic data

The focus of our experiments on synthetic data is to explore how PLM performs covering a wide range of less/non-RIP conditions. To be precise, there are two ways to cause a random matrix being less/non-RIP. One way is to use fewer number of measurements, according to Section 2.1.2. The other way is that some column vectors are close to each other, far away from being orthogonal. We consider both types of less/non-RIP scenarios in our experiments. To identify a successful reconstruction, support detection is our main concern, rather than the reconstruction error measured by $\|\mathbf{x}-\mathbf{u}\|_p$. We propose the support-detection-based evaluation metrics and detail it in Section 5.1.2.

The experimental group consists of:

- ℓ_1 minimization, as the baseline,
- PLM-2, in which $|\Omega_{\mathcal{P}2}| = 2$,
- PLM-3, in which $|\Omega_{\mathcal{P}3}| = 3$, consisting of $\Omega_{\mathcal{P}2}$ and a spurious index,
- PLM-s, in which $\Omega_{\mathcal{P}s} = supp(\mathbf{u})$. It is the best case that one can expect on reconstructing sparse signals through any optimization. Note, to do that, we assume that an oracle tells us the $supp(\mathbf{u})$.

5.1.1 Less/non-RIP Matrix Generation

RIP with δ_{2s} is a property such that, for all the subsets of column vectors in $\mathbf{A}$ with the cardinality not larger than $2s$, any pair in such subset should be almost perpendicular to each other. Therefore, we consider that the smaller the angle between two column vectors is, the less RIP the matrix is. Therefore, we make use of the spherical coordinate system [32] to generate less/non-RIP matrices as follows:

1. generate the random matrix $\mathbf{A} \in \mathbb{R}^{m\times N}$ with independent Gaussian entries;
2. select a set $\mathcal{C}$ of column vectors from $\mathbf{A}$ uniformly at random;
3. select one column vector in $\mathcal{C}$ as the central vector, give it polar angle $\mathbf{a}_0$ with $m-1$ entries $\overset{i.i.d.}{\sim} U(0,\pi/2)$;
4. generate $|\mathcal{C}|-1$ polar angles, each $\mathbf{a}_j$ $(j \neq 0)$ with $m-1$ entries $\overset{i.i.d.}{\sim} U(0,\pi/2)$;

5. set the radius $r = 1$, entries in the central vector is the Cartesian coordinates retrieved from r and $\mathbf{a}_0$;

6. select a constant $0 \leq \mu \leq 1$ as distance indicator, for the rest vectors in $\mathcal{C}$, entries of each is the Cartesian coordinates retrieved from r and $\mathbf{a}_0 + \mu\,\mathbf{a}_j$.

We would like to point out that the degree of less/non-RIP is not deterministic. And we call columns in $\mathcal{C}$ as *corrupted columns*.

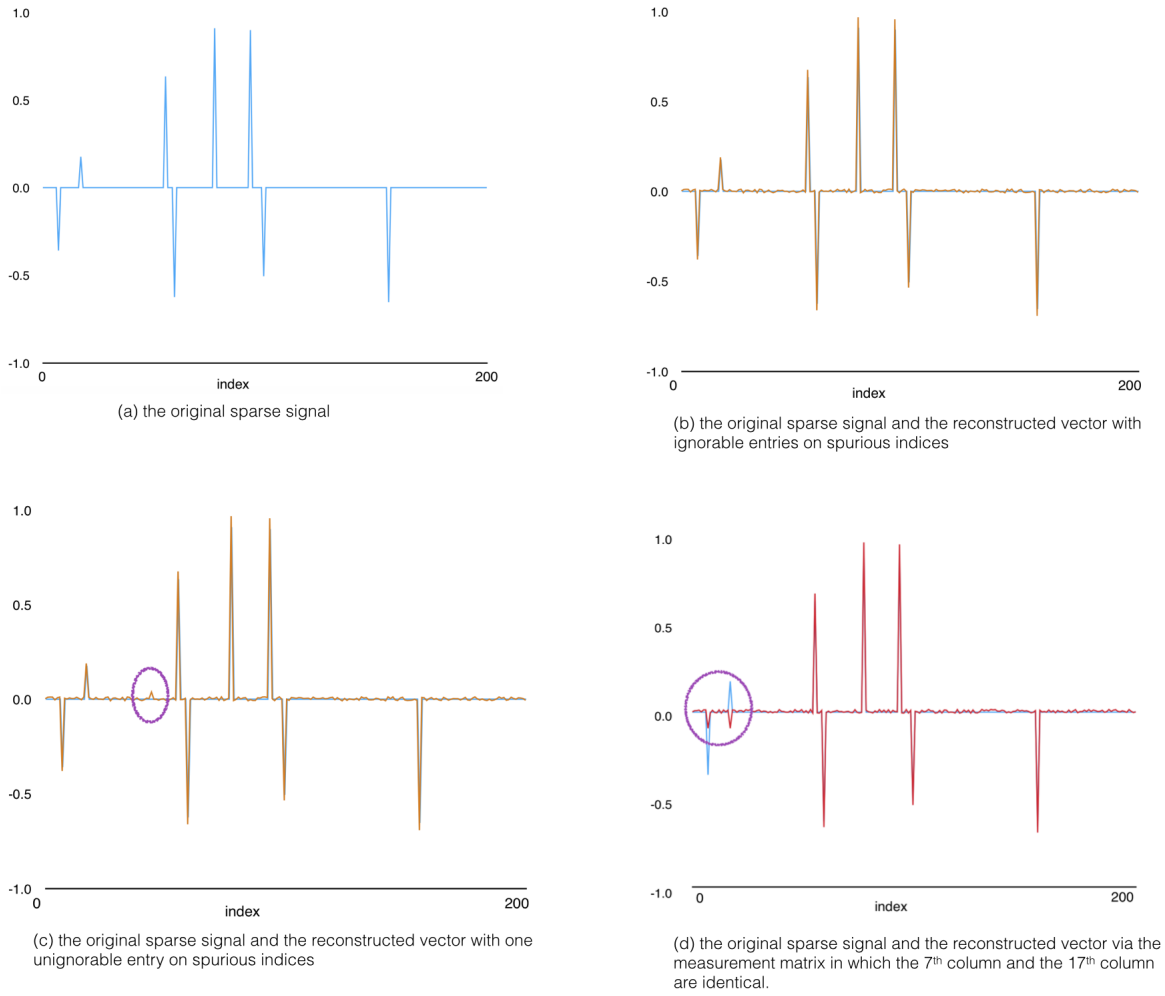
5.1.2 Evaluation Metric

Recognizing the support of the interested signal is the most important thing for our approach. However, considering that the iteration time and the return from CVXPY includes ignorable entries rather than absolute 0 (as shown in Figure 5.1 (b)), we need to provide tolerances for the approximations rather than exact reconstructions and the threshold for such tolerances should be considered (as shown in Figure 5.1 (c)). Furthermore, we notice that when all the columns in $\mathcal{C}$ are identical and $\mathcal{C} \cap \mathbb{J} = \mathbf{Q} \neq \varnothing$, the reconstructed vector through CVXPY is a vector in which all the entries with indices in $\mathcal{C}$ have the same value. Especially, if $\mathcal{C} == \mathbb{J}$, the reconstructed vector may have the right support but also large errors (as shown in Figure 5.1 (d)).

With the above concerns, we give our evaluation metric as follows: Let $\mathbb{J}_{\mathbf{u}}$ denote the support of the actual signal $\mathbf{u}$, $\mathbb{J}_{\mathbf{x}}$ denote the support of the minimizer $\mathbf{x}$. We define the minimizer $\mathbf{x}$ as the reconstruction of $\mathbf{u}$ if the following properties are satisfied:

$$\begin{aligned}
&\mathbb{J}_{\mathbf{u}} \subseteq \mathbb{J}_{\mathbf{x}};\\
&\max_{i \in \mathbb{J}_{\mathbf{x}} \setminus \mathbb{J}_{\mathbf{u}}} |x_i| \,/\, \min_{i \in \mathbb{J}_{\mathbf{x}}} |x_i| \leq \alpha;\\
&\text{no identical items in} \{x_i, i \in \mathcal{C}\}.
\end{aligned}$$

We would like to point out that the first two properties for evaluating an acceptable reconstruction are available for other cases, while the third one is mainly with respect to the behavior pattern of CVXPY.



(a) the original sparse signal

(b) the original sparse signal and the reconstructed vector with ignorable entries on spurious indices

(c) the original sparse signal and the reconstructed vector with one unignorable entry on spurious indices

(d) the original sparse signal and the reconstructed vector via the measurement matrix in which the 7th column and the 17th column are identical.

Figure 5.1: The original sparse signal and the reconstructed vectors.

5.1.3 Experiments Setting and Results

Parameters in all the experiments are set to be $\rho = 1.01$, $\epsilon = 10^{-8}$, iteration ≤ 1000, $\alpha = 0.01$. PLM-2 is given $\Omega_{\mathcal{P}2} = \{7, 17\}$. PLM-3 is given $\Omega_{\mathcal{P}3} = \{0, 7, 17\}$, in which the item "0" is a spurious index. We conduct the experiments as follows:

1. Experiment on m/s: While s increases, m must also increase sufficiently for satisfying the RIP. For a random matrix $\mathbf{A} \in \mathbb{R}^{m \times N}$, when m/s is not sufficient, it no longer satisfies the RIP with high probability. Accordingly, the reconstruction frequency via ℓ_1 minimization is supposed to decrease when s increases. We want to use this experiment to explore how PLM performs when s increases, and whether it is better than ℓ_1 minimization in this scenario.

   (a) fix the random matrix $\mathbf{A} \in 50 \times 200$ with entries $\overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$;

   (b) given $|\mathbb{J}| = s$, enforce $\Omega_{\mathcal{P}} \subseteq \mathbb{J}$, then select the remaining portion of the support,

$\mathbb{J} \setminus \Omega_{\mathcal{P}}$, uniformly at random, and sample a $\mathbf{u}$ on $\mathbb{J}$ with entries $\overset{i.i.d.}{\sim} U(-1,1)$, then run the experimental group on $\mathbf{u}$ and $\mathbf{A}$;

(c) repeat 50 times for each $s \in \{2, \ldots, 20\}$;

2. Experiment on angle distance: While the angle distance decreases, $\mathbf{A}$ becomes less/non-RIP. In this experiment, we are interested in two things. First, how PLM and ℓ_1 minimization perform when the angle distance decreases. To explore the first interest, we make the distance indicator μ change from 0 to 1 evenly and the step size be 0.02. Second, how PLM performs when the intersection of $\mathcal{C}$ and $\Omega_{\mathcal{P}}$ becomes smaller. To explore the second interest, we discuss four settings, $\mathcal{C} == \Omega_{\mathcal{P}}$, $\Omega_{\mathcal{P}} \subset \mathcal{C}$, $(\Omega_{\mathcal{P}} \cap \mathcal{C}) \subset \Omega_{\mathcal{P}}$, and $\Omega_{\mathcal{P}} \cap \mathcal{C} = \varnothing$.

(a) make $\mathcal{C}_1 = \Omega_{\mathcal{P}2}$;

(b) select $\mathcal{C}_2$ such that $\Omega_{\mathcal{P}2} \subset \mathcal{C}_2$, $|\mathcal{C}_2| = 25$;

(c) make $\mathcal{C}_3 = \mathcal{C}_2 \setminus \{7\}$, $|\mathcal{C}_2| = 24$;

(d) make $\mathcal{C}_4 = \mathcal{C}_2 \setminus \Omega_{\mathcal{P}2}$, $|\mathcal{C}_2| = 23$;

(e) changing μ from 0 to 1 with the step size being 0.02, correspondingly generate a sequence of less/non-RIP $\mathbf{A}'_{\mathcal{C}_1}$ from $\mathbf{A}$ with respect to $\mathcal{C}_1$, then $\mathbf{A}'_{\mathcal{C}_2}$, $\mathbf{A}'_{\mathcal{C}_3}$, $\mathbf{A}'_{\mathcal{C}_4}$;

(f) fix $s = 8$, generate 50 $\mathbf{u}$ following the way in 2(b);

(g) run the experimental group with the above 50 $\mathbf{u}$ on each matrix in $\mathbf{A}'_{\mathcal{C}_1}$, $\mathbf{A}'_{\mathcal{C}_2}$, $\mathbf{A}'_{\mathcal{C}_3}$ and $\mathbf{A}'_{\mathcal{C}_4}$;

3. Experiment on $\sum_{i \in \Omega_{\mathcal{P}}} |x_i| / \|\mathbf{x}\|$: A small ratio indicates entries in $\Omega_{\mathcal{P}}$ is not significant in $\mathbf{x}$. Through this experiment, we want to explore whether PLM prefers an $\Omega_{\mathcal{P}}$ with indices corresponding to larger entries.

(a) fix $\mathbf{A}$, and make $s = 15$;

(b) generate $\mathbf{u}$ as the way in 2(b), then run the experimental group on $\mathbf{A}$. Repeat 500 times.

The results are shown in Figures 5.2-5.7. Given a random matrix, partial ℓ_1 minimization is more capable than ℓ_1 minimization when m is too small compared with s. When the measurement matrix is no more a random matrix and corrupted columns are involved, partial ℓ_1

minimization is more tolerant than ℓ_1 minimization, within the conditions of interest (highlighted in grey in Figure 5.2-5.5). Here, *the condition of interest* means a condition that PLM-s is able to fulfill uniform reconstruction, that is the reconstruction frequency equals to 100%. As the intersection between $\Omega_{\mathcal{P}}$ and $\mathcal{C}$ decreases, the performance of partial ℓ_1 minimization gradually approaches ℓ_1 minimization. The fluctuation of ℓ_1 minimization in Figure 5.3 between $\mu = 0.42$ and $\mu = 0.7$, and the steeper slope of the partial-ℓ_1-reconstruction in Figure 5.5 ($\mu \leq 0.18$) and Figure 5.6 ($\mu \leq 0.08$) may be because $\mathbf{u}$ is randomly generated, in terms of the support and entries, and the value of μ is just an approximation to the degree of less/non-RIP. The advantage of PLM shown in Figure 5.2 is not as obvious as the one exhibited in Figures 5.3-5.6. A possible reason is that fewer number of measurements causes the measurement matrix to become less/non-RIP globally, while corrupted columns only have regional influence. It is the implication of "*restricted*" in RIP. A good feature of PLM as shown in Figure 5.7 is that it does not discriminate $\Omega_{\mathcal{P}}$ with indices corresponding to small entries.

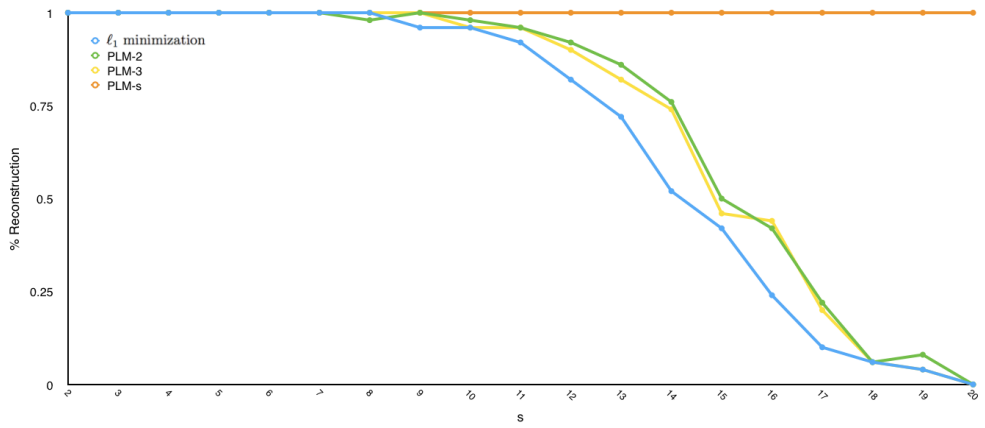


Figure 5.2: The reconstruction frequency (50 runs) versus the sparsity s. When s increases, the fixed Gaussian matrix no longer satisfies satisfy RIP with high probability. PLM-2 and PLM-3 both have higher reconstruction frequency than ℓ_1 minimization as m/s get smaller. PLM-3 performs almost as well as PLM-2, even though it is provided a spurious index. PLM-s is able to keep 100% reconstruction even when all the other three optimization solvers completely fails. In fact, PLM-s exhibits the capability of ℓ_0 minimization if the computational complexity to solve it were not a issue. The gap between PLM-s and ℓ_1 minimization shows that ℓ_1 minimization asks for more resources to guarantee a uniform reconstruction. PLM-2 and PLM-3 both display improvements on ℓ_1 minimization.

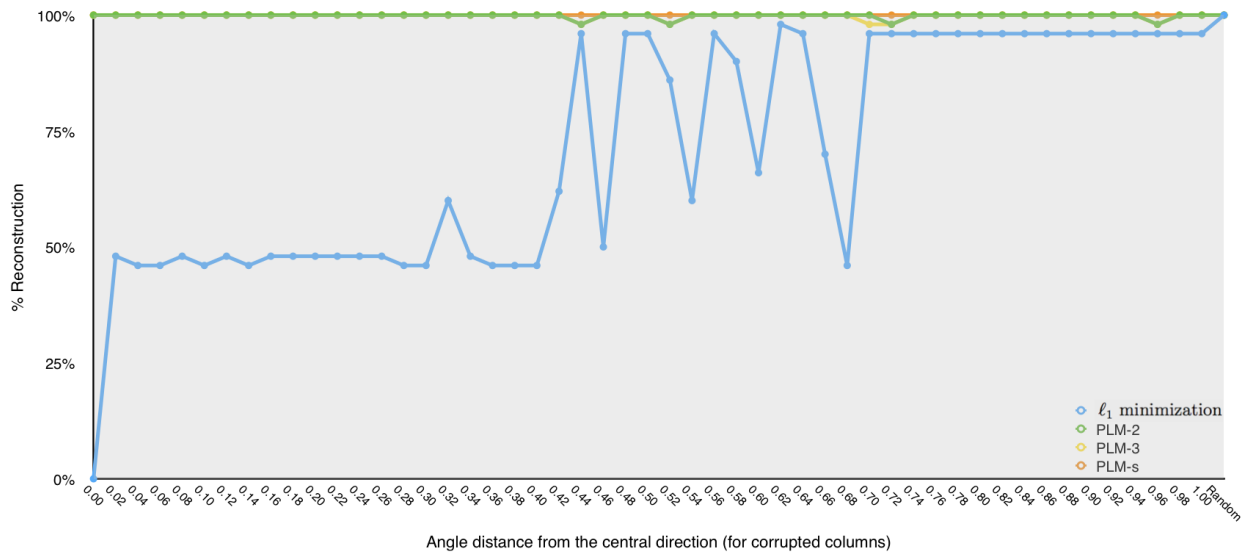


Figure 5.3: The reconstruction frequency (50 runs) versus the angle distance indicator μ, when $\mathcal{C} == \Omega_{\mathcal{P}2}$ and $s = 8$. PLM-2, PLM-3 and PLM-s all perform extremely well than ℓ_1 minimization. PLM-2 and PLM-3 perform as well as PLM-s. This exhibits how the $\Omega_{\mathcal{P}}$ can relax the RIP requirement on $\mathbf{A}$. The reconstruction frequency of ℓ_1 minimization has obvious fluctuation between $\mu = 0.42$ and $\mu = 0.7$.

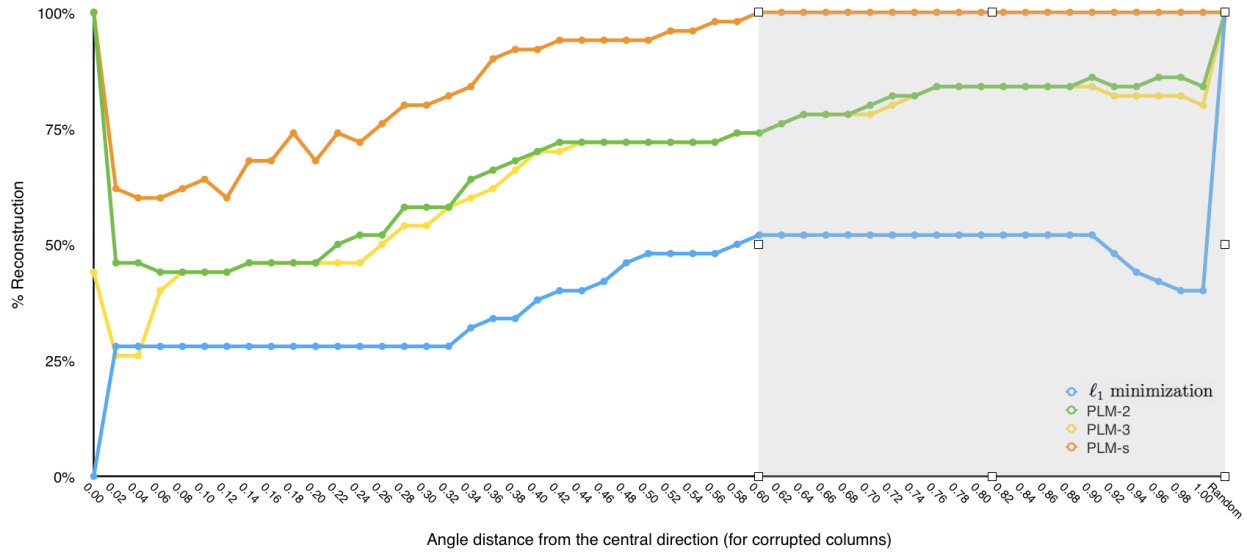


Figure 5.4: The reconstruction frequency (50 runs) versus the angle distance indicator μ, when $\Omega_{\mathcal{P}2} \subset \mathcal{C}$ and $s = 8$. In this experiment, $\Omega_{\mathcal{P}2}$ are included in $\mathcal{C}$. To compare the performance among the experimental group, we only concern the range that PLM-s has stable uniform reconstruction, highlighted in grey. PLM-2 and PLM-3 perform overwhelming better than ℓ_1 minimization between $\mu = 0.6$ and $\mu = 1$. However, the steep jump at $\mu = 0.02$ shows in PLM-2, PLM-3 and PLM-s needs further study. Note, the range of the grey area is narrower compared to Figure 5.3, Figure 5.5, and Figure 5.6. This is because in this setting, more columns are corrupted than other settings. This is exactly the expected phenomena in the less/non-RIP case.

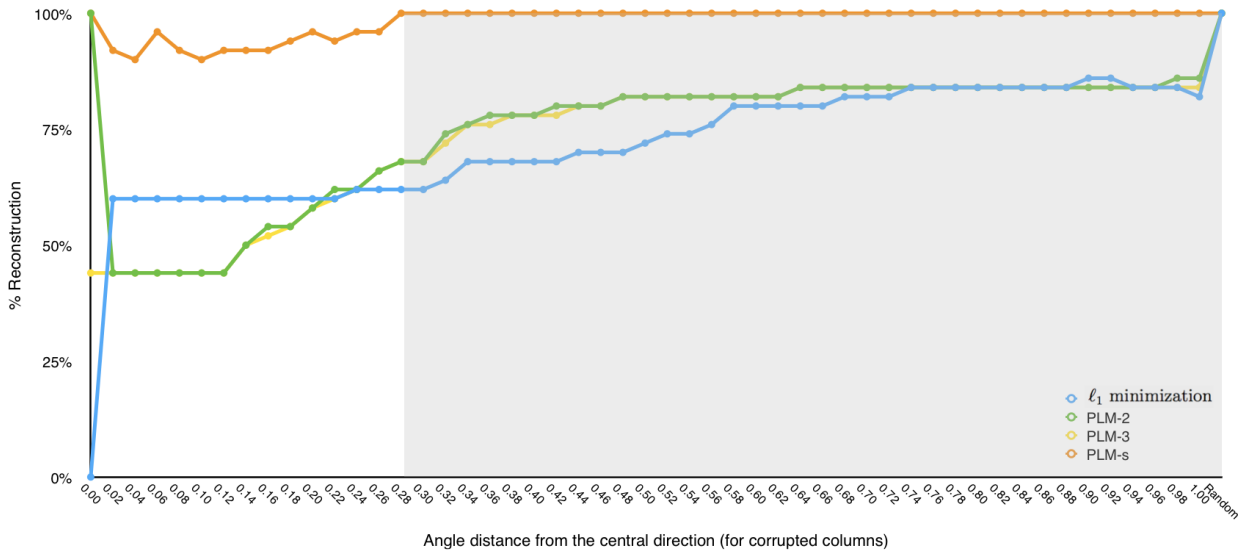


Figure 5.5: The reconstruction frequency (50 runs) versus the angle distance indicator μ, when $\mathcal{C} \cap \Omega_{\mathcal{P}2} = 7$ and $s = 8$. In this experiment, $\mathcal{C}$ and $\Omega_{\mathcal{P}2}$ only overlap. The other freed direction is supposed to be useless. However, the reconstruction frequency of PLM-2 is 100% when the 7^{th} column is identical to other 23 columns rather than the 17^{th} column. For comparison, we only concern the range that PLM-s has stable uniform reconstruction, highlighted in grey. PLM-2 and PLM-3 reach higher reconstruction frequency than ℓ_1 minimization between $\mu = 0.28$ and $\mu = 0.56$. After that, PLM-2, PLM-3 and ℓ_1 minimization are almost the same. The general trend in this experiment shows the reconstruction frequency increases while the angle distance indicator μ increases as well. This shows that our method to generate less/non-RIP matrices is reasonable.

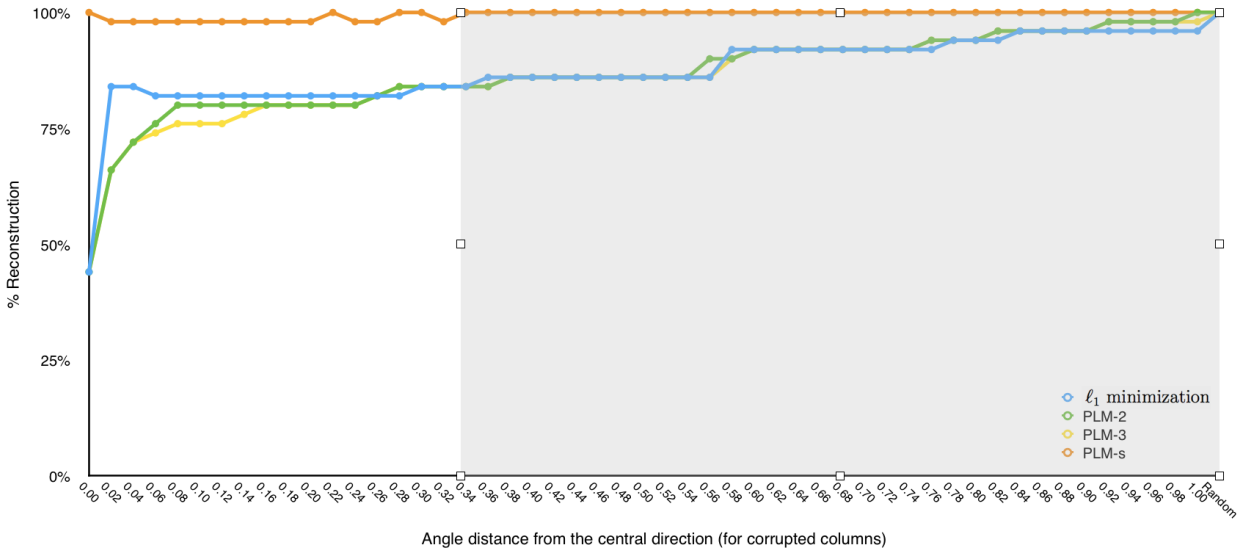


Figure 5.6: The reconstruction frequency (50 runs) versus the angle distance indicator μ, when $\mathcal{C} \cap \Omega_{\mathcal{P}2} = \varnothing$ and $s = 8$. For comparison, we only concern the range that PLM-s has stable uniform reconstruction, highlighted in grey. When $\Omega_{\mathcal{P}2}$ has no relations with $\mathcal{C}$, PLM-2, PLM-3 and ℓ_1 minimization perform almost the same. This shows that setting part of directions free may at least do not make things worse, and as long as the freed directions have relations with the corrupted sets, it can help the optimization solver be more reliable.

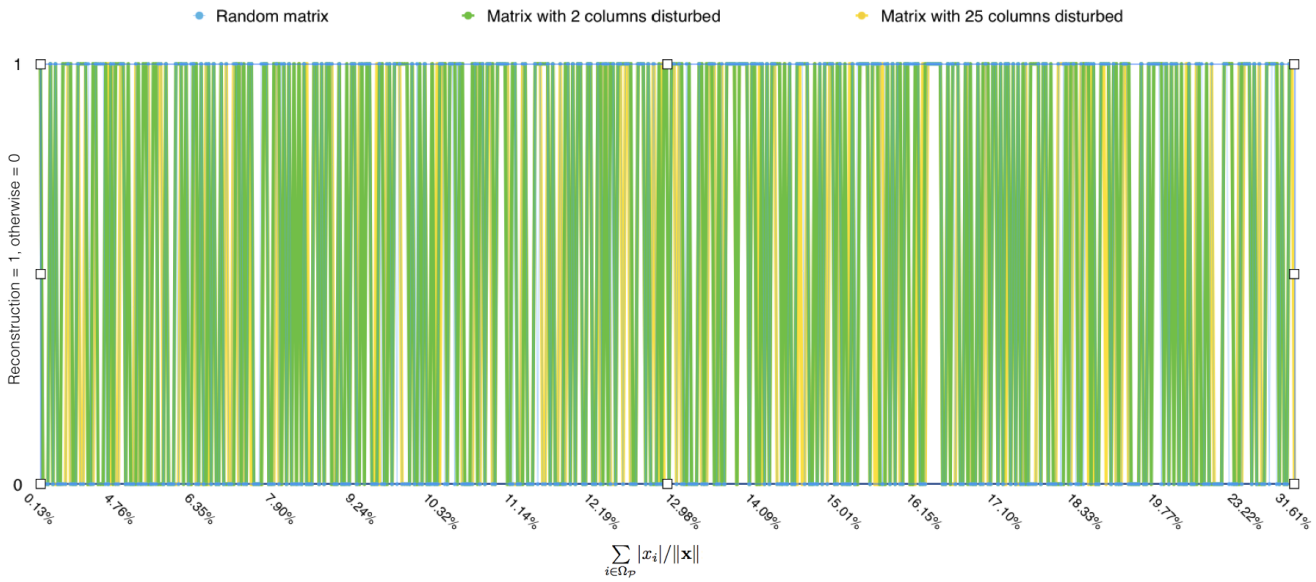


Figure 5.7: The PLM-reconstruction versus $\sum_{i\in\Omega_{\mathcal{P}}} |x_i|/\|\mathbf{x}\|$, with $s = 15$. As the ratio increases, there is no trend on the PLM-reconstruction, no matter under a random matrix or less/non-RIP matrix. As the reconstructed vector $\mathbf{u}$ is randomly generated, the upper bound of the ratio in 500 runs is 31.6%. It is safe to say that at the range $\sum_{i\in\Omega_{\mathcal{P}}} |x_i|/\|\mathbf{x}\| \leq 31.6\%$, the partial-$\ell_1$-reconstruction is not sensitive to how large the entries with indices in $\Omega_{\mathcal{P}}$.

5.2 Image Data

This section is an implementation of the general sampling process (see details in Section 2.3) in CS. Many of the wavelet transforms are very efficient to compress images. As our formulation and solver are limited in real number domain, here we use the discrete cosine tranform(DCT) as Ψ, a compression methodology used in JPG (a commonly used method of lossy compression for digital images) [39]. The Gaussian matrix with entries $\overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ is used as Φ. Then m rows are uniformly, and randomly picked up from Φ to formulate Φ^* as our sensing matrix. The correspondingly less/non-RIP $\mathbf{A}'$ for this experiment is generated by enforcing part of column vectors to be identical. Before acquiring measurements, we preprocess the 2D-image data as follows:

1. transform a 2D pixel image into the frequency domain via 2D-DCT, to get a compressible representation;

2. preserve the s significant coefficient in the 2D compressible representation, and shrink all the other entries to be 0;

3. flatten the 2D sparse representation as a vector, that is $\mathbf{u}$.

Here, $\mathbf{u}$ is our input signal to be reconstructed. In this experiment, $N = 2000$, $s = 168 = 0.084N$, $m = 504 = 3s$, and $|\Omega_{\mathcal{P}}| = 5 = 0.03s$. The reconstruction from a corrupted measurement matrix via our optimization solver exactly shows the original pattern, while the reconstruction via CVXPY blurs the shape. Compared with the number of significant coefficients, the cardinality of $\Omega_{\mathcal{P}}$ is very small, but it is enough to improve the performance of the classic ℓ_1 minimization.

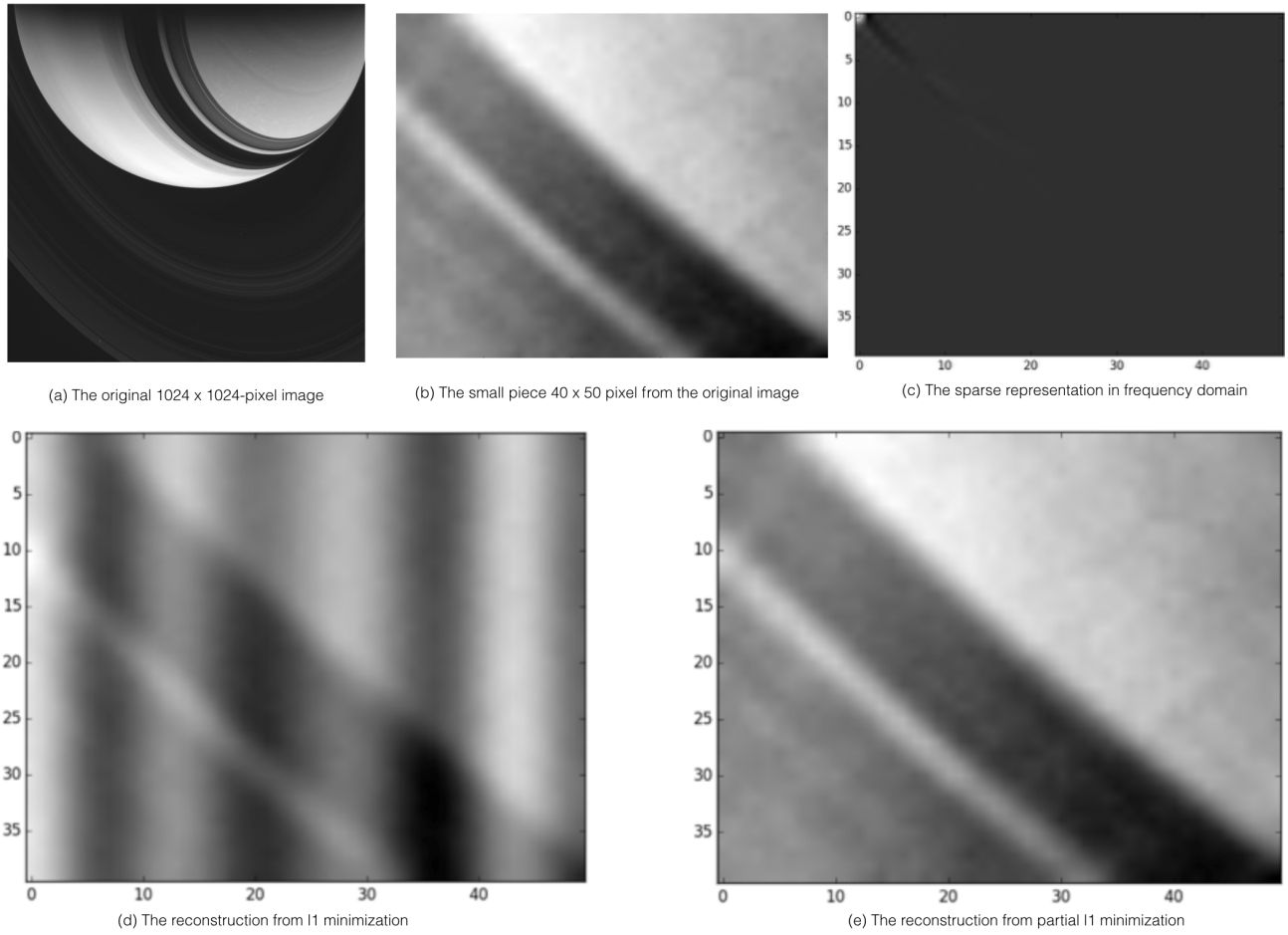


Figure 5.8: The reconstruction of preprocessed image data. The disturbed column set in $\mathbf{A}'$ is $\{1, \ldots, 10\}$. (a) the original 1024×1024 image data in the time domain. (b) limited by the memory and computing efficiency of the algorithms, only part of the original image is kept as the signal of interest, $N = 40 \times 50$. (c) locations of significant coefficients in frequency domain. (d) $\Omega_{\mathcal{P}} = \{1, \ldots, 5\}$, reconstruction via PLM-ALM, with $\rho = 1.01$, $\epsilon = 1e{-}8$, iteration ≤ 1000. (e) reconstruction via CVXPY. While the reconstruction via CVXPY is blurry, the reconstruction via PLM-ALM is clearly recognizable.

Chapter 6

Discussion

6.1 Conclusion

We have presented a novel optimization approach, called partial ℓ_1 minimization (PLM), for a sparse signal reconstruction from undersampled measurements. It is an approach to combine the convexity of ℓ_1 minimization and the robustness of ℓ_0 minimization. The empirical results demonstrate that PLM is more reliable than ℓ_1 minimization when RIP is not satisfied. To make such performance improvements, PLM requires only minor information about the support, which is allowed to include spurious indices.

6.2 Future Work

We propose two directions of the future work:

First, the pattern as shown in the synthetic experiments should explained in a theoretical way. We would like to prove the δ_{2s} upper bound for guaranteeing uniform reconstruction via PLM.

Second, the limitation of the proposed optimization solver has been shown based on the experiments on image data. As PLM can be solved by most existing solvers for ℓ_1 minimization with straightforward modifications, we can try another scheme to speed up the computing.

Third, as shown in the experiments on the synthetic data, PLM-s displays the best case that we can expect to deal with less/non-RIP $\mathbf{A}$. How to modify TLM to be a convex optimization problem is a key goal. Prof. Randy Paffenroth has conjectured that given a sufficiently large

$\varepsilon \in \mathbb{R}^+$, the nonconvex TLM can be transformed into a convex problem as follows:

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^N} \|\mathbf{x}\|_1 - \max_{\mathbf{z}\in\mathbb{R}^N} \|\mathcal{D}^r(\mathbf{z})\,\mathbf{x}\|_1 \\ &\text{subject to} \quad |\mathbf{A}\mathbf{x} - \mathbf{y}| \dot{\leq} \varepsilon, \end{aligned}$$

where, $\dot{\leq}$ means "point-wise smaller than". We will attempt to further verify this conjecture empirically and theoretically.

Bibliography

[1] Candès, Emmanuel J. "The restricted isometry property and its implications for compressed sensing." *Comptes Rendus Mathematique 346.9 (2008): 589-592.*

[2] Candès, Emmanuel J., and Michael B. Wakin. "An introduction to compressive sampling." *IEEE Signal Processing Magazine 25.2 (2008): 21-30.*

[3] Foucart, Simon, and Holger Rauhut. *A mathematical Introduction to Compressive Sensing. Vol. 1. No. 3. Boston: Birkhäuser, 2013.*

[4] Hu, Yao, et al. "Fast and accurate matrix completion via truncated nuclear norm regularization." *IEEE Transactions on Pattern Analysis and Machine Intelligence 35.9 (2013): 2117-2130.*

[5] Candès, Emmanuel J., and Terence Tao. "Decoding by linear programming." *IEEE Transactions on Information Theory 51.12 (2005): 4203-4215.*

[6] Lustig, Michael, David Donoho, and John M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging." *Magnetic Resonance in Medicine 58.6 (2007): 1182-1195.*

[7] Egiazarian, Karen, Alessandro Foi, and Vladimir Katkovnik. "Compressed sensing image reconstruction via recursive spatially adaptive filtering." *Image Processing, 2007. ICIP 2007. IEEE International Conference on. Vol. 1. IEEE, 2007.*

[8] Candès, Emmanuel, and Justin Romberg. "Sparsity and incoherence in compressive sampling." *Inverse problems 23.3 (2007): 969.*

[9] Takhar, Dharmpal, et al. "A new compressive imaging camera architecture using optical-domain compression." *Electronic Imaging 2006. International Society for Optics and Photonics, 2006.*

[10] Tutorial of compressed sensing. https://www.codeproject.com/Articles/852910/Compressed-Sensing-Intro-Tutorial-w-Matlab

[11] Wikipedia. Linear isometry. https://en.wikipedia.org/wiki/Isometry#Linear_isometry

[12] Daubechies. *Ten Lectures of Wavelets, SIAM, 1992.*

[13] S. A. Vavasis. *Nonlinear Optimization: Complexity Issues, Oxford University Press, New York, 1991.*

[14] Candès, Emmanuel J., Justin Romberg, and Terence Tao. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information." *IEEE Transactions on Information Theory 52.2 (2006): 489-509.*

[15] Boyd, Stephen, et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers." *Foundations and Trends ® in Machine Learning 3.1 (2011): 1-122.*

[16] Wikipedia. Spike-and-Slab. https://en.wikipedia.org/wiki/Spike-and-slab_variable_selection#cite_note-5

[17] Ishwaran, Hemant, and J. Sunil Rao. "Spike and slab variable selection: frequentist and Bayesian strategies." *Annals of Statistics (2005): 730-773.*

[18] Brown, Philip J., Marina Vannucci, and Tom Fearn. "Multivariate Bayesian variable selection and prediction." *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 60.3 (1998): 627-641.*

[19] Prugovecki, Eduard. *Quantum Mechanics in Hilbert Space. Vol. 92. Academic Press, 1982.*

[20] Beck, Amir, and Marc Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems." *SIAM Journal on Imaging Sciences 2.1 (2009): 183-202.*

[21] Yang, Allen Y., et al. "Fast ℓ_1-minimization algorithms and an application in robust face recognition: A review." *Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE, 2010.*

[22] Hale, Elaine T., Wotao Yin, and Yin Zhang. "A fixed-point continuation method for ℓ_1-regularized minimization with applications to compressed sensing." *CAAM TR07-07, Rice University 43 (2007): 44.*

[23] Yun, Sangwoon, and Kim-Chuan Toh. "A coordinate gradient descent method for ℓ_1-regularized convex minimization." *Computational Optimization and Applications 48.2 (2011): 273-307.*

[24] Subgradient. `https://see.stanford.edu/materials/lsocoee364b/01-subgradients_notes.pdf`

[25] Bertsekas, Dimitri P. "Multiplier methods: a survey." *Automatica 12.2 (1976): 133-145.*

[26] Diamond, Steven, and Stephen Boyd. "CVXPY: A Python-embedded modeling language for convex optimization." *Journal of Machine Learning Research 17.83 (2016): 1-5.*

[27] Donoho, David L. "Compressed sensing." *IEEE Transactions on Information Theory 52.4 (2006): 1289-1306.31.*

[28] Brady, David J., et al. "Compressive holography." *Optics Express 17.15 (2009): 13040-13049.*

[29] Provost, Jean, and Frédéric Lesage. "The application of compressed sensing for photo-acoustic tomography." *IEEE Transactions on Medical Imaging 28.4 (2009): 585-594.*

[30] Lensless Imaging with Compressive Ultrafast Sensing. `http://web.media.mit.edu/~guysatat/singlepixel/`

[31] Wikipedia. Spherical coordinate system.`https://en.wikipedia.org/wiki/Spherical_coordinate_system`

[32] Nesterov, Yurii. *Introductory Lectures on Convex Optimization: A Basic Course. Vol. 87. Springer Science & Business Media, 2013.*

[33] Wikipedia. Strongly NP-hard. `https://en.wikipedia.org/wiki/Strong_NP-completeness`

[34] Wikipedia. Overcompletness. `https://en.wikipedia.org/wiki/Overcompleteness`

[35] Wikipedia. Convex function. `https://en.wikipedia.org/wiki/Convex_function`

[36] Tillmann, Andreas M., and Marc E. Pfetsch. "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing." *IEEE Transactions on Information Theory 60.2 (2014): 1248-1259.*

[37] Baraniuk, Richard, et al. "A simple proof of the restricted isometry property for random matrices." *Constructive Approximation 28.3 (2008): 253-263.*

[38] Wikipedia. JPEG. `https://en.wikipedia.org/wiki/JPEG`

[39] Shor, N. Z. *Minimization Methods for Non-Differentiable Functions (Vol. 3). Springer Science & Business Media. 2012.*