Newton-type and Partial Gradient Thresholding Algorithms for Sparse Signal Reconstruction

By

Nan Meng

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
University of Birmingham
December 2021

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Sparse signal recovery, one of the main tasks in signal processing, can be formulated as specific sparse optimization problems. Due to the NP-hardness of such problems, fruitful methods have been proposed at the current stage of development, among which thresholding-type methods have received considerable attention due to their structural simplicity and low complexity. The key step of thresholding methods can be regarded as integrating a search direction in traditional nonlinear optimization method and a specific thresholding technique that generates sparse vectors. The primary work of the thesis is to propose several novel thresholding-type algorithms by merging some existing thresholding techniques with new search directions.

(a) We first develop the Newton-step-based hard thresholding by introducing a modified Newton-like direction to replace the steepest descent direction that appears in many hard thresholding methods in the literature. Our method is motivated by the classic optimization theory which suggests that the Newton-like method often has a numerical advantage over the classic gradient method from the perspective of local convergence.

(b) We propose the Newton-type optimal k-thresholding algorithms, motivated by the appreciable performance of Newton-type methods and the latest more powerful optimal thresholding technique which overcomes the inherent weakness of hard thresholding by simultaneously performing objective minimization and thresholding.

(c) In commonly used thresholding methods, the iterate to be thresholded is generally not sparse due to the adoption of full descent directions, thus a large number of important pieces of information are usually thrown out during the thresholding process. We thereby adopt the partial gradient direction to generate a sparser iterate and combine the more efficient optimal k-thresholding method to propose the so-called partial gradient optimal thresholding algorithms.

(d) It was known that the use of hard thresholding usually increases the objective value in the iterative process. To avoid such a potential increase of the objective value, one can develop the optimal partial gradient direction that minimizes the objective function at a current iterate, leading to the so-called optimal partial gradient hard thresholding methods.

We will show the efficiency of the proposed algorithms from a theoretical point of view. The sufficient conditions for the efficiency of our algorithms are established in terms of the restricted isometry property (RIP), which is one of the fundamental tools for analyzing compressed sensing algorithms. Empirical results show that our algorithms are stable and robust for sparse signal reconstruction and comparable to some mainstream compressed sensing algorithms.

Acknowledgements

It has been an honor to complete my Ph.D. study at the School of Mathematics, University of Birmingham. I am deeply grateful to all those who have kindly supported me and provided excellent guidance and assistance.

I would like to express my sincere gratitude to my supervisor, Prof. Yun-Bin Zhao, for his guidance, encouragement, and patience throughout my Ph.D. study. He has deeply influenced my academic career with his dedicated work style, distinctive personality, profound knowledge, and strong responsibility. I would not have been able to present this thesis in this form without his invaluable guidance. My sincere gratitude also goes out to my supervisor, Prof. Michal Kočvara, for his helpful suggestions, patience, and assistance through my Ph.D. study.

In addition, I wish to express my appreciation to Dr. Sándor Zoltán Németh for his guidance and suggestions in my postgraduate research project. I would like to thank Dr. Daniel Loghin for his guidance and advice, and I have also learned from him to be conscientious, even if it is just a punctuation mark. Special thanks to Prof. Houduo Qi for his helpful suggestions, which helped improve the thesis. I am grateful for the company of my friends and for the wonderful time that we shared outside of work. Thank you all for the support, patience, encouragement, and friendship you have shown me.

Lastly, but above all, I would like to express my appreciation to my family members whose support and encouragement are valued beyond what can be expressed in words. Having my loved ones by my side has enabled me to achieve this most challenging academic career, and I would like to wish them a happy future.

I hope that this thesis will not be the end of the academic road, and that the preceding sentence is not empty words.

Table of Notation

Notation	Definition
$\mathbb{R}$	The real space.
$\|x\|_0$	The number of nonzero components of x.
$\|x\|_1$	$\sum_i^n \|x_i\|$, where $x \in \mathbb{R}^n$.
$\|x\|_2$	$\sqrt{\sum_i^n \|x_i\|^2}$, where $x \in \mathbb{R}^n$.
$\|x\|_\infty$	$\max_{i\in[n]} \|x_i\|$.
$\|A\|_2$	$\sqrt{\lambda_{\max}(A^T A)}$, where $A \in \mathbb{R}^{m\times n}$.
$[N]$	$\{1, \ldots, n\}$.
$\overline{S}$	The complementary set of S.
e	Euler's number.
$\mathbf{e}$	The vector of ones, i.e., $(1, \ldots, 1)^T$.
I	The identity matrix.
$\nabla f(x)$	The gradient of $f(x)$.
$\nabla^2 f(x)$	The Hessian of $f(x)$.
$\text{supp}(x)$	$\{i : x_i \neq 0\}$.
$\text{card}(S)$	The cardinality of a set $S \subseteq [N]$.
$\mathcal{T}_k(\cdot)$	A thresholding operator.
$\mathcal{H}_k(\cdot)$	The hard thresholding operator which retains k largest magnitudes of a vector and zeroes out others.
$\mathcal{L}_k(\cdot)$	The support set of k largest magnitudes of a vector.
$\mathcal{Z}_k^\#(\cdot)$	The optimal k-thresholding operator.
$\langle \cdot, \cdot \rangle$	The inner product in Euclidean space.
$\sigma_1(A)$	The largest singular value of matrix $A \in \mathbb{R}^{m\times n}$.
$\sigma_m(A)$	The smallest singular value of matrix $A \in \mathbb{R}^{m\times n}$.
Σ_k	The set of k-sparse vectors.
$\otimes$	The Hadamard product.

CONTENTS

Table of Notation ii

1. Introduction **1**
1.1. Sparse optimization problems 2
1.2. Main contribution 6
1.3. Organization of thesis 10
1.4. Notation 10
1.4.1. Sets, vectors and matrices 11
1.4.2. Norms for vectors and matrices 11
1.4.3. Convex sets and functions 13

2. Literature Review **14**
2.1. Properties of measurement matrices 14
2.1.1. Minimal number of measurements 14
2.1.2. Restricted isometry property (RIP) 16
2.1.3. Other properties 18
2.2. Optimization methods 19
2.2.1. Steepest descent method 20
2.2.2. Newton's method 21
2.3. Compressive sensing algorithms 23
2.3.1. The ℓ_1-minimization 23
2.3.2. Hard thresholding algorithms 26
2.3.3. Optimal k-thresholding algorithms 29
2.3.4. Heuristic algorithms 32

3. Newton-step-based Hard Thresholding Algorithms **36**
3.1. Introduction 36
3.2. Algorithms 38
3.2.1. Extension to general problem 40
3.3. Theoretical Analysis 41
3.3.1. Analysis of NSIHT in noisy scenarios 42
3.3.2. Analysis of NSHTP in noisy scenarios 53

3.4. Numerical Experiments . . . 61
3.4.1. Residual Reduction . . . 62
3.4.2. Performance of sparse signal recovery . . . 64
3.4.3. Phase transition . . . 70
3.4.4. Main findings . . . 71
3.5. Conclusion . . . 74

4. Newton-Type Optimal k-Thresholding Algorithms 75
4.1. Introduction . . . 75
4.2. Algorithms . . . 77
4.2.1. Extension to the general case . . . 80
4.3. Theoretical Analysis . . . 81
4.3.1. Analysis of NTOT in noisy scenarios . . . 81
4.3.2. Analysis of NSROT in noisy scenarios . . . 86
4.3.3. Analysis of NTROTP in noisy scenarios . . . 91
4.4. Numerical Performance . . . 95
4.4.1. Residual reduction . . . 95
4.4.2. The number of iterations . . . 96
4.4.3. The performance of signal recovery . . . 99
4.4.4. Phase transition . . . 101
4.5. Conclusion . . . 104

5. Partial Gradient Optimal k-Thresholding Algorithms 105
5.1. Introduction . . . 105
5.2. Algorithms . . . 107
5.2.1. Extension to general problem . . . 110
5.3. Theoretical Analysis . . . 110
5.3.1. PGOT in noisy scenarios . . . 111
5.3.2. PGROT in noisy scenarios . . . 116
5.3.3. PGROTP in noisy scenarios . . . 120
5.3.4. Solving sparse optimization problem . . . 122
5.4. Numerical Experiments . . . 124
5.4.1. Residual reduction . . . 124
5.4.2. Number of iterations . . . 125
5.4.3. Sparse signal recovery . . . 127
5.4.4. Phase transition . . . 128
5.4.5. Main findings . . . 133
5.5. Conclusion . . . 133

6. Optimal Partial Gradient Hard Thresholding Algorithms **135**
6.1. Introduction . . . 135
6.2. Algorithms . . . 137
6.3. Theoretical Analysis . . . 139
6.3.1. Analysis of OPGHT in noisy scenarios . . . 139
6.3.2. Analysis of ROPGHT in noisy scenarios . . . 141
6.3.3. Analysis of ROPGHTP in noisy scenarios . . . 148
6.4. Numerical Experiments . . . 150
6.4.1. Residual reduction . . . 151
6.4.2. Phase transition . . . 152
6.5. Conclusion . . . 155

7. Conclusions and Future Work **156**

List of References **158**

LIST OF ALGORITHMS

2.1. Steepest Descent Method 20
2.2. Newton's Method (pure form) 22
2.3. Iterative Hard Thresholding (IHT) 27
2.4. Hard Thresholding Pursuit (HTP) 28
2.5. Optimal k-thresholding (OT) 30
2.6. Relaxed Optimal k-thresholding (ROT) 31
2.7. Relaxed optimal k-thresholding pursuit (ROTP) 32
2.8. Orthogonal Matching Pursuit (OMP) 33
2.9. Compressive Sampling Matching Pursuit (CoSaMP) 34
2.10. Subspace Pursuit (SP) 34
3.1. Newton-Step-Based Iterative Hard Thresholding (NSIHT) 39
3.2. Newton-Step-Based Hard Thresholding Pursuit (NSHTP) 39
3.3. Generalized Newton-Step-Based Iterative Hard Thresholding (GNSIHT) . . 40
3.4. Generalized Newton-Step-Based Hard Thresholding Pursuit (GNSHTP) . . 41
4.1. Newton-Type Optimal k-Thresholding (NTOT) 78
4.2. Newton-Type Relaxed Optimal k-Thresholding (NTROT) 79
4.3. Newton-Type Relaxed Optimal k-Thresholding Pursuit (NTROTP) 79
4.4. Generalized Newton-Type Relaxed Optimal k-Thresholding Pursuit (GNTROT) 80
5.1. Partial Gradient Optimal k-Thresholding (PGOT) 109
5.2. Partial Gradient Relaxed Optimal k-Thresholding (PGROT) 109
5.3. Partial Gradient Relaxed Optimal k-thresholding Pursuit (PGROTP) . . . 109
5.4. Generalized Partial Gradient Relaxed Optimal k-thresholding Pursuit (GRPGOTP) 110
6.1. Optimal Partial Gradient Hard Thresholding (OPGHT) 138
6.2. Relaxed Optimal Partial Gradient Hard Thresholding (ROPGHT) 138
6.3. Relaxed Optimal Partial Gradient Hard Thresholding Pursuit (ROPGHTP) 138

LIST OF FIGURES

1.1. Example of compressed sensing model. The unknown signal is assumed to be sparse, containing few entries that are non-zero, whereas the measurement matrix is not necessarily sparse. 3

1.2. Graphs of the ℓ_0-norm, $\ell_{0.1}$-norm, $\ell_{0.5}$-norm, $\ell_{0.9}$-norm and the ℓ_1-norm. The horizontal axis represents the value of x, and the vertical axis represents the values of functions. 5

1.3. The intersection of a straight line and the unit ball defined by the ℓ_1-norm, ℓ_2-norm, $\ell_{0.5}$-norm and $\ell_{0.1}$-norm. 5

1.4. Process of thresholding-type methods 10

2.1. Classification of mainstream compressed sensing algorithms 24

3.1. The residual reduction of HTP, NSIHT and NSHTP in the course of iterations. The size of measurement matrices is 256×512. The measurements are exact $y = Ax^*$ with sparsity level $k = 50$. 62

3.2. Performance of NSIHT and NSHTP in residual reduction with $\epsilon = \sigma_1^2 + 1$ and different step lengths: $\lambda = 1, 2, 4, 5$. The size of measurement matrices is 256×512. The measurements are exact $y = Ax^*$ with sparsity level $k = 50$. 64

3.3. Performance of NSIHT and NSHTP in residual reduction with different choices of ϵ. The size of measurement matrices is 256×512. The measurements are exact $y = Ax^*$ with sparsity level $k = 50$. 65

3.4. Recovery performance of NSIHT and NSHTP with different dimensions of matrices. The sizes of measurement matrices are 150×500, 200×500 and 250×500. The measurements are exact: $y = Ax^*$. 66

3.5. Recovery performance of NSIHT and NSHTP with a fixed step length ($\lambda^* = 5$) and different choices of ϵ ($\epsilon_1 = \epsilon^*$, $\epsilon_2 = 1.1\epsilon^*$, $\epsilon_3 = 1.5\epsilon^*$, $\epsilon_4 = 2\epsilon^*$). The size of measurement matrices is 256×512, and measurements are exact $y = Ax^*$. Maximum iterations for NSIHT and NSHTP are 35 and 20, respectively. 67

3.6. Recovery performance of NSIHT and NSHTP with different choices of stepsizes. The size of measurement matrices is 256×512, and measurements are exact $y = Ax^*$. 67

3.7. Comparison of NSIHT and NSHTP by performing different numbers of iterations. The size of measurement matrices is 256×512, and measurements are exact $y = Ax^*$. 68
3.8. Comparison of NSIHT and NSHTP with different algorithmic parameters (ϵ, λ). The size of measurement matrix is 256×512, and the measurements are exact, i.e., $y = Ax^*$. 69
3.9. Comparison of NSHTP and some existing algorithms, such as ℓ_1-minimization, OMP, CoSaMP and SP. The size of generated Gaussian random matrix is 256×512. The number of iterations for CoSaMP, SP and NSHTP is 30, whereas OMP is terminated after k iterations. 70
3.10. Phase transition of IHT, HTP, NSIHT and NSHTP with respect to success rate for sparse signal recovery. 72
3.11. Phase transition of HTP and NSHTP with respect to time consumption. . 73

4.1. (a) Comparison of residual-reduction performances of several algorithms; (b) Residual change in the course of iterations using different thresholding operators. Experimental settings: $A \in \mathbb{R}^{256\times512}, y = Ax^*$ and $k = 70$. 96
4.2. Residual reduction by NTROT and NTROTP with $\epsilon = \sigma_1^2 + 1$ and different stepsize $\lambda = 1, 2, 5, 10$. Experimental settings: $A \in \mathbb{R}^{256\times512}, y = Ax^*$ and $k = 70$. 97
4.3. Residual reduction by NTROT and NTROTP with $\lambda = 10$ and different parameter ϵ: $\epsilon_1 = \epsilon^*$, $\epsilon_2 = 1.1\epsilon^*$, $\epsilon_3 = 1.5\epsilon^*$, $\epsilon_4 = 2\epsilon^*$. Experimental settings: $A \in \mathbb{R}^{256\times512}, y = Ax^*$ and $k = 70$. 97
4.4. Comparison of the average number of iterations required by different algorithms in terms of different sparsity levels. Experimental settings: the size of A is 256×512, $y = Ax^*$, and the ratio k/n varies from 0.01 to 0.35. . . 98
4.5. Comparison of the average number of iterations required by different algorithms in terms of different number of measurements. Experimental settings: $n = 500$, $y = Ax^*$, and the number of measurements m varies from 50 to 300. 99
4.6. Comparison of success frequencies of signal recovery via different algorithms 100
4.7. Comparison of success frequencies of signal recovery via different Newton-type algorithms 101
4.8. Phase transition of NSHTP and NTROTP with respect to success rate and time consumption under inexact measurements 102
4.9. Phase transition curve of NSHTP and NTROTP under inexact measurements in terms of success rates 103
4.10. Examples of using logistic model to estimate the success rate for NTROTP at a certain value of δ 103

5.1. Residual change in the course of iterations for PGROTP with different q, i.e., $q = k, 2k, 3k, n$. The size of measurement matrix is 500×1000, and the measurements are exact, i.e., $y = Ax^*$. 125
5.2. Comparison of the average number of iterations required by different q. The length of the target signal is 1000, and measurements are exact, i.e., $y = Ax^*$.126
5.3. Comparisons of success rate of sparse signal reconstruction between several algorithms. 127
5.4. Phase transition of ℓ_1-minimization, SP, PGROTP and ROTP2 algorithms in terms of success rates 129
5.5. Phase transition curve of ℓ_1-minimization, SP, PGROTP and ROTP2 under inexact measurements in terms of success rates 130
5.6. Time consumption of ℓ_1-minimization, SP, PGROTP and ROTP2 algorithms for successful recovery 131
5.7. Time consumption ratio of PGROTP and ROTP2 133

6.1. Residual change in the course of iterations between HTP, ROPGHT, ROPGHTP, PGROT and PGROTP algorithms. For partial gradient algorithms, $q = k$. 151
6.2. Residual change in the course of iterations between HTP, ROPGHT, ROPGHTP, PGROT and PGROTP algorithms. For partial gradient algorithms, $q = 2k$. 152
6.3. Phase transition of HTP and ROPGHTP ($q = k$) with different stepsizes in terms of success rates. 153
6.4. Phase transition curves of HTP and ROPGHTP ($q = k$) with $\lambda = 1$ in terms of different success rates. 154
6.5. Phase transition curves of HTP and ROPGHTP ($q = 2k$) with $\lambda = 1$ in terms of different success rates. 154

CHAPTER 1

INTRODUCTION

In signal processing, a commonly used technique to obtain the information of a signal that cannot be directly acquired is reconstructing the signal from a specific type of measurements. The classical Nyquist-Shannon sampling theory [102], which has dominated the field of signal processing for the past few decades, claims that the minimum sampling rate required to reconstruct a time-continuous signal successfully should be at least twice its highest frequency. However, perfect signal reconstruction is also possible without meeting such a criterion. For instance, the compressed sensing (CS) technique can recover a signal using a small number of linear measurements if the target signal is sparse or can be sparsely approximated. A signal is said to be sparse if the majority of its components are zero. In the real world, many types of signals can be sparsely represented or approximated, such as medical images, wireless communication signals, and so on. The CS technique, pioneered by Donoho [49] and Candés, Romberg, and Tao [29, 33] around 2006, breaks through the bottleneck of Nyquist-Shannon sampling theory in significantly reducing the required number of samples for signal reconstruction when the signal has an additional structure. Since then, a large amount of work has been published in both the fields of sparse optimization and signal processing. Moreover, sparse optimization problems arising from CS have been widely studied and applied in diverse areas, such as statistical regression [68], image processing [54, 21], pattern recognition [95], sparse channel estimation [10], to name but a few. Since around 2004, the CS theory has been systematically developed and

its application has been extended to a wide range of practical areas. Some excellent books have been published in this area including the ones by Foucart [63] and Zhao [136], Elad [54] and Eldar [56].

In the following sections, we introduce the fundamental models in CS, main contributions of the thesis, and the notations used throughout the thesis.

1.1 Sparse optimization problems

Linear measurement is one of the simplest methods to capture a signal, i.e., $y = Ax^*$, where $A \in \mathbb{R}^{m\times n}$ is usually called the measurement matrix (sensing matrix) and $y \in \mathbb{R}^m$ is called the measurements of an unknown signal $x^* \in \mathbb{R}^n$. Traditional signal processing theory and classic linear algebra indicate that, to guarantee the reconstruction success, the number of measurements should be at least equal to the length of x^*, i.e. $m \geq n$. Clearly, if the number of measurements is much lower than its length, i.e. $m \ll n$, it is impossible to recover x^* without additional information since an underdetermined linear system has infinitely many solutions. However, the purpose of CS technology is to realize this idea when the signal has a sparsity structure. An example of the compressed sensing model is shown in Fig. 1.1.

An essential task in CS is finding the sparsest solution of an underdetermined linear system, which can be formulated as

$$\min_x \{\|x\|_0 : y = Ax\}, \tag{1.1}$$

where the ℓ_0-norm $\|\cdot\|_0$ counts the number of nonzero entries of a vector. In the signal processing literature, the problem (1.1) is usually referred to as the standard ℓ_0-minimization. In general, the constraint $y = Ax$ represents the exact or noiseless measurements. In many practical situations, however the measurements may be contaminated by noises rendering inaccurate measurements $y := Ax + \nu$, where $\nu \in \mathbb{R}^m$ denotes the noise vector.

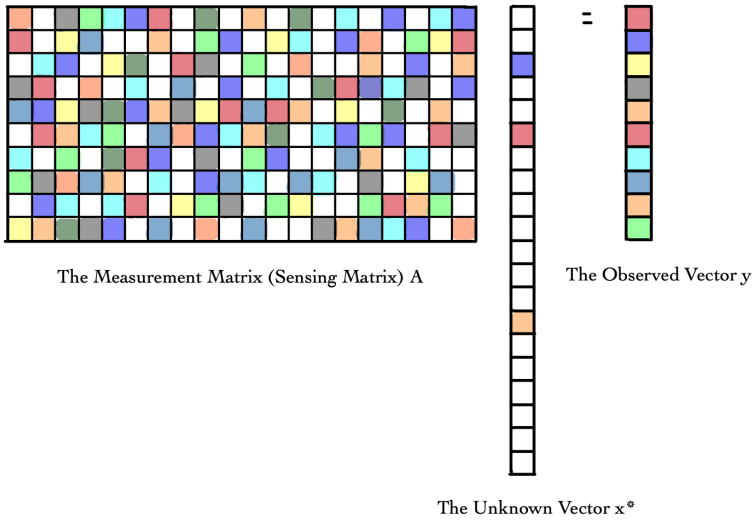


Figure 1.1: An example of compressed sensing model. The unknown signal is assumed to be sparse, containing few entries that are non-zero, whereas the measurement matrix is not necessarily sparse.

Suppose the noise level is bounded by $\eta > 0$, i.e., $\|\nu\|_2^2 \leq \eta$. Then the problem by taking into account the effect of noise can be formulated as the following quadratic constrained problem:

$$\min_x \left\{ \|x\|_0 : \|y - Ax\|_2^2 \leq \eta \right\}. \tag{1.2}$$

Depending on application scenarios, an alternative of the CS problems can be also modeled as

$$\min_x \left\{ \|y - Ax\|_2^2 : \|x\|_0 \leq k \right\}, \tag{1.3}$$

where the integer $k > 0$ is a given sparsity level. Throughout the dissertation, a vector x is said to be k-sparse if $\|x\|_0 \leq k$. Clearly, the purpose of (1.3) is to find a k-sparse vector at which the minimum of the objective can be achieved. The solution to (1.3) is the best k-term approximation of the signal to recover. A more general problem than (1.3) is the following sparsity constrained optimization problem:

$$\min_x \{ f(x) : \|x\|_0 \leq k \},$$

where $f(x)$ is a differentiable function. Although the measurement of the signal is linear,

the constructing process for the signal is highly nonlinear. Recent study shows that these sparsity optimization problems are generally NP-hard [63]. The main difficulty in solving these problems is locating the position of nonzero entries of the target signal, i.e., identifying the support of the signal.

Definition 1.1.1. *The* ***support*** *of a vector $x \in \mathbb{R}^n$ is the index set of its nonzero entries, i.e.,*

$$\mathrm{supp}(x) := \{i : x_i \neq 0, i \in [N]\},$$

where $[N] = \{1, \ldots, n\}$.

It is challenging to solve ℓ_0-minimization problems whether the sparsity level is known or not. For example, if we use the exhaustive search method, it would require checking all possible k columns of the matrix A to find the best representation of the signal, which is time-consuming. In addition, when the sparsity level of the target vector is unknown, problems (1.1) and (1.2) are more challenging to solve than the case where the sparsity level of the signal is clear. Moreover, the unknown support of the target signal also makes the problem difficult to solve. In fact, if the support is given, then the signal recovery can be achieved simply by solving a deterministic linear system or a least square problem.

Directly working on the ℓ_0-norm is hard. It makes sense to find some proper functions which can approximate the ℓ_0-norm in a certain sense. Two commonly used choices are the ℓ_1-norm, $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, which is the convex function closest to the ℓ_0-norm over the set $\|x\|_\infty \leq 1$, where $\|x\|_\infty = \max_{i\in[n]} |x_i|$, and the ℓ_q-norm, $\|x\|_q = \left(\sum_{i=1}^{n} |x_i|^q\right)^{1/q}$ with $0 < q < 1$, which can approximate the ℓ_0-norm very well. Fig. 1.2 demonstrates several norms in $\mathbb{R}$. It can be seen that, for the ℓ_q-norm, the smaller the q, the closer the ℓ_q-norm is to the ℓ_0-norm. Replacing the objective function in (1.2) by $\|x\|_1$ and $\|x\|_q$, respectively, leads to the ℓ_1-minimization problem

$$\min_x \{\|x\|_1 : \ \|y - Ax\|_2^2 \leq \eta\}, \tag{1.4}$$

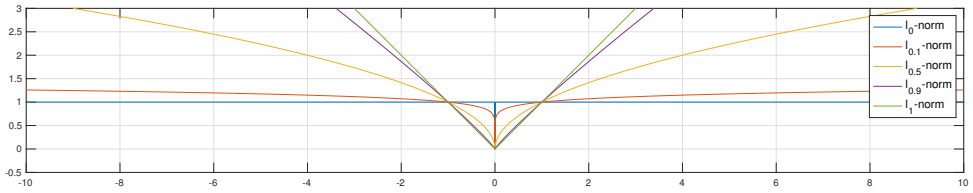


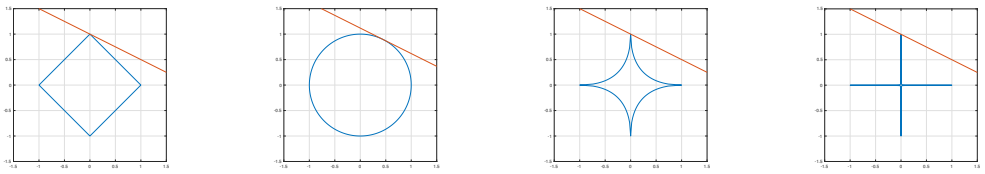
Figure 1.2: Graphs of the ℓ_0-norm, $\ell_{0.1}$-norm, $\ell_{0.5}$-norm, $\ell_{0.9}$-norm and the ℓ_1-norm. The horizontal axis represents the value of x, and the vertical axis represents the values of functions.

and the ℓ_q-minimization problem

$$\min_x \{\|x\|_q : \ \|y - Ax\|_2^2 \leq \eta\}.$$

When $\eta = 0$, these two problems reduce to the standard ℓ_1-minimization and ℓ_q-minimization problems.

The ℓ_1-norm and the ℓ_q-norm are introduced to replace the ℓ_0-norm because they still have certain ability to locate the sparse solution of the problem. Fig. 1.3 shows why the ℓ_1-norm and ℓ_q-norm are generally better than the ℓ_2-norm to approximate the ℓ_0-norm. Four different unit balls and an objective function are plotted in Fig. 1.3. It can be seen that in (a), (c) and (d) the point where the straight line touches the unit ball is sparse, whereas the point in (b) is not, since the intersection points are located along the axis. Compared to the ℓ_2-norm, the ℓ_1-norm and ℓ_q-norm ($0 < q < 1$) clearly have a higher probability for seeking sparse solutions.



(a) ℓ_1-unit ball (b) ℓ_2-unit ball (c) $\ell_{0.5}$-unit ball (d) $\ell_{0.1}$-unit ball

Figure 1.3: The intersection of a straight line and the unit ball defined by the ℓ_1-norm, ℓ_2-norm, $\ell_{0.5}$-norm and $\ell_{0.1}$-norm.

1.2 Main contribution

In this thesis, we focus on solving the specific optimization problem

$$\min_x \left\{ \|y - Ax\|_2^2 : \|x\|_0 \leq k \right\} \tag{1.5}$$

for sparse signal recovery, where $A \in \mathbb{R}^{m \times n}$ with $m \ll n$ is a measurement matrix, y denotes the measurements and k is a given sparsity level. This problem is a special case of the following problem

$$\min_x \{ f(x) : \|x\|_0 \leq k \}. \tag{1.6}$$

Plenty of methods for solving these problems have been proposed in the past two decades, including convex methods, heuristic methods, and thresholding algorithms. The detailed discussion for these methods can be found in Section 2.3.

Thresholding-type methods have gained a great deal of attention due to their low complexity, clear structures and easy implementations. The key step of thresholding methods for solving (1.6) can be simply stated as

$$x^{p+1} = \mathcal{T}_k(x^p + \lambda \, \mathbf{d}(x^p)), \tag{1.7}$$

where $\mathcal{T}_k(\cdot)$ represents a thresholding operator, $\lambda > 0$ is the stepsize and $\mathbf{d}(x^p)$ denotes a descent direction of the objective function at the current iterate x^p. Depending on the type of thresholding operators, thresholding methods can be classified roughly as hard-thresholding (see, e.g., [18], [17], [19], [16], [23]), soft-thresholding (see, e.g., [50], [58]) and optimal-thresholding (see, e.g., [134], [139]). The widely used search directions are the steepest descent direction (see, e.g., [18, 17, 19]) and the Newton-type direction (see, e.g., [84, 85, 145, 144]). Many existing algorithms are a mixture of thresholding and optimization methods.

Based on the general iterative scheme (1.7), various thresholding methods can be developed by selecting the thresholding operator and the search direction. This thesis is

dedicated to proposing four novel thresholding-type algorithms by incorporating the new search directions into the existing thresholding strategies. We now briefly describe these four approaches.

Newton-step-based hard thresholding

The hard thresholding denoted by $\mathcal{H}_k(\cdot)$ retains the largest k magnitudes of a vector and zeroes out others. A class of traditional hard thresholding algorithms (see Section 2.3.2 in detail) adopts the steepest descent direction as a search direction, and its key step can be written as

$$x^{p+1} = \mathcal{H}_k(x^p + \lambda\ \mathbf{d}_{\mathrm{SD}}(x^p)), \tag{1.8}$$

where λ is a stepsize, and$\mathbf{d}_{\mathrm{SD}}(x^p)$ denotes the gradient descent direction at x^p of $f(x)$, namely, $\mathbf{d}_{\mathrm{SD}}(x^p) = -\nabla f(x^p)$. The theoretical performance of this method can be shown under some standard assumptions, and empirical results indicate that the numerical performance of this method can be also guaranteed if the stepsize is carefully chosen. However, the practical performance of this algorithm is quite sensitive to the choice of stepsize, and sometimes it is not easy to find a proper stepsize when solving specific problems. Classical optimization theory indicates that Newton's method usually has a faster local convergence than the steepest descent method when solving an unconstrained nonlinear optimization problem. Motivated by this fact, we propose the so-called Newton-step-based hard thresholding method to which the key step can be formulated as

$$x^{p+1} = \mathcal{H}_k(x^p + \lambda\ \mathbf{d}_{\mathrm{Newton}}(x^p)), \tag{1.9}$$

where $\mathbf{d}_{\mathrm{Newton}}(x^p)$ is a Newton-type direction at x^p. It is worth mentioning that the Hessian of the objective function in (1.5) is singular, so the standard Newton's method cannot be applied. But one can still develop a Newton-like method by introducing a certain modification to the Hessian, leading to the algorithms proposed in Chapter 3.

Newton-type optimal k-thresholding

As pointed out in [134] and [139], the hard thresholding may cause a dramatic increase of the objective function compared to the objective value at its unthresholded counterpart. This is because the hard thresholding is independent of the objective function. Recently, Zhao [134, 139] has proposed a novel thresholding technique, called optimal k-thresholding to improve the performance of (1.8), resulting in the iterative scheme

$$x^{p+1} = \mathcal{Z}_k^{\#}(x^p + \lambda\ \mathbf{d}_{\mathrm{SD}}(x^p)),$$

where $\mathcal{Z}_k^{\#}(\cdot)$ denotes the optimal k-thresholding operator which selects the k entries that minimizes the objective function. Numerical results in [134] indicate that the optimal-thresholding-type algorithms usually transcend the traditional hard thresholding methods for sparse signal recovery, and they are comparable to ℓ_1-minimization which is one of the efficient approaches for ℓ_0-minimization problems.

Motivated by the appreciable performance of Newton-type methods and the optimal k-thresholding technique, we propose the so-called Newton-type optimal k-thresholding algorithms

$$x^{p+1} = \mathcal{Z}_k^{\#}(x^p + \lambda\ \mathbf{d}_{\mathrm{Newton}}(x^p)), \tag{1.10}$$

where $\mathbf{d}_{\mathrm{Newton}}(x^p)$ is the same type of direction used in (1.9). Empirical results show that the Newton-type optimal thresholding methods often perform better than Newton-type hard thresholding methods in sparse signal recovery.

Partial gradient optimal k-thresholding

In a typical thresholding algorithms, the iterate generated by a specific thresholding technique is by adopting the full steepest descent direction or a Newton-like direction, e.g., (1.8), (1.9) and (1.10). A large part of a vector has been discarded when performing a hard thresholding. This motivates us to propose the so-called partial gradient optimal

k-thresholding algorithm, which is an integration of the partial gradient and optimal k-thresholding. The key iterative scheme can be described as

$$x^{p+1} = \mathcal{Z}_k^{\#}\left(x^p + \lambda\ \mathcal{H}_q\left(\mathbf{d}_{\mathrm{SD}}(x^p)\right)\right), \tag{1.11}$$

where the integer $q > 0$ is used to control the number of elements extracted from the steepest descent direction $\mathbf{d}_{\mathrm{SD}}(x^p)$. The partial direction $\mathcal{H}_q\left(\mathbf{d}_{\mathrm{SD}}(x^p)\right)$ is still a descent direction, except that it does not make the objective function decrease so rapidly as the full direction. However, the use of this direction can benefit the thresholding step as only a small number of elements will be discarded, and thus the objective value would not increase much. Numerical experiments indicate that the partial gradient could maintain the similar performance of the optimal thresholding algorithms proposed in [134] while using much less time.

Optimal partial gradient hard thresholding

We also study a combination of the partial gradient method and the hard thresholding strategy. Different from the partial gradient direction used in (1.11), we adopt the so-called optimal partial gradient direction which is also extracted from the full gradient such that the objective function at the current iterate is minimized. The key step can be stated as

$$x^{p+1} = \mathcal{H}_k\left(x^p + \lambda\ \mathbf{d}_{\text{opt-par}}(x^p)\right),$$

where the optimal partial gradient direction is given by $\mathbf{d}_{\text{opt-par}} = w^p \otimes \mathbf{d}_{\mathrm{SD}}$ with

$$w^p := \arg\min_w \{f\left(x^p + \lambda(w \otimes \mathbf{d}_{\mathrm{SD}}(x^p))\right) : \mathbf{e}^T w = q, \mathbf{e} \in \{0,1\}^n\},$$

where $\otimes$ denotes the Hadamard product. The integer $q > 0$ is used to adjust the number of elements of the partial direction. Empirical results indicate that such algorithms are less sensitive to the choice of stepsize λ than the traditional hard thresholding algorithm (1.8).

The major work of the thesis is to show the efficiency of the above algorithms for sparse signal recovery from both theoretical and practical view points. Theoretical analysis including convergence analysis is carried out in terms of the restricted isometry property (RIP) which is one of the fundamental tools in analyzing compressed sensing algorithms. We have shown that the proposed algorithms are guaranteed to recover sparse signals when the RIP conditions are satisfied and algorithmic parameters are properly chosen. Numerical experiments are carried out via synthetic data to demonstrate the efficiency and robustness of the proposed algorithms for sparse signal recovery.

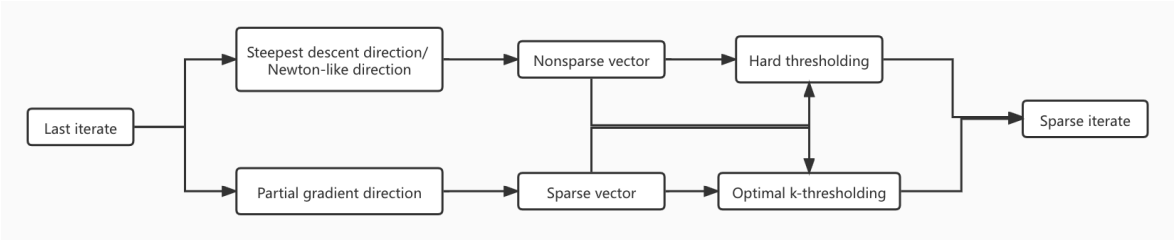


Figure 1.4: Process of thresholding-type methods

1.3 Organization of thesis

The rest of the thesis is organized as follows. In the next section, we introduce some notation used throughout the thesis. Chapter 2 introduces fundamental assumptions, optimization methods and compressed sensing algorithms that are closely related to the thesis. The proposed algorithms are discussed in Chapters 3-6, respectively. The conclusions and future work are summarized in the last chapter.

1.4 Notation

For clarity, we introduce some of the notation used in the thesis.

1.4.1 Sets, vectors and matrices

The thesis is based on the real space $\mathbb{R}$. We use $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ to denote the vector space and matrix space, respectively. The identity matrix is denoted by I, $[N]$ the index set $\{1,\ldots,n\}$, and $\emptyset$ the empty set. Given a set $\Omega \subseteq [N]$, $\overline{\Omega} := [N]\backslash\Omega$ is the complement set in $[N]$ of Ω, and $\mathrm{card}(\Omega)$ is the cardinality of the set Ω. For a given vector $x \in \mathbb{R}^n$, $x_\Omega \in \mathbb{R}^n$ denotes the vector derived from x by retaining the entries indexed by Ω and zeroing out others. The vector $[x_\Omega] \in \mathbb{R}^{\mathrm{card}(\Omega)}$ denotes the subvector of x which retains the entries indexed by Ω and deletes others. $|a|$ denotes the absolute value of number a. Given two vectors $x, y \in \mathbb{R}^n$, the inner product is $\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i = x^T y$, and the Hadamard product is $x \otimes y := [x_1 y_1, \ldots, x_n y_n]^T$. A vector x is called k-sparse if $\|x\|_0 \leq k$.

For a matrix A, we use A^T to represent its transpose. We say A is symmetric if A is a square matrix and satisfies $A^T = A$. The matrix $A \in \mathbb{R}^{n\times n}$ is orthogonal if $A^T A = AA^T = I$, which is equivalent to $A^T = A^{-1}$, where A^{-1} denotes the inverse of A. A diagonal matrix $A \in \mathbb{R}^{n\times n}$ is denoted by $\mathrm{diag}(a_1, \ldots, a_n)$, where $a_i, i = 1, \ldots, n$ is the i-th entry on the diagonal. Given a matrix $A \in \mathbb{R}^{m\times n}$, $A_\Omega \in \mathbb{R}^{m\times n}$ denotes the matrix derived from A by retaining its columns indexed by set Ω and zeroing out others. $[A_\Omega]$ denotes the submatrix of A which retains the columns indexed by Ω and deletes others. The matrix $A_{\Omega\times\Omega} \in \mathbb{R}^{m\times n}$ is obtained by zeroing out the columns and rows of $A \in \mathbb{R}^{m\times n}$ indexed by $\overline{\Omega}$. $[A_{\Omega\times\Omega}] \in \mathbb{R}^{\mathrm{card}(\Omega)\times\mathrm{card}(\Omega)}$, denotes the submatrix of A obtained by deleting the rows and columns indexed by $\overline{\Omega}$, if $\mathrm{card}\,\Omega \leq m$ and $\Omega \subseteq [m]$.

1.4.2 Norms for vectors and matrices

Definition 1.4.1 (vector norm). For any $x, y \in \mathbb{R}^n$ and any $c \in \mathbb{R}$, a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is a norm (vector norm) if the following properties are satisfied:

(i) Nonnegativity: $\|x\| \geq 0$;

(ii) Homogeneity: $\|cx\| = |c|\|x\|$;

(iii) Triangle inequality: $\|x+y\| \leq \|x\| + \|y\|$.

The ℓ_p-norm of a vector $x \in \mathbb{R}^n$ is defined as $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$. Several commonly used vector norms include the ℓ_1-norm $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, the ℓ_2-norm (Euclidean norm) $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$ and the ℓ_∞-norm (infinity norm) $\|x\|_\infty = \max_{i\in[N]} |x_i|$. When $0 < p < 1$, the function $\|x\|_p$ violates the triangle inequality property, so it is not a norm but referred to as a quasi-norm. The ℓ_0-norm $\|\cdot\|_0$, counting the number of nonzero entries, is also not a norm since it does not satisfy the homogeneity property. In particular, it can be approximated as

$$\|x\|_0 = \lim_{p\to 0} \sum_{i=1}^{n} |x_i|^p.$$

Definition 1.4.2 (matrix norm). For any $A, B \in \mathbb{R}^{m\times n}$ and any $c \in \mathbb{R}$, a function $\|\cdot\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ is a norm (matrix norm) if the following properties are satisfied:

(i) Nonnegativity: $\|A\| \geq 0$;

(ii) Homogeneity: $\|cA\| = |c|\|A\|$;

(iii) Triangle inequality: $\|A+B\| \leq \|A\| + \|B\|$;

The ℓ_p-norm for a matrix $A \in \mathbb{R}^{n\times n}$ is defined as

$$\|A\|_p := \sup_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p \leq 1} \|Ax\|_p.$$

The commonly used ℓ_2-norm of a matrix A is given as $\|A\|_2 = \sqrt{\lambda_{\max}\left(A^T A\right)} = \sigma_{\max}(A)$, where $\lambda_{\max}(A^T A)$ denotes the maximum eigenvalue of matrix $A^T A \in \mathbb{R}^{n\times n}$ and $\sigma_{\max}(A)$ represents the maximum singular value of A. A singular value of a matrix A can be derived by taking the positive square root of an eigenvalue of the symmetric matrix $A^T A$ or AA^T, i.e., $\sigma_i = \sqrt{\lambda_i\left(A^T A\right)} = \sqrt{\lambda_i\left(AA^T\right)}$, $i = 1, \ldots,$ where $\lambda_i(A^T A)$ is the i-th eigenvalue of $A^T A$. Other matrix-norm definitions and properties, such as ℓ_1- and ℓ_∞-norm, can be found in the literature [71, Chapter 5] and [63, Appendix A].

The singular value decomposition (SVD) is one of the commonly used methods to decompose a matrix, which is defined as follows.

Proposition 1.4.3 (Proposition A.13 of [63]). For a matrix $A \in \mathbb{R}^{m\times n}$ with $m \leq n$, there exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and nonnegative numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$, called the singular values of A, such that $A = U\Sigma V^T$, where

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{n-m,m} \\ & & \sigma_m & \end{bmatrix}. \tag{1.12}$$

1.4.3 Convex sets and functions

Definition 1.4.4 (see Section 2.1.4 of [101]). A set C is **convex** if for any $x, y \in C$ and $0 \leq \theta \leq 1$, we have $\theta x + (1-\theta)y \in C$.

This implies that the line segment connecting x and y is within the set C. The convex hull of a set C is the tightest convex set containing C.

Definition 1.4.5 (Definition 3.1.1 of [101]). Let $\mathrm{dom}(f)$ denote the domain of f. A function $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** if $\mathrm{dom}(f)$ is a convex set and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y),$$

for any $x, y \in \mathrm{dom}(f)$ with $0 \leq \theta \leq 1$

A function f is **concave** if $-f$ is convex.

CHAPTER 2

LITERATURE REVIEW

Compressed sensing refers to three main aspects, signal sparse representation, signal measurements, and reconstruction algorithms. In the thesis, we focus on the intrinsic problems, theoretical assumptions and reconstruction methods for signal recovery. This chapter is organized as follows. Section 2.1 introduces the theoretical properties of a measurement matrix that are essential for recovery success. Section 2.2 reviews two traditional methods for solving nonlinear optimization problems. Section 2.3 reviews the mainstream sparse signal recovery algorithms at the current stage of development.

2.1 Properties of measurement matrices

We first look at how many measurements are necessary for successful signal recovery.

2.1.1 Minimal number of measurements

In CS scenarios, the number of measurements m is assumed to be far less than the vector length n, i.e., $m \ll n$. However, the number m cannot be arbitrarily small. First of all, in order to possibly recover a k-sparse solution for a linear system, it is necessary to make sure that the linear system $y = Ax$ does not have multiple such sparse solutions, i.e., $Ax' \neq Ax''$ for any two k-sparse vectors x' and x''. This is equivalent to $A(x' - x'') \neq 0$ which means that there exists no $2k$-sparse vector $z \in \mathbb{R}^n$ in the null space of A. One way

to guarantee this is by using the spark of a matrix.

Definition 2.1.1 [51]. The spark of a matrix $A \in \mathbb{R}^{m\times n}$ is the smallest number of linearly dependent columns of A.

The following theorem implies the minimal number of measurements for the uniqueness guarantee of the solution to the linear system.

Theorem 2.1.2 (Corollary 1 of [51]). For any vector $y \in \mathbb{R}^m$, there exists at most one k-sparse vector $x \in \mathbb{R}^n$ such that $y = Ax$ if $\text{spark}(A) > 2k$.

This theorem implies the minimum number of the measurements m should satisfy that $m > 2k$, since the $\text{spark}(A) \leq m+1$. This also explains why an algorithm is usually difficult to recover the signal when $k > m/2$. Sufficient conditions concerning the uniqueness of the solution to linear system in terms of some other tools, such as mutual coherence and Babel function, can be found in [136, Chapter 2]. However, $2k$ measurements may not be enough to guarantee the recovery of a signal. More requirements on the number of measurements can be found in the CS literature. In what follows, we now introduce one of such conditions in terms of the concept of instance optimality.

Definition 2.1.3 (Definition 11.1 of [63]). Let Δ denote a recovery scheme $\Delta : \mathbb{R}^m \to \mathbb{R}^n$ and let Σ_k denote the set of k-sparse vectors. Given $q \geq 1$, a pair of measurement matrix $A \in \mathbb{R}^{m\times n}$ and Δ is called ℓ_q-instance optimal of order k with constant $C > 0$ if

$$\|x - \Delta(Ax)\|_q \leq C\sigma_k(x)_q \quad \text{for all } x \in \mathbb{R}^n,$$

where

$$\sigma_k(x)_q = \min_{z\in\Sigma_k} \|x - z\|_q.$$

The ℓ_2-instance ($q = 2$) is commonly used in the literature, which indicates that the distance between the recovered vector $\Delta(Ax)$ and the target vector x stays under control. To achieve the instance optimality above, other conditions on the number of measurements are required.

Theorem 2.1.4 (Theorem 1.13 of [56]). If the pair (A, Δ) is ℓ_2-instance optimal of order $k \geq 1$ with constant C, then

$$m \geq \left(1 - \sqrt{1 - 1/C^2}\right) n,$$

where $C \geq 1$.

Theorem 2.1.5 (Theorem 11.6 of [63]). Let e denote Euler's number. If the pair (A, Δ) is ℓ_1-instance optimal of order $k \geq 1$ with constant C, then

$$m \geq ck \ln(en/k)$$

for some constant c depending only on C.

However, it is far from enough to guarantee the stable and robust reconstruction via an algorithm only by conditioning the number of measurements. Further assumptions are required for this.

2.1.2 Restricted isometry property (RIP)

The restricted isometry property (RIP), also known as uniform uncertainty principle, was first introduced by Candès and Tao [31]. It is closely related to the restricted isometry constant (RIC), a constant with respect to a rectangular matrix $A \in \mathbb{R}^{m \times n}$ with $m \ll n$.

Definition 2.1.6 [31]. The k-th restricted isometry constant (RIC) δ_k of a matrix $A \in \mathbb{R}^{m \times n}$ is defined as the smallest $\delta_k \in (0, 1]$ such that

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

for all k-sparse vectors x.

A matrix is said to be a RIP matrix if the relation (2.1) is satisfied. In this case, we say matrix A satisfies the RIP of order k. An equivalent expression of the constant δ_k is

given as

$$\delta_k = \max_{\Omega \subset [N], \mathrm{card}(\Omega) \leq k} \left\| [A_\Omega]^T [A_\Omega] - I \right\|_2,$$

where $\mathrm{card}(\Omega)$ denotes the cardinality of set Ω and I is the identity matrix. In the formula above, $[A_\Omega] \in \mathbb{R}^{m \times k}$ denotes the submatrix of A which retains the columns of A indexed by Ω and deletes others. This property indicates that the eigenvalues of matrix $[A_\Omega]^T [A_\Omega] \in \mathbb{R}^{k \times k}$ lie in the interval $[1 - \delta_k, 1 + \delta_k]$.

There are two questions about the RIP that remain open. Firstly, testing the RIP of a matrix is a combinatorial problem [5, 118]. There are $\binom{k}{n}$ combinations of k columns of matrix $A \in \mathbb{R}^{m \times n}$, which is a large number especially when k and n are large. Detailed discussion of the computational complexity of RIP can be found in [110]. Secondly, it is challenging to construct a deterministic matrix that satisfies the RIP. The sensing matrix used for sparse signal reconstruction can be random matrices or deterministic matrices. Probability theory indicates that certain types of random matrices can satisfy the RIP with an overwhelming probability, such as Bernoulli random matrices, Gaussian random matrices, sub-Gaussian random matrices and partial random Fourier matrices (see [6] and [63, Chapter 9]).

Theorem 2.1.7 (Corollary 9.3 in [63]). Let $A \in \mathbb{R}^{m \times n}$ be a Gaussian or Bernoulli random matrix. The restricted isometry constant of $A/\sqrt{m}$ satisfies $\delta_k < \delta$ with probability at least $1 - \epsilon$, if

$$m \geq C \delta^{-2} \left(k \ln(en/k) + \ln\left(2\epsilon^{-1}\right) \right),$$

where C is a universal constant and e denotes Euler's number.

The RIP of other matrices with certain structures are also discussed in the literature, including Chirp sensing matrices [4], second-order Reed-Muller sensing matrices [72], Binary Bose-Chaudhuri-Hocquenghem (BCH) matrices [1], matrices based on mutual coherence [2], matrices arising from near orthogonal systems [76], matrices satisfying a statistical isometry property [44] and matrices designed via algebraic curves over finite fields [48]. A good survey can be found in [92].

Remark 2.1.1

The formula (2.1) is often called a standard version of the restricted isometry property (RIP) in the literature. A more general form of (2.1) is stated as

$$\alpha\|x\|_2^2 \leq \|Ax\|_2^2 \leq \beta\|x\|_2^2, \tag{2.2}$$

where $0 < \alpha \leq \beta < +\infty$ and x is an arbitrary k-sparse vector. The inequality (2.2) can be simply scaled up or down to derive (2.1), for instance, multiplying A by $\sqrt{2/(\beta+\alpha)}$ yields $\delta_t = (\beta-\alpha)/(\beta+\alpha)$ [56, Chapter 1]. In addition to the symmetric versions of RIP (2.1) and (2.2), the asymmetric version is also used in the analysis of CS algorithms, see, e.g., [13]. For the convenience of discussion, we will use the standard form (2.1) to perform the analysis of proposed algorithms in the thesis.

2.1.3 Other properties

Mutual coherence, introduced by Donoho and Hu [52], is a much simpler method to measure the quality of a measurement matrix than the RIP. The mutual coherence $\mu(A)$ of a matrix A is defined as

$$\mu(A) \coloneqq \max_{1\leq i\neq j\leq n} \frac{|a_i^T a_j|}{\|a_i\|_2\|a_j\|_2},$$

where $a_i, i = 1, \ldots, n$ denotes the columns of A. In general, a reconstruction scheme performs better with smaller mutual coherence, which means that the columns of A have lower correlations. However, the mutual coherence cannot be arbitrarily small. A well-known lower bound for μ is the Welch bound [119], given as

$$\mu(A) \geq \sqrt{\frac{n-m}{m(n-1)}}.$$

The mutual coherence and related concepts have been used to analyze the uniqueness of ℓ_0-minimization problems [136, Section 1.3], and to develop sufficient conditions for

the guaranteed performance of several types of algorithms, such as orthogonal matching pursuit (OMP), basis pursuit (BP) and thresholding-type algorithms (see [63, Chapter 5]).

The null space property (NSP) in the literature is a necessary and sufficient condition for the exact recovery of sparse signals via the standard ℓ_1-minimization (1.4) (see, e.g., [52], [39], [132] and [63, Chapter 4]). The NSP with respect to a matrix $A \in \mathbb{R}^{m\times n}$ and a set $\Omega \subset [N]$ is defined as follows.

Definition 2.1.8 [63]. A matrix $A \in \mathbb{R}^{m\times n}$ is said to satisfy the null space property (NSP) relative to a set $\Omega \subset [N]$ if

$$\|v_\Omega\|_1 < \|v_{\overline{\Omega}}\|_1,$$

for all $v \in \ker(A)$, where $\ker(A) = \{u \in \mathbb{R}^n | Au = 0\}$.

The RIP is more restrictive than the NSP, but the former has a broader application in compressed sensing since it can easily be used to analyze the reconstruction algorithm in noisy environments [25]. More discussion about the relationship and the gap between NSP and RIP can be found in [39] and [110].

In addition, Zhao [135] has introduced another matrix-property that is less restrictive than RIP and NSP, called range space property (RSP) of A^T, which provides a broad understanding of the equivalence between ℓ_0-minimization and ℓ_1-minimization (see Definition 1.1 in [135] or Definition 3.1.1 in [136]), and the stability of convex optimization algorithms for signal recovery. The RSP-based analysis work can be found in [138], [133] and [121].

2.2 Optimization methods

We now review the well-known steepest descent method and Newton's method, which are the key optimization methods related to the signal recovery algorithms studied in the thesis. A more detailed introduction of optimization methods and convex optimization

theory can be found in the books by Boyd and Vandenberghe [24], Nocedal and Wright [94], Rockafellar [101] and Beck [8].

2.2.1 Steepest descent method

The steepest descent method, also known as the gradient descent method, is to find local minima of unconstrained problem $\min_x f(x)$ according to the first-order information of the objective function f. Given a continuous function f and a point x, the first-order approximation of $f(x)$ around x is

$$f(x + \lambda d) \approx f(x) + \lambda \nabla f(x)^T d, \tag{2.3}$$

where $\nabla f(x)$ denotes the gradient of f, and λ is the stepsize. The direction d is said to be a descent direction at x if $\nabla f(x)^T d < 0$ which implies $f(x + \lambda d) < f(x)$ for small λ. Finding the steepest descent direction amounts to minimizing (2.3), i.e.,

$$\min_d \left\{ \nabla f(x)^T d : \|d\|_2 = 1 \right\}.$$

Clearly, the optimal solution of the problem above is $d = -\dfrac{\nabla f(x)}{\|\nabla f(x)\|_2}$, which lies in the same direction of $-\nabla f(x)$. In general, the negative gradient $-\nabla f(x)$ is regarded as the steepest descent direction. The full description of the steepest descent method is stated as Algorithm 2.1.

Algorithm 2.1 Steepest Descent Method

- Given a current point x^p.
- Compute the steepest descent direction $-\nabla f(x^p)$.
- Choose the stepsize $\lambda > 0$.
- Update the new point $x^{p+1} = x^p - \lambda \nabla f(x^p)$.

The choice of stepsize λ plays a key role in Algorithm 2.1. Line search methods are

widely used to determine the stepsize λ. A simple choice for λ is the positive number that minimizes $f(x)$ along the direction $-\nabla f\left(x^p\right)$ at current iterate x^p, i.e.,

$$\lambda := \underset{\hat{\lambda}:\hat{\lambda}>0}{\arg\min} f(x^p - \hat{\lambda}\nabla f\left(x^p\right)).$$

Many other line search rules can be also used to find the λ, including Armijo rule [104], Goldstein rule [107] and Wolfe rule [128]. The analysis for the global convergence of gradient methods can be found in the literature [93, 127, 97, 107]. In the following chapters, we will see the importance of the stepsize in compressed sensing algorithms.

Over the past few decades, many modifications of the steepest descent method have been proposed, such as the two-point step size gradient method [7], conjugate gradient method [69, 106], stochastic gradient descent method [22], and so on. These methods have been widely used in various application scenarios, e.g., machine learning [122] and statistical learning [68]. These algorithms maintain their popularity in solving various practical optimization problems because of their low computational cost especially when dealing with large-scale problems.

2.2.2 Newton's method

Newton's method is derived from the second-order Taylor approximation of the objective function f, which is also used to solve the problem $\min_x f(x)$. Given a twice continuously differentiable function $f(x)$, the second-order approximation of $f(x+\alpha d)$ around the point x is:

$$f(x+\alpha d) \approx f(x) + \alpha\nabla f(x)^T d + \frac{\alpha^2}{2} d^T\ \nabla^2 f(x)\ d, \tag{2.4}$$

where $\nabla^2 f(x)$ denotes the Hessian matrix of $f(x)$, and α is the stepsize. When the Hessian $\nabla^2 f(x)$ is nonsingular, the minimizer of the right-hand side of (2.4) with respect to αd is given by $\alpha d = -\left(\nabla^2 f(x)\right)^{-1}\nabla f(x)$, which is usually called the Newton search direction and denoted as Δ_{Newton}. The pure form of Newton's method is defined as Algorithm 2.2. In particular, if the Hessian is positive definite, i.e., $\nabla^2 f(x) \succ 0$, the Newton step is

actually a descent direction since

$$\nabla f(x)^T \Delta_{\text{newton}} = -\nabla f(x)^T \left(\nabla^2 f(x)\right)^{-1} \nabla f(x) < 0.$$

Algorithm 2.2 Newton's Method (pure form)

- Given current point x^p.
- Compute the Newton step: $\Delta_{\text{newton}} := -\left(\nabla^2 f(x^p)\right)^{-1} \nabla f(x^p)$.
- Update the new point: $x^{p+1} = x^p + \Delta_{\text{Newton}}$.

Newton's method usually has a local quadratic convergence order when the current point is close to a local minimum, but it has several drawbacks as pointed in [67], for instance, it cannot distinguish the maxima and minima due to the Hessian $\nabla^2 f(x)$ not being positive definite, and Algorithm 2.2 cannot be applied when the Hessian is singular. Thus some modifications are usually made to Algorithm 2.2.

i. *Damped Newton's method.* Taking the stepsize λ into account leads to

$$x^{p+1} = x^p - \lambda \left(\nabla^2 f(x^p)\right)^{-1} \nabla f(x^p),$$

where $\lambda > 0$ can be either fixed or determined by a line search method. Such a modification allows Newton's method to work from any starting point and converges to a local minimum.

ii. *Modified Newton's method.* The Hessian of a twice differentiable continuous function f is not always invertible. In this case, a simple approach is to perturb the Hessian $\nabla^2 f(x^p)$ such that $\nabla^2 f(x^p) + \epsilon I$ is positive definite, where I denotes the identity matrix and $\epsilon > 0$ is a parameter. Thus the iterative step becomes

$$x^{p+1} = x^p - \lambda \left(\nabla^2 f(x^p) + \epsilon I\right)^{-1} \nabla f(x^p).$$

It is easy to verify $-\nabla f(x)^T \left(\nabla^2 f(x) + \epsilon I\right)^{-1} \nabla f(x) < 0$ when the function $f(x)$ is convex, therefore the modified Newton step is still a descent direction when $\nabla f(x) \neq 0$. Other modifications of the Hessian can be found in [94, Section 3.4].

iii. *Quasi-Newton method.* This is also a modification of Newton's method by replacing the Hessian with a positive definite matrix which can be seen as a certain approximation of the Hessian. The key iteration of the quasi-Newton method is written as

$$x^{p+1} = x^p - \lambda Q^p \nabla f(x^p),$$

where the positive definite matrix Q^p is updated iteratively by a quasi-Newton formula. More details about this method can be found in [94, Chapter 6-7].

2.3 Compressive sensing algorithms

At the current stage of development, the main compressed sensing algorithms for sparse signal recovery can be classified into three types: convex methods, thresholding algorithms and heuristic algorithms. Each category of methods contains a number of variants, for example, see Fig. 2.1. Here we only introduce some CS algorithms that are closely related to the thesis.

2.3.1 The ℓ_1-minimization

The ℓ_1-minimization

$$\min_x \{\|x\|_1 : y = Ax\}, \tag{2.5}$$

is one of the popular approaches for solving ℓ_0-minimization problems. It is one of the convex optimization methods. The problem (2.5) and the following more general problem

$$\min_x \left\{\|x\|_1 : \|y - Ax\|_2^2 \leq \eta\right\}. \tag{2.6}$$

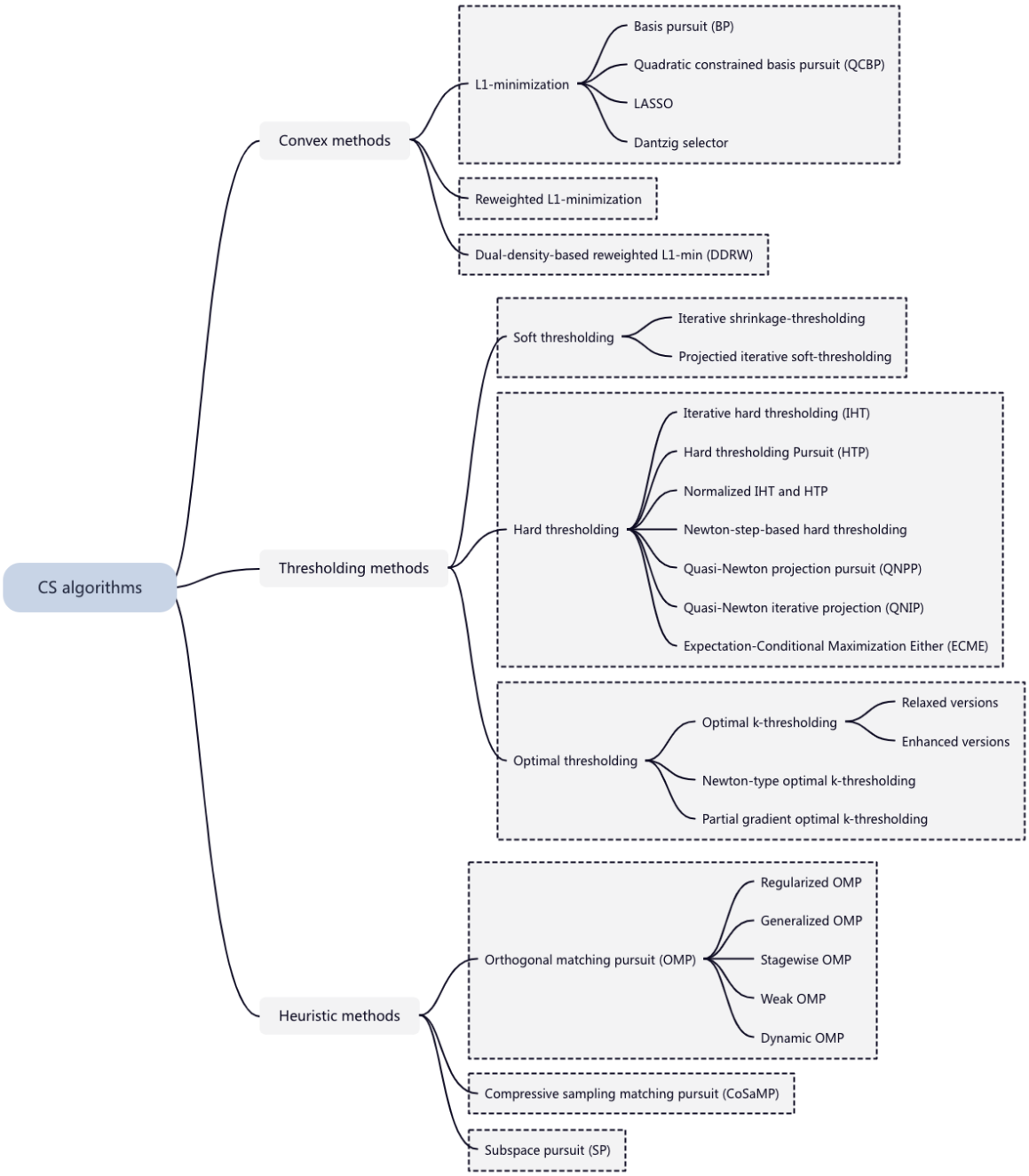


Figure 2.1: Classification of mainstream compressed sensing algorithms

are often called basis pursuit (BP) [38]. Due to the convexity, the classical convex optimization algorithms can be used to solve the ℓ_1-minimization, such as interior-point method, homotopy method [120], Chambolle and Pock's primal-dual algorithm [35] and iteratively reweighted least squares (IRLS) algorithm [70, 43].

Depending on applications, some variants of the problem (2.6) are also widely studied in the literature. For instance, the so-called LASSO (least absolute shrinkage and selection operator) [109]

$$\min_x \ \|y - Ax\|_2^2 + \alpha\|x\|_1,$$

where $\alpha > 0$ is a parameter, which can be also stated as

$$\min_x \left\{\|y - Ax\|_2^2 : \|x\|_1 \leq \tau\right\},$$

where τ is an appropriate parameter. The so-called Dantzig selector [28]

$$\min_x \left\{\|x\|_1 : \|A^T(Ax - y)\|_\infty \leq \tau\right\},$$

for some parameter $\tau > 0$, is also studied by some researchers.

Using ℓ_1-minimization for sparse recovery, one must ensure that the solution of (1.1) coincides with that of (2.5). In other words, it is important to guarantee that the problem (2.5) must admit a unique solution which is the same as the unique or one of the solutions of (1.1) (see definition of equivalence and strong equivalence in [136, Chapter 3]). Zhao [136, Chapter 2] has derived a necessary and sufficient condition for the solution uniqueness of the solution to (2.5) by utilizing the classic strict complementary slackness property (SCSP) of linear programming. Other sufficient conditions include mutual coherence [55, 51], restricted isometry property (RIP) [31, 29, 30, 46, 62, 59, 87, 27], null space property (NSP) [131] and range space property (RSP) of A^T [135, 133].

Remark 2.3.1

The RIP-based conditions for the efficiency of ℓ_1-minimization can be summarized as follows.

In [30], Candès has proven that if the restricted isometry constant (RIC) of order $2k$ satisfies $\delta_{2k} < \sqrt{2}-1 \approx 0.4142$, the solution to ℓ_1-minimization (2.5) is that of the ℓ_0-minimization problem (1.1). The upper bound of δ_{2k} has improved to $\delta_{2k} < 2(3-\sqrt{2})/7 \approx 0.4531$ by Foucart and Lai [62] in 2009, $\delta_{2k} < 0.4652$ by Foucart [59] in 2010, and $\delta_{2k} < 0.4931$ by Mo and Li [87] in 2011. Later, Cai and Zhang [27] have improved it to $\delta_{2k} < 1/\sqrt{2}$. Davies and Gribonval [46] have pointed out that ℓ_1-minimization might fail to recover signals when δ_{2k} is close to $1/\sqrt{2}$.

Remark 2.3.2

Reweighted ℓ_1-minimization [32, 138] is an enhanced version of ℓ_1-minimization, which seeks sparse solution by solving a sequence of weighted ℓ_1-minimization problems, i.e.,

$$\min_x \{\|W^p x\|_1 : y = Ax\}, \tag{2.7}$$

where W^p is a diagonal matrix whose entries are determined by a merit function for sparsity (a certain approximation of the ℓ_0-norm) at the current iterate. Problem (2.7) can be recast as a linear program, thus its dual problem can be derived. Motivated by this, dual-density-based reweighted ℓ_1-minimization (DDRW) [137, 140] has been proposed by working on the dual problem and using complementary to recover the sparse signal through the dense slack variables of the dual problem. A systematic discussion of these two methods can be found in [136, Chapter 7-8].

2.3.2 Hard thresholding algorithms

Thresholding methods are designed to solve the sparse optimization problem

$$\min_x \left\{ \frac{1}{2}\|y - Ax\|_2^2 : \|x\|_0 \le k \right\}, \tag{2.8}$$

which is clearly equivalent to (1.3). Traditional hard thresholding methods adopt the steepest descent direction as a search direction in the iterative step. The gradient of

$f(x) = \frac{1}{2}\|y - Ax\|_2^2$ is given by $\nabla f(x) = -A^T(y - Ax)$. Without the sparsity constraint, the major step of steepest descent method is

$$x^{new} = x + \lambda A^T(y - Ax), \tag{2.9}$$

where $\lambda > 0$ denotes a stepsize. However, the iterate x^{new} generated by (2.9) is possibly not feasible to the problem (2.8), since most of its entries are usually nonzero. It makes sense to use the hard thresholding technique to get a sparse vector satisfying the constraint of (2.8), leading to

$$x^{new} = \mathcal{H}_k\left(x + \lambda A^T(y - Ax)\right), \tag{2.10}$$

where the operator $\mathcal{H}_k(\cdot)$ retains the largest k magnitudes of a vector and zeroes out others. Formula (2.10) is actually the iterative hard thresholding (IHT) which was initiated by Blumensath and Davies [18] in 2008, and it is also the called gradient descent with sparsification in [64]. The full description of IHT is stated as Algorithm 2.3.

Algorithm 2.3 Iterative Hard Thresholding (IHT)

- Input: measurement matrix $A \in \mathbb{R}^{m\times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:
$$x^{p+1} = \mathcal{H}_k(x^p + \lambda A^T(y - Ax^p)).$$
- Output: The k sparse vector $x^* := x^{p+1}$.

The stepsize λ in IHT can be either a fixed number, e.g., $\lambda \in (0, 1]$, or can be adaptively changed from one iteration to another, such as the one used in traditional conjugate methods, i.e.,

$$\lambda^p = \frac{\|(A_{\Gamma^p})^T(y - Ax^p)\|_2^2}{\|A_{\Gamma^p}A^T(y - Ax^p)\|_2^2},$$

where $\Gamma^p = \mathrm{supp}(x^p)$ and A_{Γ^p} denotes the matrix obtained by retaining the columns indexed by Γ^p and zeroing out others. The IHT using such stepsize is called normalized iterative hard thresholding (NIHT) [19, 108]. Another modification of IHT is the Expectation-

Conditional Maximization Either (ECME) algorithm [99], i.e.,

$$x^{p+1} = \mathcal{H}_k\left(x^p + A^T(AA^T)^{-1}(y - Ax^p)\right).$$

The method above becomes the method in (2.10) when $AA^T = I$.

Foucart [60] has proposed another enhanced version of IHT called hard thresholding pursuit (HTP) by taking an orthogonal projection step into account to further reduce the objective function. The full HTP algorithm is described as Algorithm 2.4.

Algorithm 2.4 Hard Thresholding Pursuit (HTP)

- Input: measurement matrix $A \in \mathbb{R}^{m\times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:

$$\overline{x}^{p+1} = \mathcal{H}_k(x^p + \lambda A^T(y - Ax^p)) \qquad (\mathrm{HTP}_1)$$
$$x^{p+1} = \arg\min_w \{\|y - Az\|_2^2, \mathrm{supp}(z) \subseteq \mathrm{supp}(\overline{x}^{p+1})\}. \qquad (\mathrm{HTP}_2)$$

- Output: The k-sparse vector $x^* := x^{p+1}$.

In Algorithm 2.4, (HTP_2) is called the pursuit step, which minimizes the objective function over a specific support set. In fact, if a support set $\Omega \subseteq [N]$ is given, the original underdetermined linear system $y = Ax$ will be reduced to an overdetermined system $y = [A_\Omega][x_\Omega]$, where $[A_\Omega]$ denotes the submatrix obtained from A by retaining columns indexed by Ω and discarding others, and $[x_\Omega]$ denotes the subvector obtained from x by retaining entries indexed by Ω and deleting others. As a result, the problem in (HTP_2) can be reduced to an unconstrained optimization problem, i.e.,

$$\min_{x\subseteq\Omega} \|y - [A_\Omega]\,[x_\Omega]\|_2^2. \qquad (2.11)$$

When $t := \mathrm{card}(\Omega)$ (the cardinality of Ω) is small enough such that any t columns of A

are linearly independent, the solution to (2.11) can be expressed in an explicit way, i.e.,

$$[x_\Omega] = \left([A_\Omega]^T [A_\Omega]\right)^{-1} [A_\Omega]^T y.$$

Both the NIHT and ECME algorithms can be enhanced by introducing such a pursuit step to further reduce the value of the objective function. More accelerated versions of IHT and HTP can be found in [34] and [16].

Hard thresholding methods have been extended to other scenarios, such as low-rank tensor recovery [65, 100], matrix completion [108], and robust regression [11, 130]. The aforementioned algorithms, including IHT, HTP and NIHT, are based on the steepest descent method. However, other first-order or second-order optimization methods can also be combined with the hard thresholding technique, such as the stochastic gradient method [143], the conjugate gradient method [116, 15, 14], alternating projection methods [78, 79], and Newton's methods [73, 125, 37, 126, 117, 84, 145]. Hard thresholding methods have gained a great deal of attention due to their low computational complexity.

Remark 2.3.3

Theoretical justification of hard-thresholding-type algorithms is usually established in terms of the RIP. Sufficient conditions for the guaranteed performance of the IHT have been analyzed by several researchers, such as $\delta_{3k} < 1/\sqrt{32} \approx 0.1767$ [17], $\delta_{2k} < 1/3 \approx 0.3333$ [64], $\delta_{3k} < 1/\sqrt{3} \approx 0.5774$ [60] and $\delta_{3k} < (\sqrt{5}-1)/2 \approx 0.618$ [142] which is the tightest result to our best knowledge. The HTP by Foucart [60] is shown to be stable and efficient under the condition $\delta_{3k} < 1/\sqrt{3} \approx 0.5774$. More theoretical analysis for hard thresholding-type methods can be found in the literature, e.g., [23], [123], [103] and [124].

2.3.3 Optimal k-thresholding algorithms

As pointed in [134], an inherent drawback of hard thresholding is that, when imposed on an incompressible vector, the value of the objective function might increase compared to the objective value at its unthresholded counterpart. This is because the hard thresholding

is independent of the objective function. To overcome this weakness and improve the performance of (2.10), Zhao [134, 139] has proposed a novel vector sparsification approach called optimal k-thresholding, which simultaneously performs thresholding and objective minimization.

The key iterative scheme of optimal k-thresholding algorithm for solving (2.8) can be simply stated as

$$x^{p+1} = \mathcal{Z}_k^{\#}\left(x^p + \lambda A^T(y - Ax^p)\right),$$

where $\mathcal{Z}_k^{\#}(\cdot)$ is the so-called optimal k-thresholding operator and $\lambda > 0$ is the stepsize. Given a vector u, the optimal k-thresholding vector of u is $\mathcal{Z}_k^{\#}(u) = u \otimes w^*$ (the Hadamard product of two vectors), where the vector w^* is the optimal solution to the quadratic 0-1 optimization problem

$$w^* := \arg\min_w \left\{ \|y - A(w \otimes u)\|_2^2 : \ \mathbf{e}^T w = k, w \in \{0,1\}^n \right\}. \tag{2.12}$$

In the formula above, $\mathbf{e} = (1, \ldots, 1)^T \in \mathbb{R}^n$ is the vector of ones, and $\{0,1\}^n$ denotes the set of n-dimensional 0-1 vectors. The full optimal k-thresholding algorithm is given as Algorithm 2.5.

Algorithm 2.5 Optimal k-thresholding (OT)

- Input: measurement matrix $A \in \mathbb{R}^{m \times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:

$$\begin{aligned} u^p &= x^p + \lambda A^T(y - Ax^p), \\ w^p &= \arg\min_z \left\{ \|y - A(w \otimes u^p)\|_2 : \ \mathbf{e}^T w = k, w \in \{0,1\}^n \right\}, \\ x^{p+1} &= w^p \otimes u^p. \end{aligned}$$

- Output: The k-sparse vector $x^* := x^{p+1}$.

The problem (2.12) is a binary optimization problem which usually needs substantial computational cost. To avoid solving it, Zhao [134, 139] suggests to solve the convex

relaxation problem

$$\overline{w}^* := \arg\min_{w} \left\{ \|y - A(w \otimes u)\|_2^2 : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e} \right\},$$

where $0 \leq w \leq \mathbf{e}$ is the convex hull of $w \in \{0,1\}^n$. The relaxed optimal k-thresholding (ROT) algorithm [134] is given as Algorithm 2.6.

Algorithm 2.6 Relaxed Optimal k-thresholding (ROT)

- Input: measurement matrix $A \in \mathbb{R}^{m \times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:

$$\begin{aligned} u^p &= x^p + \lambda A^T(y - Ax^p), \\ \overline{w}^p &= \arg\min_{z} \left\{ \|y - A(w \otimes u^p)\|_2^2 : \ \mathbf{e}^T w = k, 0 \leq w \leq \mathbf{e} \right\}, \\ x^{p+1} &= \mathcal{H}_k(\overline{w}^p \otimes u^p). \end{aligned}$$

- Output: The k-sparse vector $x^* := x^{p+1}$.

In Algorithm 2.6, the vector $\overline{w}^p$ may not necessarily be k-sparse, thus the hard thresholding $\mathcal{H}_k(\cdot)$ is used to truncate the iterate $\overline{w}^p \otimes u^p$ to satisfy the sparsity constraint. We should emphasize that the use of $\mathcal{H}_k(\cdot)$ here is different from the one used in hard thresholding algorithms, because the vector $\overline{w}^p \otimes u^p$ is either k-sparse or admits a compressible feature whereas the vector u^p may not be sparse or compressible at all.

Both OT and ROT algorithms can be accelerated by taking a pursuit step into account after solving the optimization problem, leading to the so-called optimal k-thresholding pursuit (OTP) and relaxed optimal k-thresholding pursuit (ROTP) (see Algorithm 2.7), respectively. Moreover, if the vector $\overline{w}^p \otimes u^p$ in Algorithms 2.6 and 2.7 is not compressible enough, multiple relaxed optimal thresholding can be performed to generate a more compressible vector leading to the ROTP$_\omega$ algorithm, where ω is the number of times of compression performed in one iteration. We omit them here to focus on the fancy framework of the optimal k-thresholding technique. The full description of algorithms,

Algorithm 2.7 Relaxed optimal k-thresholding pursuit (ROTP)

- Input: measurement matrix $A \in \mathbb{R}^{m\times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:

$$\begin{aligned} u^p &= x^p + \lambda A^T(y - Ax^p), \\ \overline{w}^p &= \arg\min_w \left\{ \|y - A(w \otimes u^p)\|_2 : \ \mathbf{e}^T w = k, 0 \le w \le \mathbf{e} \right\}, \\ S^{p+1} &= \mathcal{L}_k(\overline{w}^p \otimes u^p) \\ x^{p+1} &= \arg\min_w \{\|y - Az\|_2^2 : \mathrm{supp}(z) \subseteq S^{p+1}\}. \end{aligned}$$

- Output: The k-sparse vector $x^* := x^{p+1}$.

theoretical analysis and empirical results are given in [134] and [139].

Remark 2.3.4

In addition to the hard and optimal thresholding strategies, there exists another one called soft thresholding, which generates a k-compressible vector by utilizing a piecewise function, i.e.,

$$S_\tau(x) = \begin{cases} x - \tau, & x > \tau \\ 0, & |x| \le \tau \\ x + \tau, & x < -\tau \end{cases}$$

where τ is a threshold value. The soft thresholding (also called shrinkage-thresholding) is also discussed by several authors including Donoho [50], Khare and Tiwary [74], Biwas and Om [12], and Liu et al. [80], Daubechies, Defrise and Mol [42], Fawzi, Davies and Frossard [57], Liu et al. [82], and Liu and Barber [81].

2.3.4 Heuristic algorithms

Three well-known algorithms are orthogonal matching pursuit (OMP) [111], compressive sampling matching pursuit (CoSaMP) [89] and subspace pursuit (SP) [41]. The first application of OMP to signal processing can be dated back to the work [96] and [47]. The

Algorithm 2.8 Orthogonal Matching Pursuit (OMP)

- Input: measurement matrix $A \in \mathbb{R}^{m\times n}$, measurement $y \in \mathbb{R}^m$, and sparsity level k.
- Initialization: $S^0 = \emptyset$, $x^0 = 0$.
- Iteration:

$$S^{p+1} = S^p \cup \arg\max_{j\in[N]} \left\{ \left| \left(A^T \left(y - Ax^p \right) \right)_j \right| \right\}$$
$$x^{p+1} = \arg\min_z \left\{ \|y - Az\|_2^2 : \mathrm{supp}(z) \subseteq S^{p+1} \right\}.$$

- Output: k-sparse vector $x^* := x^{p+1}$.

OMP is a greedy algorithm that iteratively adds one index with maximum correlation to the residual into the identified support and then performs the orthogonal projection onto the support to update the iterate. For small-sized problems, OMP is fast at finding a sparse solution since it only needs k iterations [88]. However, when dealing with large-scale problems, the number of iterations required by OMP dramatically increases. More detailed discussions about the OMP can be found in [112, 45, 26, 129, 88, 36, 137, 40, 77].

Some variants of Algorithm 2.8 have been proposed, such as regularized orthogonal matching pursuit (ROMP) [91, 90], generalized orthogonal matching pursuit (GOMP) [114, 75, 115], stagewise OMP (StOMP) [53], weak OMP (WOMP) [111] and the latest dynamic OMP (DOMP) algorithms [141]. Empirical results in [141] show that the DOMP-type methods are efficient for sparse signal recovery, and the number of iterations they require is much less than OMP.

Two matching-pursuit-type algorithms which can be also viewed as some modifications of OMP are the CoSaMP (see Algorithm 2.9) and SP (see Algorithm 2.10). These two algorithms share a similar framework: (i) support identification, (ii) orthogonal projection, (iii) k-sparse vector generation. The first sufficient condition $\delta_{4k} < 0.17157$ for the efficiency of CoSaMP in terms of the RIP for sparse vector recovery was established by Needell and Tropp [89]. After that, it has been improved to $\delta_{4k} < 0.38427$ in [61] and $\delta_{4k} < 0.5102$ in [142]. For SP, the RIP conditions for its efficiency include $\delta_{3k} < 0.205$ in [41] and

Algorithm 2.9 Compressive Sampling Matching Pursuit (CoSaMP)

- Input: measurement matrix $A \in \mathbb{R}^{m\times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:

$$\begin{aligned} U^{p+1} &= \mathrm{supp}(x^p) \cup \mathcal{L}_{2k}(A^T(y - Ax^p)), \\ u^{p+1} &= \arg\min_z \left\{ \|y - Az\|_2^2, \mathrm{supp}(z) \subseteq U^{n+1} \right\}, \\ x^{p+1} &= \mathcal{H}_k(u^{p+1}) \end{aligned}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Algorithm 2.10 Subspace Pursuit (SP)

- Input: measurement matrix $A \in \mathbb{R}^{m\times n}$, measurement $y \in \mathbb{R}^m$, sparsity level k, and an initial point $x^0 = 0$.
- Iteration:

$$\begin{aligned} U^{p+1} &= \mathrm{supp}(x^p) \cup \mathcal{L}_k(A^T(y - Ax^p)), \\ u^{p+1} &= \arg\min_z \left\{ \|y - Az\|_2^2, \mathrm{supp}(z) \subseteq U^{n+1} \right\}, \\ S^{p+1} &= \mathcal{L}_k(u^{p+1}), \\ x^{p+1} &= \arg\min_z \left\{ \|y - Az\|_2^2, \mathrm{supp}(z) \subseteq S^{p+1} \right\}. \end{aligned}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

$\delta_{3k} < 0.4859$ in [105].

Remark 2.3.5

The complexity is an important feature of an algorithm concerning their efficiency and computational cost. In Table 2.1, we summarize the complexity of some classical CS algorithms per iteration in terms of the number of measurements m, the vector length n, the sparsity level k and the size of the problem data encoding in binary L. BP is a typical convex method, the next four algorithms are generally used heuristic methods, and the last three algorithms belong to thresholding methods.

Table 2.1: Complexity of some CS algorithms

Algorithm	Computational Complexity	References
BP	$O(n^3)$	[38], [41], [98]
OMP	$O(kmn)$	[112], [91], [41]
StOMP	$O(n \log n)$	[53], [98]
CoSaMP	$O(mn)$	[89], [98]
SP	$O(kmn)$	[41], [98]
IHT	$O(mn)$	[17],[139]
HTP	$O(m^3 + mn)$	[60], [139]
ROTP$_w$	$O(m^3 + mn + n^{3.5}L)$	[139]

However, the complexity does not give a full picture of an algorithm. In practical applications, one also needs to choose among algorithms by comparing their numerical performances in terms of objective reduction, the ability of sparse signal recovery, the average number of iterations required, the average runtime, phase transition behavior, capability for real image restoration and others. In the next chapters, we formally describe the Newton-step-based hard thresholding algorithms.

CHAPTER 3

NEWTON-STEP-BASED HARD THRESHOLDING ALGORITHMS

3.1 Introduction

The hard thresholding technique incorporated with the steepest descent method shows a good theoretical performance in sparse signal recovery. The classic optimization theory indicates that Newton-like methods often have a numerical advantage over the traditional gradient methods for solving nonlinear optimization problems from the perspective of local convergence. This motivates us to propose the so-called Newton-step-based hard thresholding algorithms discussed in this chapter[1], which integrate a Newton-like search direction and the hard thresholding technique.

Consider the sparse optimization problem

$$\min_x \left\{ \frac{1}{2}\|y - Ax\|_2^2 : \|x\|_0 \leq k \right\}, \tag{3.1}$$

where the ℓ_0-norm $\|x\|_0$ counts the number of nonzero entries of x, $A \in \mathbb{R}^{m\times n}$ is a measurement matrix with $m \ll n$ and k is a given integer. The problem (3.1) is difficult to solve due to the appearance of the ℓ_0-norm without which the problem can be solved easily by one of the classic optimization methods. The main step of steepest descent method for

[1]The main result in this chapter has been published (jointly with Prof. Yun-Bin Zhao) on IEEE Transactions on Signal Processing, Vol. 68, pp. 6594-6606, 2020 [84]

solving the unconstrained problem $\min_x \frac{1}{2}\|y - Ax\|_2^2$ can be formulated as

$$x^{\text{new}} = x^p + \lambda A^T (y - Ax^p), zhao2021analysis \tag{3.2}$$

where $\lambda > 0$ denotes a stepsize at the current iterate x^p. Clearly, the iterate generated by (3.2) may not be sparse, so it might not be feasible for problem (3.1). The hard thresholding technique, as one of the simplest methods for generating sparse vectors, is introduced to derive a vector satisfying the constraint of (3.1), leading to the iterative hard thresholding (IHT) method [18]:

$$x^{p+1} = \mathcal{H}_k \left(x^p + \lambda A^T (y - Ax^p) \right),$$

where the hard thresholding operator $\mathcal{H}_k(\cdot)$ retains the largest k magnitudes of a vector and zeroes out the other entries. The earlier works, e.g., [18] and [20], provide solid theoretical guarantees under the RIP condition for the IHT, yet its practical performance is not as good as expected. For instance, empirical results indicate that, when the stepsize λ is inappropriate, IHT may terminate at a local minima, or the objective function sometimes increases sharply during the course of iterations, both may cause reconstruction failure. The classic optimization theory [24, 94] has shown that the Newton-like methods usually have a faster local convergence than the gradient method. Thus it is generally more efficient than the steepest descent method for solving nonlinear optimization problems, which motivates one to develop a Newton-like hard thresholding method.

The main contribution of this chapter can be summarized as follows.

i. We introduce a class of Newton-step-based hard thresholding methods for solving the problem (3.1), which are described as Algorithms 3.1 and 3.2, respectively. Extensions of the algorithms to a generic sparse optimization problem are also provided, as stated in Algorithms 3.3 and 3.4.

ii. We establish some sufficient conditions in terms of restricted isometry property (RIP)

for the guaranteed success of the proposed algorithms for sparse signal recovery. The main results are summarized in Theorems 3.3.5 and 3.3.9.

iii. Experiments have been conducted to investigate the numerical behaviors of the proposed algorithms in residual reduction and sparse signal reconstruction. Empirical results through a large number of random sparse optimization problems indicate that the proposed algorithms are generally robust and efficient for signal recovery, and they are also comparable to several mainstream CS algorithms when the algorithmic parameters are suitably chosen.

3.2 Algorithms

Given a twice continuously differentiable function f with gradient $\nabla f(x)$ and non-singular Hessian matrix $\nabla^2 f(x)$, the traditional Newton-type methods use the following iterative scheme:

$$x^{p+1} = x^p - \lambda \left(\nabla^2 f\left(x^p\right)\right)^{-1} \nabla f\left(x^p\right),$$

where $\lambda > 0$ is a stepsize. The gradient and Hessian of $f(x) = \dfrac{1}{2}\|y - Ax\|_2^2$ are given, respectively, by

$$\nabla f(x) = -A^T(y - Ax), \quad \left(\nabla^2 f(x)\right)^{-1} = A^T A.$$

However, the Hessian is a singular matrix in compressed sensing scenarios since $m \ll n$. In this case, a direct use of the pure form of Newton's method is difficult, but one can develop a Newton-like method by introducing a suitable modification of the Hessian. A simple idea is to modify the Hessian with a parameter $\epsilon > 0$ so that $A^T A + \epsilon I$ is positive definite, where $I \in \mathbb{R}^{n\times n}$ is the identity matrix. Such a modification of the Hessian leads to the following Newton-like iterative method:

$$x^{new} = x + \lambda(A^T A + \epsilon I)^{-1} A^T (y - Ax^p)), \tag{3.3}$$

where $\lambda > 0$ denotes the stepsize, and $\epsilon > 0$ is a parameter to modify the Hessian. Similar to (3.2), the iterate generated by (3.3) may not be sparse so that the constraint of (3.1) cannot be satisfied. To develop an iterative method that generates sparse iterate, it makes sense to consider the following iterative scheme

$$x^{p+1} = \mathcal{H}_k\left(x^p + \lambda\left(A^TA + \epsilon I\right)^{-1} A^T\left(y - Ax^p\right)\right),$$

which only retains the largest k components generated by (3.3). It is actually the key step of the so-called Newton-step-based hard thresholding algorithm, and a generic description can be found in Algorithm 3.1.

Algorithm 3.1 Newton-Step-Based Iterative Hard Thresholding (NSIHT)

- Input: measurement matrix A, measurement vector y, sparsity level k, parameter $\epsilon > 0$, and stepsize λ.
- Iteration:
$$x^{p+1} = \mathcal{H}_k\left(x^p + \lambda(A^TA + \epsilon I)^{-1}A^T(y - Ax^p)\right).$$
- Output: The k-sparse vector $x^* := x^{p+1}$.

The objective function of problem (3.1) can be further reduced by performing a pursuit step over the support set generated by NSIHT, leading to the so-called Newton-step-based hard thresholding pursuit (NSHTP) described in Algorithm 3.2.

Algorithm 3.2 Newton-Step-Based Hard Thresholding Pursuit (NSHTP)

- Input: measurement matrix A, measurement vector y, sparsity level k, parameter $\epsilon > 0$, and stepsize λ.
- Iteration:
$$\begin{aligned} \overline{x}^p &= \mathcal{H}_k\left(x^p + \lambda(A^TA + \epsilon I)^{-1}A^T(y - Ax^p)\right), \\ x^{p+1} &= \arg\min_z\{\|y - Az\|_2^2 : \ \mathrm{supp}(z) \subseteq \mathrm{supp}(\overline{x}^p)\}. \end{aligned}$$
- Output: The k-sparse vector $x^* := x^{p+1}$.

3.2.1 Extension to general problem

Problem (3.1) is one of the special cases of the following general sparse optimization problem:

$$\min_x \{f(x) : \|x\|_0 \leq k\}, \tag{3.4}$$

where $f(x)$ is a twice differentiable function. The NSIHT and NSHTP can also be generalized to an ordinary form for solving the problem above with slight differences. If the Hessian of $f(x)$ is positive definite, the modification of Hessian is superfluous, thus the damped Newton's method (see Section 2.2.2) can be directly used. However, if the Hessian of $f(x)$ is not positive definite, we can introduce a parameter $\epsilon > 0$ to slightly change $\nabla^2 f(x)$ so that $\nabla^2 f(x) + \epsilon I \succ 0$ (positive definite). The generalized versions of NSIHT and NSHTP for solving (3.4), based on such Hessian modifications, are described in Algorithms 3.3 and 3.4, respectively.

Algorithm 3.3 Generalized Newton-Step-Based Iterative Hard Thresholding (GNSIHT)

- Input: objective function $f(x)$, sparsity level k, parameter $\epsilon \geq 0$, and stepsize λ.
- Iteration: If $\nabla^2 f(x^p) \succ 0$, set $\epsilon = 0$; Otherwise, choose a suitable $\epsilon > 0$ such that $\nabla^2 f(x^p) + \epsilon I \succ 0$. Then set

$$x^{p+1} = \mathcal{H}_k \left(x^p - \lambda (\nabla^2 f(x^p) + \epsilon I)^{-1} \nabla f(x^p) \right).$$

- Output: The k-sparse vector $x^* := x^{p+1}$.

Remark 3.2.1

Some progress of Newton-type algorithms have recently been achieved for solving the sparse optimization problem (3.1) and (3.4), e.g., Jing, Zhou and Qi [73], Yuan and Liu [125], Yuan [126], Wang and Qu [117], Zhou, Xiu and Qi [145] and the references therein. However, our method is not a special case of the existing sparsity Newton-type algorithms. A significant difference between our method and the Newton hard thresholding pursuit method in [145] is that our method adopts an invariant Hessian, rather than iteratively extracting submatrices of $A^T A$. The Newton greedy pursuit algorithm in [126] solves a

Algorithm 3.4 Generalized Newton-Step-Based Hard Thresholding Pursuit (GNSHTP)

- Input: objective function $f(x)$, sparsity level k, parameter $\epsilon \geq 0$, and stepsize λ.
- Iteration: If $\nabla^2 f(x^p) \succ 0$, set $\epsilon = 0$; Otherwise, choose a suitable $\epsilon > 0$ such that $\nabla^2 f(x^p) + \epsilon I \succ 0$. Then set

$$\begin{aligned} \overline{x}^{p+1} &= \mathcal{H}_k \left(x^p - \lambda (\nabla^2 f(x^p) + \epsilon I)^{-1} \nabla f(x^p) \right), \\ x^{p+1} &= \arg\min_z \{ f(z) : \ \mathrm{supp}(z) \subseteq \mathrm{supp}(\overline{x}^{p+1}) \}. \end{aligned}$$

- Output: The k-sparse vector $x^* := x^{p+1}$.

local quadratic problem constructed by the second-order Taylor expansion around the current iterate, while we provide an explicit iterative scheme which is much simpler than it. The method in [73] called quasi-Newton iterative projection (QNIP) also illustrates a clear expression, i.e.,

$$x^{p+1} = \mathcal{H}_k \left(x^p + \lambda \left(2I - A^T A \right) A^T \left(y - A x^p \right) \right). \tag{3.5}$$

However, it differs from ours in the construction of Newton-type search direction, where (3.5) approximates the Hessian inverse directly, whereas our approach modifies the singular Hessian.

3.3 Theoretical Analysis

In this section, we establish the standard sufficient conditions including convergence in terms of the restricted isometry property (RIP) for the guaranteed success of the NSIHT and NSHTP algorithms for sparse signal recovery. We first analyze the NSIHT algorithm in noisy scenarios, i.e., $y = Ax + \eta$, where η denotes a noise vector. Without loss of generality, we assume the target vector x is k-compressible, that is x can be well approximated by its largest k magnitudes, i.e., $x \approx x_S$ where $S = \mathcal{L}_k(x)$. But our results can be simplified to the case where x is k-sparse or measurements are exact, i.e., $x = x_S$ and $y = Ax$. We will

show that signal recovery via the proposed algorithms can be guaranteed under a standard RIP assumption and the suitable choice of the algorithmic parameter (ϵ, λ).

3.3.1 Analysis of NSIHT in noisy scenarios

Before showing the main result of this section, we need some useful technical lemmas.

Lemma 3.3.1 (Lemma 6.16 of [63]). Let $A \in \mathbb{R}^{m\times n}$ with $m \ll n$ be a RIP matrix. Given vectors $u, v \in \mathbb{R}^n$ and an index set $\Omega \subset [N]$, one has

$$\begin{aligned} &\left|\langle u, (I - A^T A)v\rangle\right| \leq \delta_t \|u\|_2 \|v\|_2 && \text{if } \operatorname{card}(\operatorname{supp}(u) \cup \operatorname{supp}(v)) \leq t, \\ &\left\|\left(\left(I - A^T A\right) u\right)_\Omega\right\|_2 \leq \delta_t \|u\|_2 && \text{if } \operatorname{card}(\operatorname{supp}(v) \cup \Omega) \leq t. \end{aligned}$$

The following proof can be found in [63, Proof 6.16].

Proof. Denote $\Lambda \coloneqq \operatorname{supp}(u) \cup \operatorname{supp}(v)$, we have

$$\begin{aligned} \left|\langle u, (I - A^T A)v\rangle\right| &= |\langle u, v\rangle - \langle Au, Av\rangle| \\ &= |\langle u_\Lambda, v_\Lambda\rangle - \langle A_\Lambda u_\Lambda, A_\Lambda v_\Lambda\rangle| \\ &= \left|\langle u_\Lambda, (I - (A_\Lambda)^T A_\Lambda)v_\Lambda\rangle\right| \\ &\leq \|u_\Lambda\|_2 \left\|(I - (A_\Lambda)^T A_\Lambda)v_\Lambda\right\|_2 \\ &\leq \left\|(I - (A_\Lambda)^T A_\Lambda)\right\|_2 \|u_\Lambda\|_2 \|v_\Lambda\|_2, \end{aligned}$$

where $A_\Lambda \in \mathbb{R}^{m\times n}$ and denotes the matrix which retains the columns indexed by Λ and zeroes out the others, and $u_\Lambda \in \mathbb{R}^n$ denotes the vector which retains the entries indexed by Λ and zeroes the out others. An equivalent expression of the restricted isometry constant of order t is given as

$$\left\|(I - (A_\Gamma)^T A_\Gamma)\right\|_2 \leq \max_{\Gamma\subset[N], \operatorname{card}(\Gamma)\leq t} \left\|(A_\Gamma)^T A_\Gamma - I\right\|_2 = \delta_t.$$

Therefore, we have

$$\left|\langle u, (I - A^T A)v\rangle\right| \leq \delta_t \|u\|_2 \|v\|_2.$$

The second inequality follows from

$$\left\|((I - A^T A)v)_\Omega\right\|_2^2 = \left\langle ((I - A^T A)v)_\Omega, (I - A^T A)v \right\rangle \leq \delta_t \left\|((I - A^T A)v)_\Omega\right\|_2 \|v\|_2.$$

■

Lemma 3.3.2 (Lemma 6.20 of [63]). Let $A \in \mathbb{R}^{m\times n}$ with $m \ll n$ be a RIP matrix. Given a vector $v \in \mathbb{R}^m$ and a support set $\Omega \subset [N]$, one has

$$\left\|(A^T v)_\Omega\right\|_2 \leq \sqrt{1+\delta_t}\|v\|_2 \qquad \text{if } \operatorname{card}(\Omega) \leq t.$$

Proof. We have

$$\begin{aligned}\left\|(A^T v)_\Omega\right\|_2^2 = \left\langle A^T v, (A^T v)_\Omega \right\rangle = \left\langle v, A^T\left((A^T v)_\Omega\right)\right\rangle &\leq \|v\|_2 \left\|A\left((A^T v)_\Omega\right)\right\|_2 \\ &\leq \|v\|_2 \sqrt{1+\delta_t}\left\|(A^T v)_\Omega\right\|_2.\end{aligned}$$

The desired result follows. ■

The next lemma is key to our theoretical analysis concerning the property of Newton-like search direction.

Lemma 3.3.3. *Let $A \in \mathbb{R}^{m\times n}$ with $m \ll n$ be a measurement matrix satisfying the RIP. Given vectors $u, v \in \mathbb{R}^n$ and an index set $\Omega \subset [N]$, if $\sigma_1^2 < \epsilon$ and $\lambda \leq \epsilon + \sigma_m^2$, where σ_1, σ_m are the largest and smallest singular values of the matrix A, respectively, then one has*

(i) if $\operatorname{card}(\operatorname{supp}(u) \cup \operatorname{supp}(v)) \leq t$

$$\left|\left\langle u, (I - \lambda(A^T A + \epsilon I)^{-1} A^T A)v\right\rangle\right| \leq \left(\delta_t + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|u\|_2\|v\|_2,$$

(ii) if $\mathrm{card}(\Omega \cup \mathrm{supp}(v)) \leq t$

$$\left\| \left[(I - \lambda(A^T A + \epsilon I)^{-1} A^T A) v \right]_\Omega \right\|_2 \leq \left(\delta_t + \sigma_1^2 - \frac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2} \right) \|v\|_2.$$

Proof. The matrix $(\epsilon I + A^T A)^{-1}$ can be written as

$$(\epsilon I + A^T A)^{-1} = \frac{1}{\epsilon}(I + \frac{1}{\epsilon} A^T A)^{-1}. \tag{3.6}$$

It is well known that [71] if a square matrix M satisfies $\|M\|_2 < 1$, then

$$(I + M)^{-1} = I - M + M^2 - M^3 + \cdots = I + \sum_{i=1}^{\infty} (-1)^i M^i. \tag{3.7}$$

In order to use the matrix inverse expansion formula above, we choose the parameter ϵ such that

$$\left\| \frac{1}{\epsilon} A^T A \right\|_2 = \frac{1}{\epsilon} \lambda_{\max}(A^T A) < 1,$$

where $\lambda_{\max}(A^T A) = \sigma_1^2$ is the largest eigenvalue of matrix $A^T A \in \mathbb{R}^{n \times n}$. It implies $\epsilon > \sigma_1^2$. By formula (3.7), the matrix $(I + \frac{1}{\epsilon} A^T A)^{-1}$ therefore can be expanded as

$$\left(I + \frac{1}{\epsilon} A^T A \right)^{-1} = I + \sum_{i=1}^{\infty} (-1)^i \left(\frac{1}{\epsilon} A^T A \right)^i = I + \sum_{i=1}^{\infty} \left(-\frac{1}{\epsilon} \right)^i (A^T A)^i. \tag{3.8}$$

By using (3.6) and (3.8), we have

$$\begin{aligned}
I - \lambda(A^T A + \epsilon I)^{-1} A^T A &= I - \frac{\lambda}{\epsilon}(I + \frac{1}{\epsilon} A^T A)^{-1} A^T A \\
&= I - \frac{\lambda}{\epsilon} \left(I + \sum_{i=1}^{\infty} \left(-\frac{1}{\epsilon} \right)^i (A^T A)^i \right) A^T A \\
&= I + \lambda \left(-\frac{1}{\epsilon} A^T A + \sum_{i=1}^{\infty} \left(-\frac{1}{\epsilon} \right)^{i+1} \left(A^T A \right)^{i+1} \right) \\
&= I + \lambda \sum_{j=1}^{\infty} \left(-\frac{1}{\epsilon} \right)^j (A^T A)^j \\
&= (I - A^T A) + B,
\end{aligned}$$

where matrix $B \in \mathbb{R}^{n\times n}$ is given as

$$B = A^TA + \lambda \sum_{j=1}^{\infty} \left(-\frac{1}{\epsilon}\right)^j (A^TA)^j.$$

Using the triangle inequality, we can write

$$\begin{aligned}
\left|\langle u, (I - \lambda(A^TA + \epsilon I)^{-1}A^TA)v\rangle\right| &= \left|\langle u, (I - A^TA + B)v\rangle\right| \\
&= \left|\left\langle u, (I - A^TA)v\right\rangle + \langle u, Bv\rangle\right| \\
&\leq \left|\left\langle u, (I - A^TA)v\right\rangle\right| + \left|\langle u, Bv\rangle\right|. \qquad (3.9)
\end{aligned}$$

By Lemma 3.3.1, the first term of (3.9) on the right-hand side can be bounded as

$$\left|\left\langle u, (I - A^TA)v\right\rangle\right| \leq \delta_t \|u\|_2 \|v\|_2, \qquad (3.10)$$

where $t = \text{card}(\text{supp}(u) \cup \text{supp}(v))$. We now bound the term $|\langle u, Bv\rangle|$. Note that

$$|\langle u, Bv\rangle| \leq \|B\|_2 \|u\|_2 \|v\|_2, \qquad (3.11)$$

To bound $\|B\|_2$, we utilize the singular value decomposition of matrix $A \in \mathbb{R}^{m\times n}$, i.e., $A = U\Sigma V$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are two orthogonal matrices, and $\Sigma \in \mathbb{R}^{m\times n}$ is the singular value matrix (e.g., see (1.12)). By using this decomposition, the matrix B can be denoted as

$$B = VCV^T,$$

where

$$C = \text{diag}\left(\sigma_1^2 + \lambda \sum_{i=1}^{\infty} \left(-\frac{\sigma_1^2}{\epsilon}\right)^i,\ \ldots,\ \sigma_m^2 + \lambda \sum_{i=1}^{\infty} \left(-\frac{\sigma_m^2}{\epsilon}\right)^i,\ 0,\ \ldots,\ 0\right).$$

Under the choice $\epsilon > \sigma_1^2$, we can calculate the geometric sequence, i.e.,

$$\sum_{i=1}^{\infty} \left(-\frac{\sigma_i^2}{\epsilon}\right)^i = -\frac{\sigma_i^2}{\epsilon + \sigma_i^2}.$$

Therefore, the eigenvalues of matrix B are given as

$$\begin{aligned} \mathrm{eig}(B)_i &= \sigma_i^2 - \frac{\lambda \sigma_i^2}{\epsilon + \sigma_i^2}, \quad i = 1, \ldots, m, \\ \mathrm{eig}(B)_i &= 0, \qquad\qquad\qquad i = m+1, \ldots, n. \end{aligned}$$

Under the condition $\lambda \leq \epsilon + \sigma_1^2$, the ℓ_2-norm of matrix B is given as

$$\|B\|_2 = \sigma_1^2 - \frac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2}. \tag{3.12}$$

Substituting (3.12) into (3.11) leads to

$$|\langle u, Bv \rangle| \leq \left(\sigma_1^2 - \frac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2}\right) \|u\|_2 \|v\|_2. \tag{3.13}$$

Combining (3.9), (3.10) and (3.13), we have

$$\left|\left\langle u, (I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right\rangle\right| \leq \left(\delta_t + \sigma_1^2 - \frac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2}\right) \|u\|_2 \|v\|_2,$$

which is item (i) in the lemma. Note that the coefficient $\delta_t + \sigma_1^2 - \dfrac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2} > 0$ is guaranteed by $\lambda \leq \epsilon + \sigma_m^2$. Item (ii) in the lemma follows immediately from item (i). In fact, for any set $\Omega \subset [N]$, we see that

$$\begin{aligned} &\left\| \left((I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right)_\Omega \right\|_2^2 \\ &= \left\langle \left((I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right)_\Omega, \left[(I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right]_\Omega \right\rangle \\ &= \left\langle \left((I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right)_\Omega, (I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right\rangle \\ &\leq \left(\delta_t + \sigma_1^2 - \frac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2}\right) \left\| \left((I - \lambda (A^T A + \epsilon I)^{-1} A^T A) v \right)_\Omega \right\|_2 \|v\|_2. \end{aligned} \tag{3.14}$$

where the final inequality follows from the first inequality in the lemma that we just shown above by setting $u = (I - \lambda(A^TA + \epsilon I)^{-1}A^TA)v]_\Omega$. Desired result follows by dividing through (3.14) by $\left\|\left((I - \lambda(A^TA + \epsilon I)^{-1}A^TA)v\right)_\Omega\right\|_2$. ■

The next lemma concerns the property of the hard thresholding operator $\mathcal{H}_k$. This was zshown by Foucart in [63, Chapter 6]. We include the proof here for completeness.

Lemma 3.3.4 [63]. For any vector $z \in \mathbb{R}^n$ and any k-sparse vector $x \in \mathbb{R}^n$, one has

$$\|\mathcal{H}_k(z) - x\|_2 \leq \sqrt{3}\|(z - x)_{S\cup S^*}\|_2,$$

where $S = \text{supp}(x)$ and $S^* = \text{supp}(\mathcal{H}_k(z))$.

Proof. (see proof 6.18 in [63]). Let S and S^* be defined as above. Since S^* is the support for the largest k magnitudes of z, we have $\|z_S\|_2^2 \leq \|z_{S^*}\|_2^2$. Eliminating those elements in $S \cap S^*$ leads to

$$\|z_{S\setminus S^*}\|_2 \leq \|z_{S^*\setminus S}\|_2 = \|(z - x)_{S^*\setminus S}\|_2,$$

where the equality follows from the fact $x_{S^*\setminus S} = 0$. Since $[\mathcal{H}_k(z)]_{S\setminus S^*} = 0$, we also have

$$\|z_{S\setminus S^*}\|_2 = \|(x - \mathcal{H}_k(z) - x + z)_{S\setminus S^*}\|_2 \geq \|(x - \mathcal{H}_k(z))_{S\setminus S^*}\|_2 - \|(z - x)_{S\setminus S^*}\|_2.$$

Combining the above two relations and rearranging terms, we get

$$\|(x - \mathcal{H}_k(z))_{S\setminus S^*}\|_2 \leq \|(z - x)_{S\setminus S^*}\|_2 + \|(z - x)_{S^*\setminus S}\|_2 \leq \sqrt{2}\|(x - z)_{S\cup S^*}\|_2.$$

Therefore, we have

$$\begin{aligned}
\|x - \mathcal{H}_k(z)\|_2^2 &= \|(x - \mathcal{H}_k(z))_{S^*}\|_2^2 + \|(x - \mathcal{H}_k(z))_{\overline{S^*}}\|_2^2 \\
&= \|(x - z)_{S^*}\|_2^2 + \left\|(x - \mathcal{H}_k(z))_{S\setminus S^*}\right\|_2^2 \\
&\leq \|(x - z)_{S^*}\|_2^2 + 2\,\|(x - z)_{S\cup S^*}\|_2^2 \\
&\leq 3\,\|(x - z)_{S\cup S^*}\|_2^2,
\end{aligned}$$

as desired. ■

The next theorem is the main result for NSIHT for sparse signal recovery.

Theorem 3.3.5. *Let the restricted isometry constant δ_{3k} of the measurement matrix $A \in \mathbb{R}^{m\times n}$ satisfy that*

$$\delta_{3k} < \frac{1}{\sqrt{3}}.$$

Let $y := Ax + \eta$ be the measurements of the signal $x \in \mathbb{R}^n$. Let ϵ be a given parameter satisfying

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{\sqrt{3}} - \delta_{3k}} - 1\right)\sigma_1^2\right\}. \tag{3.15}$$

If the stepsize λ in NSIHT is chosen such that

$$\epsilon + \sigma_1^2 - \left(\frac{1}{\sqrt{3}} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} \;<\; \lambda \;\leq\; \epsilon + \sigma_m^2, \tag{3.16}$$

then the sequence $\{x^n\}$ generated by the NSIHT algorithm satisfies

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho\left\|x^p - x_S\right\|_2 + \tau\left\|\eta + Ax_{\overline{S}}\right\|_2,$$

where

$$\rho = \sqrt{3}\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right) < 1 \quad \textit{and} \quad \tau = \frac{\sqrt{3}\lambda\sigma_1}{\epsilon + \sigma_1^2}.$$

Note that $S := \mathcal{L}_k(x)$ is the support set of the largest k magnitudes of x.

The above theorem claims that when the measurements of the signal are slightly inaccurate and the signal can be sparsely approximated, i.e., $x \approx x_S$, the recovery of the major information of the signal can be achieved by the algorithm NSIHT. We now provide the proof of Theorem 3.3.5.

Proof. Let $\overline{S}$ denote the complementary set of S, where $S = \mathcal{L}_k(x)$. The measurements

can be represented as $y = Ax + \eta = Ax_S + \eta'$, where $\eta' = Ax_{\overline{S}} + \eta$. We define

$$\begin{aligned} u^p &\coloneqq x^p + \lambda(A^TA + \epsilon I)^{-1}A^T(y - Ax^p) \\ &= x^p + \lambda(A^TA + \epsilon I)^{-1}A^TA(x_S - x^p) + \lambda(A^TA + \epsilon I)^{-1}A^T\eta'. \end{aligned} \tag{3.17}$$

Denote $S^{p+1} = \operatorname{supp}(x^{p+1})$. We immediately have the following relation by Lemma 3.3.4:

$$\left\|x^{p+1} - x_S\right\|_2 = \|\mathcal{H}_k(u^p) - x_S\|_2 \leq \sqrt{3}\|(u^p - x_S)_{S^{p+1}\cup S}\|_2. \tag{3.18}$$

By (3.17), we can write

$$u^p - x_S = \left(I - \lambda\left(A^TA + \epsilon I\right)^{-1}A^TA\right)(x^p - x_S) + \lambda\left(A^TA + \epsilon I\right)^{-1}A^T\eta'.$$

Thus by triangle inequality, one has

$$\begin{aligned} \|(u^p - x_S)_{S^{p+1}\cup S}\|_2 \leq &\left\|\left(\left(I - \lambda(A^TA + \epsilon I)^{-1}A^TA\right)(x^p - x_S)\right)_{S^{p+1}\cup S}\right\|_2 \\ &+ \left\|\left(\lambda\left(A^TA + \epsilon I\right)^{-1}A^T\eta'\right)_{S^{p+1}\cup S}\right\|_2. \end{aligned}$$

Since $\operatorname{card}(S^{p+1}) \leq k$, $\operatorname{card}(S) \leq k$ and $\operatorname{card}(\operatorname{supp}(x^p)) \leq k$, we see that

$$\operatorname{card}(S^{p+1} \cup S \cup \operatorname{supp}(x^p - x_S)) \leq 3k.$$

Under the condition $\epsilon > \sigma_1^2$ and $\lambda \leq \epsilon + \sigma_m^2$, applying Lemma 3.3.3 leads to

$$\left\|\left(\left(I - \lambda\left(A^TA + \epsilon I\right)^{-1}A^TA\right)(x^p - x_S)\right)_{S^{p+1}\cup S}\right\|_2 \leq \left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x^p - x_S\|_2. \tag{3.19}$$

We now make an observation of matrix $(A^TA + \epsilon I)^{-1}A^T$. By singular value decomposition

of $A = U\Sigma V^T$ as aforementioned, we immediately have that

$$(A^TA + \epsilon I)^{-1}A^T = V^T \begin{bmatrix} \frac{\sigma_1}{\epsilon + \sigma_1^2} & & & \\ & \ddots & & 0_{m\times(n-m)} \\ & & \frac{\sigma_m}{\epsilon + \sigma_m^2} & \end{bmatrix} U^T.$$

Using $\sigma_1^2 < \epsilon$, it is easy to verify

$$\frac{\sigma_1}{\epsilon + \sigma_1^2} \geq \frac{\sigma_2}{\epsilon + \sigma_2^2} \geq \cdots \geq \frac{\sigma_m}{\epsilon + \sigma_m^2},$$

hence we have

$$\left\| \left(A^TA + \epsilon I\right)^{-1} A^T \right\|_2 = \frac{\sigma_1}{\epsilon + \sigma_1^2}.$$

We then can write

$$\left\| \left(A^TA + \epsilon I\right)^{-1} A^T\eta' \right\|_2 \leq \left\| \left(A^TA + \epsilon I\right)^{-1} A^T \right\|_2 \|\eta'\|_2 \leq \frac{\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2. \tag{3.20}$$

Combining (3.18)-(3.20), we obtain

$$\left\|x^{p+1} - x_S\right\|_2 \leq \sqrt{3}\left[\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x^p - x_S\|_2 + \frac{\lambda\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2\right] \tag{3.21}$$

where the coefficient satisfies

$$\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2} > 0$$

which is guaranteed under the condition $\lambda \leq \epsilon + \sigma_m^2$. To ensure the convergence of NSIHT, we require

$$\rho := \sqrt{3}\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right) < 1, \tag{3.22}$$

which is guaranteed if the stepsize λ is taken such that

$$\lambda > \epsilon + \sigma_1^2 - \left(\frac{1}{\sqrt{3}} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2}.$$

So, together with $\lambda \leq \epsilon + \sigma_m^2$, the range of λ is determined as

$$\epsilon + \sigma_1^2 - \left(\frac{1}{\sqrt{3}} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \lambda \leq \epsilon + \sigma_m^2.$$

To ensure the existence of such a range, it is sufficient to require that

$$\frac{1}{\sqrt{3}} - \delta_{3k} > 0,$$

and

$$\epsilon + \sigma_m^2 > \epsilon + \sigma_1^2 - \left(\frac{1}{\sqrt{3}} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2}. \tag{3.23}$$

Inequality (3.23) is guaranteed under the following choice of ϵ:

$$\epsilon > \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{\sqrt{3}} - \delta_{3k}} - 1\right)\sigma_1^2.$$

Together with $\epsilon > \sigma_1^2$, the range of ϵ is given as

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{\sqrt{3}} - \delta_{3k}} - 1\right)\sigma_1^2\right\}.$$

The desired results follow. ■

Our result can be easily simplified to the case when the signal is k-sparse and when the measurements are accurate, i.e., $y := Ax$. Under such a circumstance, the relation (3.21) will be reduced to

$$\left\|x^{p+1} - x\right\|_2 \leq \rho \left\|x^p - x\right\|_2 \leq \rho^{p+1} \left\|x^0 - x\right\|_2,$$

where ρ is the same as (3.22). It implies the sequence $\{x^p\}$ generated by NSIHT converges to x as $p \to \infty$. The result is summarized as follows.

Theorem 3.3.6. *Let the restricted isometry constant δ_{3k} of the measurement matrix*

$A \in \mathbb{R}^{m\times n}$ *satisfy*

$$\delta_{3k} < \frac{1}{\sqrt{3}}.$$

Let ϵ *be a given parameter satisfying*

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{\sqrt{3}} - \delta_{3k}} - 1\right)\sigma_1^2\right\}.$$

If the stepsize λ *in NSIHT is chosen such that*

$$\epsilon + \sigma_1^2 - \left(\frac{1}{\sqrt{3}} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \lambda \leq \epsilon + \sigma_m^2,$$

then for any k*-sparse signal* $x \in \mathbb{R}^n$ *with accurate measurements* $y \coloneqq Ax$*, the sequence* $\{x^n\}$ *generated by the NSIHT algorithm converges to the signal* x*.*

For the special choice of stepsize $\lambda = \epsilon$, following the same proof of Theorem 3.3.5, it can be seen that to ensure the convergence of the algorithm, ϵ should be chosen such that the following inequality is guaranteed:

$$\sigma_1^2 - \left(\frac{1}{\sqrt{3}} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < 0,$$

which implies that

$$\epsilon > \left(\frac{\sigma_1^2}{\frac{1}{\sqrt{3}} - \delta_{3k}} - 1\right)\sigma_1^2.$$

Thus, together with the requirement $\epsilon > \sigma_1^2$, we can summarize the result of NSIHT when $\lambda = \epsilon$ as follows.

Corollary 3.3.7. *Let the restricted isometry constant* δ_{3k} *of the matrix* $A \in \mathbb{R}^{m\times n}$ *satisfy*

$$\delta_{3k} < \frac{1}{\sqrt{3}},$$

and let ϵ be a given parameter satisfying

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2}{\frac{1}{\sqrt{3}} - \delta_{3k}} - 1\right)\sigma_1^2\right\},$$

where σ_1 is the largest singular value of matrix A. Then for any k-sparse signal $x \in \mathbb{R}^n$ with accurate measurements $y := Ax$, the sequence $\{x^n\}$ generated by the NSIHT algorithm with $\lambda = \epsilon$ converges to x.

3.3.2 Analysis of NSHTP in noisy scenarios

We now establish the recovery error bound in noisy situations for the NSHTP which can be regarded as a combination of the NSIHT and the pursuit step. To this purpose, we first state a lemma concerning the property of the pursuit step.

Lemma 3.3.8 [23, 134]. Let $y := A\widehat{x} + \nu$ be the noisy measurements of the k-sparse signal $\widehat{x}$, and let $u \in \mathbb{R}^n$ be an arbitrarily given k-sparse vector. Then the solution of the pursuit step

$$z^* = \arg\min_z \{\|y - Az\|_2^2 : \mathrm{supp}(z) \subseteq \mathrm{supp}(u)\}$$

satisfies that

$$\|z^* - \widehat{x}\|_2 \leq \frac{1}{\sqrt{1 - (\delta_{2k})^2}}\|\widehat{x} - u\|_2 + \frac{\sqrt{1+\delta_k}}{1 - \delta_{2k}}\|\nu\|_2.$$

The proof of this lemma can be found in the proof of Theorem 4.8 in [134], yet we include it for the completeness of the thesis.

Proof. Let $\Omega := \mathrm{supp}(u)$ and $\overline{\Omega}$ denote its complementary set. We see that $(z^*)_{\overline{\Omega}} = 0$ and $u_{\overline{\Omega}} = 0$, thus

$$\begin{aligned}\|\widehat{x} - z^*\|_2^2 &= \|(\widehat{x} - z^*)_\Omega\|_2^2 + \|(\widehat{x} - z^*)_{\overline{\Omega}}\|_2^2 \\ &= \|(\widehat{x} - z^*)_\Omega\|_2^2 + \|(\widehat{x} - u)_{\overline{\Omega}}\|_2^2. \end{aligned} \tag{3.24}$$

The first term of (3.24) on the right-hand side can be bounded as follows. The solution z^*

to pursuit step satisfies $\left(A^T(y-A)z^*\right)_\Omega = 0$ by orthogonality. By using $y = A\widehat{x} + \nu$, we can write

$$\begin{aligned}
\|(\widehat{x}-z^*)_\Omega\|_2 &= \left\|\left(\widehat{x}-z^*+A^T(y-A)z^*\right)_\Omega\right\|_2 \\
&= \left\|\left((I-A^TA)(\widehat{x}-z^*)\right)_\Omega + (A^T\nu)_\Omega\right\|_2 \\
&\leq \left\|\left((I-A^TA)(\widehat{x}-z^*)\right)_\Omega\right\|_2 + \left\|(A^T\nu)_\Omega\right\|_2 \\
&\leq \delta_{2k}\|\widehat{x}-z^*\|_2 + \sqrt{1+\delta_k}\|\nu\|_2, \qquad (3.25)
\end{aligned}$$

where the last inequality follows from Lemma 3.3.1 and 3.3.2 due to the fact that $\text{card}(\Omega \cup \text{supp}(\widehat{x})) \leq 2k$ and $\text{card}(\Omega) \leq k$. Substituting (3.25) into (3.24) leads to

$$\|\widehat{x}-z^*\|_2^2 \leq \delta_{2k}^2\|\widehat{x}-z^*\|_2^2 + 2\delta_{2k}\sqrt{1+\delta_k}\|\widehat{x}-z^*\|_2\|\nu\|_2 + (1+\delta_k)\|\nu\|_2^2 + \|(\widehat{x}-u)_{\overline{\Omega}}\|_2^2,$$

which implies

$$(1-\delta_{2k}^2)\|\widehat{x}-z^*\|_2^2 - 2\delta_{2k}\sqrt{1+\delta_k}\|\widehat{x}-z^*\|_2\|\nu\|_2 - (1+\delta_k)\|\nu\|_2^2 - \|(\widehat{x}-u)_{\overline{\Omega}}\|_2^2 \leq 0.$$

The left-hand side of the formula above actually is a quadratic function in terms of $\|\widehat{x}-z^*\|_2$, thus one has

$$\begin{aligned}
\|\widehat{x}-z^*\|_2 &\leq \frac{2\delta_{2k}\sqrt{1+\delta_k}\|\nu\|_2 + \sqrt{4\left(1+\delta_k\right)\|\nu\|_2^2 + 4\left(1-\delta_{2k}^2\right)\|(\widehat{x}-u)_{\overline{\Omega}}\|_2^2}}{2\left(1-\delta_{2k}^2\right)} \\
&\leq \frac{2\delta_{2k}\sqrt{1+\delta_k}\|\nu\|_2 + 2\sqrt{1+\delta_k}\|\nu\|_2 + 2\sqrt{1-\delta_{2k}^2}\|(\widehat{x}-u)_{\overline{\Omega}}\|_2}{2\left(1-\delta_{2k}^2\right)},
\end{aligned}$$

where the last inequality holds since $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for any $a, b \geq 0$. Rearranging

terms, we have

$$\begin{aligned}\|\widehat{x} - z^*\|_2 &\leq \frac{1}{\sqrt{1-\delta_{2k}^2}} \|(\widehat{x}-u)_{\overline{\Omega}}\|_2 + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}\|\nu\|_2 \\ &\leq \frac{1}{\sqrt{1-\delta_{2k}^2}} \|\widehat{x}-u\|_2 + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}\|\nu\|_2.\end{aligned}$$

The desired result follows. ■

Using this property, we now prove the result for NSHTP.

Theorem 3.3.9. *Let σ_1 and σ_m be the largest and smallest singular values of the matrix $A \in \mathbb{R}^{m\times n}$ with $m \ll n$. Suppose that the restricted isometry constant of A satisfies that*

$$\delta_{3k} < \frac{1}{2}.$$

Let $y \coloneqq Ax + \eta$ be the measurements of the signal $x \in \mathbb{R}^n$. Let ϵ be a given parameter satisfying

$$\epsilon > \max\left\{\sigma_1^2,\ \left(\frac{\sigma_1^2-\sigma_m^2}{\frac{1}{2}-\delta_{3k}} - 1\right)\sigma_1^2\right\}. \tag{3.26}$$

If the stepsize λ in NSHTP is chosen such that

$$\epsilon + \sigma_1^2 - \left(\frac{1}{2}-\delta_{3k}\right)\frac{\epsilon+\sigma_1^2}{\sigma_1^2} \ < \ \lambda \ \leq \ \epsilon + \sigma_m^2, \tag{3.27}$$

then the sequence $\{x^n\}$, generated by the NSHTP, satisfies

$$\|x^{p+1} - x_S\|_2 \leq \rho^*\|x^p - x_S\|_2 + \tau^*\|Ax_{\overline{S}} + \eta\|_2, \tag{3.28}$$

where

$$\rho^* = \sqrt{\frac{3}{1-(\delta_{2k})^2}}\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon+\sigma_1^2}\right) < 1,$$

and

$$\tau^* = \frac{\sqrt{3}\lambda\sigma_1}{(\epsilon+\sigma_1^2)\sqrt{1-(\delta_{2k})^2}} + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}.$$

In particular, when x is k-sparse and $\eta = 0$ (exact measurements), then the sequence $\{x^p\}$ converges to x.

Proof. We still write $y = Ax + \eta = Ax_S + \eta'$ with $\eta' = Ax_{\overline{S}} + \eta$, where $S = \mathcal{L}_k(x)$ and $\overline{S}$ is the complement set of S. Note that the intermediate point $\overline{x}^p$ in NSHTP was generated by the NSIHT. Based on a result for NSIHT, i.e., the relation (3.21), we immediately obtain the following inequality:

$$\|\overline{x}^p - x_S\|_2 \leq \sqrt{3}\left[\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x^p - x_S\|_2 + \frac{\lambda\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2\right]. \tag{3.29}$$

Again the relation $\delta_{3k} + \sigma_1^2 - \dfrac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2} > 0$ is guaranteed by $\lambda \leq \epsilon + \sigma_m^2$. Applying Lemma 3.3.8 by setting $\hat{x} = x_S$, $u = \overline{x}^p$ and $z^* = x^{p+1}$ yields

$$\left\|x^{p+1} - x_S\right\|_2 \leq \frac{1}{\sqrt{1 - (\delta_{2k})^2}}\|x_S - \overline{x}^p\|_2 + \frac{\sqrt{1+\delta_k}}{1 - \delta_{2k}}\|\eta'\|_2. \tag{3.30}$$

Combining (3.29) and (3.30) leads to

$$\begin{aligned}\left\|x^{p+1} - x_S\right\|_2 \leq \sqrt{\frac{3}{1 - (\delta_{2k})^2}}&\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x^p - x_S\|_2 \\ &+ \left(\frac{\sqrt{3}\lambda\sigma_1}{\sqrt{1 - (\delta_{2k})^2}\,(\epsilon + \sigma_1^2)} + \frac{\sqrt{1+\delta_k}}{1 - \delta_{2k}}\right)\|\eta'\|_2.\end{aligned} \tag{3.31}$$

To guarantee the performance, it is sufficient to require, under the condition (3.28) and (3.27), that

$$\rho^* := \frac{\sqrt{3}}{\sqrt{1 - (\delta_{2k})^2}}\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right) < 1. \tag{3.32}$$

By the definition of the restricted isometry constant, we see that $\delta_{2k} \leq \delta_{3k}$. This together with the condition $\delta_{3k} < 1/2$ implies that

$$\frac{\sqrt{3}}{2} < \sqrt{1 - (\delta_{3k})^2} \leq \sqrt{1 - (\delta_{2k})^2}.$$

Therefore, to ensure (3.32) is satisfied, it is sufficient to require that

$$\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2} < \frac{1}{2},$$

which is equivalent to

$$\lambda > \epsilon + \sigma_1^2 - \left(\frac{1}{2} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2}.$$

Note that $\lambda \leq \epsilon + \sigma_m^2$ is also required. Therefore the relation (3.32) is guaranteed provided that the stepsize λ satisfies the following condition:

$$\epsilon + \sigma_1^2 - \left(\frac{1}{2} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \lambda \leq \epsilon + \sigma_m^2.$$

In order to guarantee the existence of this range for the stepsize λ, it suffices to require that the right-hand side of the above inequality is strictly larger than its left-hand side, i.e.,

$$\epsilon + \sigma_1^2 - \left(\frac{1}{2} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \epsilon + \sigma_m^2,$$

which is guaranteed by

$$\epsilon > \max\left\{\sigma_1^2,\ \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{2} - \delta_{3k}} - 1\right)\sigma_1^2\right\}.$$

Therefore the desired bound (3.28) is established.

In particular, if $\eta = 0$ and x is k-sparse, the relation (3.31) is reduced to

$$\left\|x^{p+1} - x\right\|_2 \leq \rho^* \left\|x^p - x\right\|_2 \leq (\rho^*)^{p+1}\left\|x^0 - x\right\|_2,$$

where ρ^* is the same as (3.32). It implies that the sequence $\{x^p\}$ generated by NSHTP converges to x as $p \to \infty$. ■

Similar to Theorem 3.3.5, the above result claims that when the measurements and the sparse signal are contaminated with noises, the algorithm NSHTP can still recover

the signal to a certain level of quality. In fact, our results can be further improved by utilizing the newest result concerning the property of hard thresholding operator, i.e., the coefficient $\sqrt{3}$ in Lemma 3.3.4 has been improved to $\frac{\sqrt{5}+1}{2}$ by Zhao in [142]. We omit this minor work here. But in the following chapters, we will use the new result to perform the analysis of hard-thresholding related algorithms.

Before discussing their numerical performances, let us first give some remarks concerning the computation of matrix $(A^TA+\epsilon I)^{-1}$, potential improved analysis, the choices of algorithmic parameter (ϵ, λ), and the convergence rate of the algorithms.

Remark 3.3.1

From the structure of NSIHT and NSHTP, it is sufficient to perform a one-off singular value decomposition (SVD) of the matrix $A \in \mathbb{R}^{m\times n}$ to obtain the inverse of $(A^TA+\epsilon I)^{-1}$. In fact, let $A = U\Sigma V^T$ be the SVD of A where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are two orthogonal matrices, and

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0_{m\times(n-m)} \\ & & \sigma_m & \end{bmatrix}$$

with $\sigma_i, i = 1, \ldots, m$ being the singular values of A satisfying a non-increasing order, i.e., $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$. Then

$$(A^TA+\epsilon I)^{-1} = V \operatorname{diag}\left(\frac{1}{\sigma_1^2+\epsilon}, \cdots, \frac{1}{\sigma_m^2+\epsilon}, \frac{1}{\epsilon}, \cdots, \frac{1}{\epsilon}\right) V^T. \tag{3.33}$$

Thus for any given $\epsilon > 0$, the computation of $(A^TA+\epsilon I)^{-1}$ is straightforward based on the SVD of A. When ϵ is changed, the inverse can be updated easily by changing the diagonal entries in (3.33). The SVD of A is also used for the choice of the algorithmic parameter (ϵ, λ). Alternatively, the computation of matrix inverse $\left(A^TA+\epsilon I\right)^{-1}$ can be avoided, since the Newton-like search direction can be obtained by solving a linear system through the LU decomposition of the matrix. In compressed sensing scenarios, the use of both SVD and LU decomposition for solving the linear system are generally stable,

since the measurement matrix A is supposed to be RIP and the choice of ϵ is not that too small indicating that the problem is well-conditioned in general. In the case of an ill-posed matrix in the linear system, some preconditioning techniques can be applied (e.g., [9, 83]). The key step of the proposed algorithm amounts to $x^{p+1} = \mathcal{H}_k(x^p + \lambda v^p)$, where v^p is the solution to the linear system $\left(A^TA + \epsilon I\right) v^p = A^T(y - Ax^p)$. Each step of NSIHT involves $O\left(mn^2 + n^3\right)$ operations, where A^TA requires mn^2 and LU decomposition of $A^TA + \epsilon I$ requires $O(n^3)$ operations. Solving the pursuit step is equivalent to solving the linear system $[A_{S^{p+1}}]^T y = \left([A_{S^{p+1}}]^T[A_{S^{p+1}}]\right) z$, where $S^{p+1} = \text{supp}(\overline{x}^p)$ and the size of $[A_{S^{p+1}}]^T[A_{S^{p+1}}]$ is $k \times k$. Therefore, each iteration of NSHTP involves a total of $O\left(mn^2 + n^3 + k^3\right)$ operations. The main computational complexity is determined by the target signal length n since we assume $k \ll m \ll n$ in compressed sensing scenarios. Thus our Newton-type algorithms are usually efficient and convenient for medium-sized sparse signal recovery. It has been pointed out in [139] that the IHT and HTP require about $O(mn)$ and $O(mn + n^3)$ operations, respectively. Compared with our algorithms, the gradient-based hard thresholding methods might be more convenient to implement from the viewpoint of computational complexity.

Remark 3.3.2

The convergence results stated in Theorems 3.3.5 and 3.3.9 indicate that the algorithms in this chapter can successfully recover a sparse signal if the algorithmic parameter (ϵ, λ) satisfies (3.15) and (3.16) for NSIHT and (3.26) and (3.27) for NSHTP. However, we cannot expect there is a straightforward way to determine the choice of (ϵ, λ) according to the above conditions due to the difficulty of computing the RICs. In applications, a simple and empirical rule implied from these conditions is to choose (ϵ, λ) such that $\epsilon > \sigma_1^2$ and $\lambda \leq \epsilon + \sigma_m^2$. This simple choice is used in practical implementations of the proposed algorithms (see more discussions in Section 3.4).

Remark 3.3.3

Our analysis is performed based on the condition $\epsilon > \sigma_1^2$ to apply the matrix expansion (3.7). It is worth mentioning that to apply the matrix inverse expansion, i.e., formula (3.7),

the parameter ϵ has to satisfy the condition $\epsilon > \sigma_1^2$ on which our analysis is performed. However, we can see that for any $\epsilon > 0$ the modified Hessian $A^TA + \epsilon I$ is positive definite, which indicates the modified Newton step is always a descent direction. The parameter ϵ can be very small to slightly perturb the matrix A^TA such that $A^TA + \epsilon I \approx A^TA$. Numerical experiments with such a choice of ϵ have been carried out in later sections, while the theoretical analysis needs further investigation.

Remark 3.3.4

The main results in Section 3.3 claim that the proposed algorithms are globally convergent to the best k-term approximation x_S of the original signal x with a linear convergence rate under the standard RIP assumption. The linear convergence ratio ρ can be explicitly calculated under a specific RIP bound. For example, under exact measurements and assuming $\delta_{3k} < 1/3$, the iterates generated by NSHTP are linearly convergent to x_S with rate $\rho = \sqrt{6}/4$, i.e., $\left\|x^{p+1} - x_S\right\|_2 \leq \dfrac{\sqrt{6}}{4} \left\|x^p - x_S\right\|_2$, provided that ϵ and λ are chosen such that

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{3} - \delta_{3k}} - 1\right)\sigma_1^2\right\}$$

and

$$\epsilon + \sigma_1^2 - \left(\frac{1}{3} - \delta_{3k}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \lambda \leq \epsilon + \sigma_m^2.$$

This can be immediately verified by the same analysis at the end of the proof of Theorem 3.3.9. We note that Theorem 3.3.9 only claims the global convergence of the NSHTP. *What can be said about the local convergence rate of NSHTP?* For simplicity, assume that the original signal x is k-sparse, i.e., $x_S = x$, and that the measurements $y = Ax$ are accurate. We now point out that the global convergence stated in Theorem 4.3 actually implies that under the same assumption of the theorem, the NSHTP terminates at x after a finite number of iterations, i.e., the NSHTP finitely stops at $x^p = x$ for some p. Thus

in the language of local convergence, the algorithm terminates at the solution x after a finite number of iterations possesses the fastest local convergence rate, faster than any locally superlinear or quadratic convergence rate. The sequence $\{x^p\}$ generated by NSHTP satisfies that

$$\|x^p - x\|_2 \leq \rho^p \left\|x^0 - x\right\|_2 \to 0,$$

when measurements are exact. The iterate x^p is always k-sparse, this implies that there exists a large integer number $\hat{p}$ such that $\text{supp}(x^{\hat{p}}) = \text{supp}(x)$. In fact, if $\text{supp}(x^{\hat{p}}) \neq \text{supp}(x)$ for any large integer number $\hat{p}$, then there is an index $i_0 \in \text{supp}(x)$ and $i_0 \notin \text{supp}(x^{\hat{p}})$. This means $x_{i_0} \neq 0$ and $(x^{\hat{p}})_{i_0} = 0$, which implies that $\left\|x^{\hat{p}} - x\right\|_2 \geq \left|(x^{\hat{p}})_{i_0} - x_{i_0}\right| = |x_{i_0}| > 0$, contradicting the convergence of x^p to x. Therefore, there must exist an iteration index $\hat{p}$ such that $\text{supp}(x^{\hat{p}}) = \text{supp}(x)$. Since $x^{\hat{p}}$ is the solution to the pursuit step, by optimality one has $\left\|y - Ax^{\hat{p}}\right\|_2 \leq \|y - Ax\|_2 = 0$. This means both $x^{\hat{p}}$ and x are k-sparse solutions to the system $y = Az$. The RIP assumption $\delta_{3k} < 1$ ensures that the linear system $y = Az$ has a unique k-sparse solution. Hence, $x^{\hat{p}} = x$, i.e., the pursuit step in NSHTP terminates at $\hat{p}$-th step. This is quite different from most Newton-type methods for general sparse optimization problems in the literature, e.g., [145], which in general cannot guarantee the Newton-type algorithm finitely terminates at the solution of the problem.

3.4 Numerical Experiments

In this section, we demonstrate some simulation results for the NSIHT and NSHTP algorithms in terms of residual reduction, success frequencies for sparse vector reconstruction, and phase transition in terms of success rates. The measurement matrices generated for experiments are Gaussian random matrices whose entries are independent and identically distributed and follow the standard normal distribution $\mathcal{N}(0, 1)$. Nonzero entries of random sparse vectors also follow $\mathcal{N}(0, 1)$, and their positions are chosen according to a uniform distribution.

3.4.1 Residual Reduction

The value of $r = \|y - Ax\|_2$ is often called the residual. The capability of an algorithm in residual reduction is directly related to the efficiency of the algorithm for vector reconstruction. This experiment is to compare the residual-reduction performance between gradient-based and Newton-type hard thresholding algorithms, and test how the NSIHT and NSHTP algorithms behave in residual reduction as the parameter (ϵ, λ) change.

The first experiment was carried out to compare the residual-reduction performance between gradient-type and Newton-type hard thresholding algorithms. Fig. 3.1 demonstrates an example of the residual change during the iterations of HTP, NSIHT, and NSHTP. The result was obtained from a random matrix $A \in \mathbb{R}^{256\times 512}$ and the exact measurements $y = Ax^*$ where $x^* \in \mathbb{R}^{512}$ is a 50-sparse vector. All algorithms start from $x^0 = 0$. In this experiment, we set

$$\lambda = 5, \quad \epsilon = \max\{\sigma_1^2 + 1, \lambda - \sigma_m^2\}, \tag{3.34}$$

which ensures that $\epsilon > \sigma_1^2$ and $\epsilon + \sigma_m^2 \geq \lambda$. The σ_1 and σ_m are the largest and smallest singular values of the realized measurement matrix, respectively.

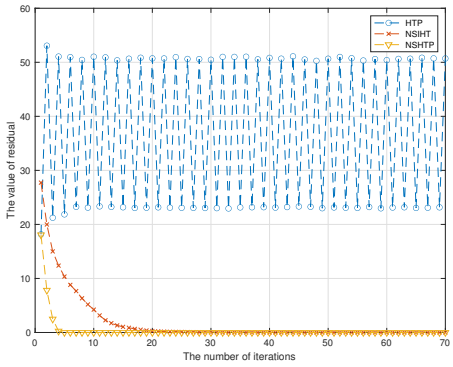


Figure 3.1: The residual reduction of HTP, NSIHT and NSHTP in the course of iterations. The size of measurement matrices is 256×512. The measurements are exact $y = Ax^*$ with sparsity level $k = 50$.

It has been pointed out in [134] that the traditional IHT and HTP algorithms may suffer dramatic numerical oscillation in residual reduction during the course of iterations. In Fig. 3.1, we see that the HTP oscillates and fails to reduce the residual iteratively.

Noting that IHT is not included in this figure due to the fact that its residual diverges very fast. For instance, we take a 3×6 Gaussian random matrix as follows:

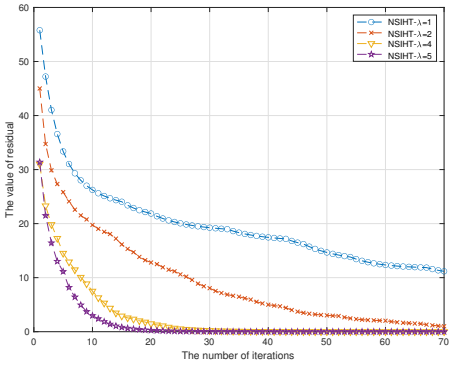
$$A = \begin{bmatrix} 1.5326 & -0.2256 & 0.0326 & 1.5442 & -0.7423 & -0.6156 \\ -0.7697 & 1.1174 & 0.5525 & 0.0859 & -1.0616 & 0.7481 \\ 0.3714 & -1.0891 & 1.1006 & -1.4916 & 2.3505 & -0.1924 \end{bmatrix},$$

and a 1-sparse vector $x = [0, 0, 0, 0, 0, 0.1966]^T$. The sequence of the value of $\|y - Ax\|_2$ generated by IHT with $\lambda = 1$ is 0.3785, 0.9929, 5.0982, 31.6019, 196.0110, 1.2158×10^3, ... The simulations on random examples show that the algorithms NSIHT and NSHTP are generally stabler than their IHT and HTP counterparts, and they can iteratively reduce the residual in broad situations provided the parameters (ϵ, λ) in the algorithms are properly selected.

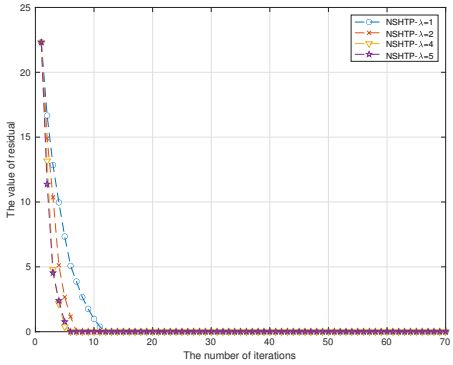
The next experiment demonstrates how the algorithms behave in residual reduction as the parameter ϵ and the step length λ change. The size of measurement matrix A and the sparsity level k are the same as the example mentioned above. In this case, we performed the NSIHT and NSHTP with a few different step lengths: $\lambda = 1, 2, 4, 5$. For every such step length, ϵ is set accordingly as

$$\epsilon = \max\left\{\sigma_1^2 + 1, \lambda - \sigma_m^2\right\}.$$

The result is given in Fig. 3.2. It shows that the algorithms with a relatively larger stepsize usually perform better in residual reduction, and that the change of λ affects the performance of NSIHT more noticeably than the NSHTP.



(a) NSIHT: $\lambda = 1, 2, 4, 5$



(b) NSHTP: $\lambda = 1, 2, 4, 5$

Figure 3.2: Performance of NSIHT and NSHTP in residual reduction with $\epsilon = \sigma_1^2 + 1$ and different step lengths: $\lambda = 1, 2, 4, 5$. The size of measurement matrices is 256×512. The measurements are exact $y = Ax^*$ with sparsity level $k = 50$.

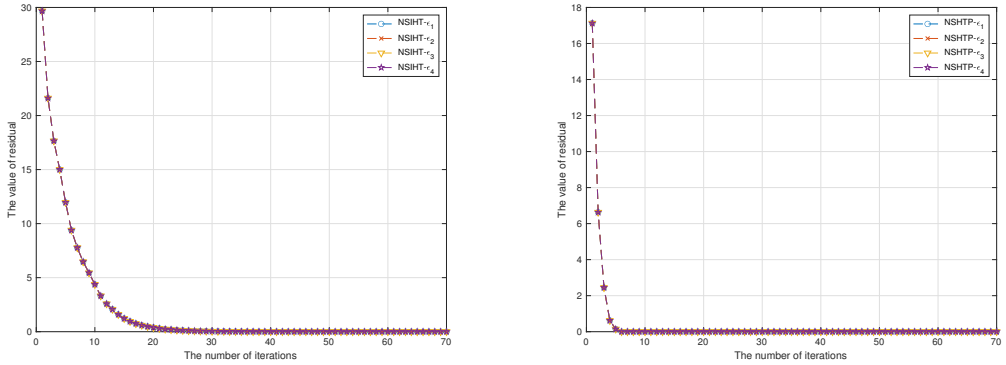
To see the influence of ϵ, the NSIHT and NSHTP were applied to the same random example above with a fixed step length $\lambda^* = 5$ and several different choices of $\epsilon : \epsilon_1 = \epsilon^*, \epsilon_2 = 1.1\epsilon^*, \epsilon_3 = 1.5\epsilon^*$ and $\epsilon_4 = 2\epsilon^*$ where

$$\epsilon^* = \max\{\sigma_1^2 + 1, \lambda^* - \sigma_m^2\}.$$

Such choices satisfy that $\epsilon_i > \sigma_1^2$ and $\epsilon_i + \sigma_m^2 > \lambda^*$, for $i = 1, 2, 3, 4$. The outcome was summarized in Fig. 3.3 which indicates that the residual-reduction performance of NSIHT and NSHTP is insensitive to the above change of ϵ.

3.4.2 Performance of sparse signal recovery

Several experiments were carried out to investigate the vector-reconstruction performance of the proposed algorithms in terms of the influence of the ratio m/n, the influence of the algorithmic parameter (ϵ, λ) and the influence of the number of iterations. In these experiments, we use (3.34) as a default choice of (ϵ, λ). Our algorithms start from the



(a) NSIHT: $\epsilon_1 = \epsilon^*, \epsilon_2 = 1.1\epsilon^*, \epsilon_3 = 1.5\epsilon^*, \epsilon_4 = 2\epsilon^*$ (b) NSHTP: $\epsilon_1 = \epsilon^*, \epsilon_2 = 1.1\epsilon^*, \epsilon_3 = 1.5\epsilon^*, \epsilon_4 = 2\epsilon^*$

Figure 3.3: Performance of NSIHT and NSHTP in residual reduction with different choices of ϵ. The size of measurement matrices is 256×512. The measurements are exact $y = Ax^*$ with sparsity level $k = 50$.

initial point $x^0 = 0$ and adopt the rule

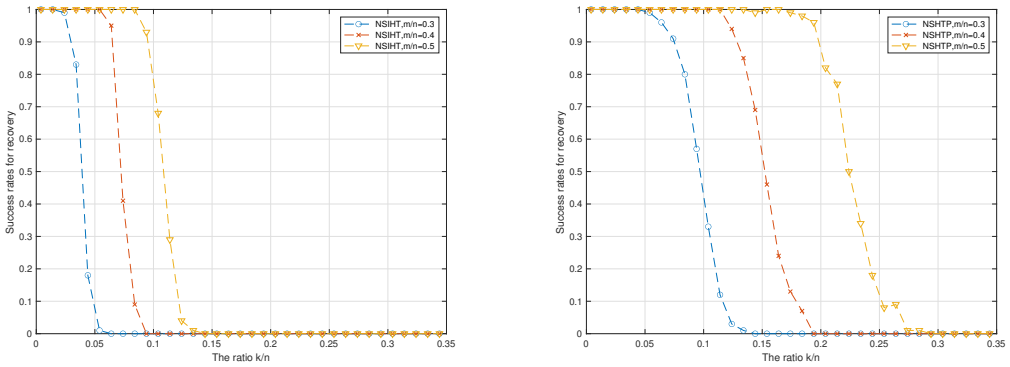
$$\frac{\|x^p - x^*\|_2}{\|x^*\|_2} \leq 10^{-3} \tag{3.35}$$

as the criterion for the success of the recovery of x^*. The success rate is calculated based on 100 randomly realized examples of recovery problems.

Influence of the ratio m/n

To see the influence of the ratio of the number of measurements and signal length on the performance of the algorithms, we performed the experiments with three different ratios: $m/n = 0.5$, 0.4, and 0.3. The random matrices with these ratios are taken as 250×500, 200×500 and 150×500, respectively, and the experiments were carried out with exact measurements $y = Ax$.

The result was given in Fig. 3.4 in which the the maximum number of iterations for NSIHT and NSHTP are set as 35 and 20, respectively. Since the NSHTP usually converges faster than the NSIHT, the maximum number of iterations for NSHTP is set lower than NSIHT. The horizontal axes in Fig. 3.4 are the ratios of the sparsity level k and the



(a) NSIHT: The maximum number of iterations = 35

(b) NSHTP: The maximum number of iterations = 20

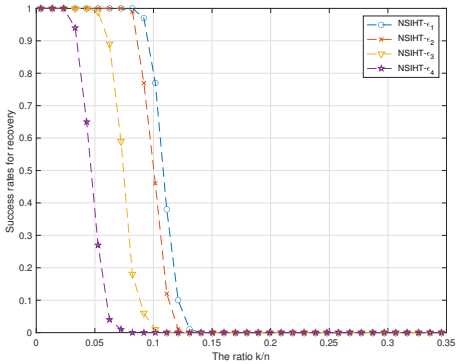
Figure 3.4: Recovery performance of NSIHT and NSHTP with different dimensions of matrices. The sizes of measurement matrices are 150×500, 200×500 and 250×500. The measurements are exact: $y = Ax^*$.

signal length n, and the vertical axes record the success rate of recovery. For each given ratio k/n, the success rate is calculated based on 100 randomly generated examples of the recovery problems. It can be observed from the figure that the lower the ratio m/n (i.e., the less number of measurements are available), the narrower the range of sparse signals can be recovered via these algorithms, and the NSHTP usually performs better than NSIHT for each given ratio m/n, and the NSHTP requires less number of iterations than NSIHT to recover the signal.

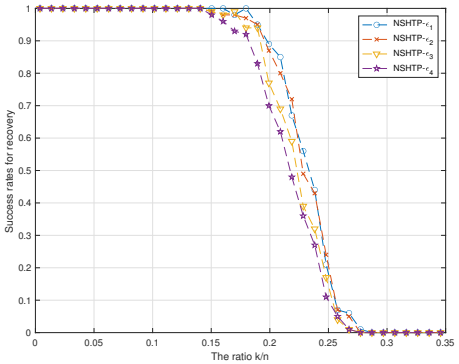
Influence of (ϵ, λ)

The recovery performance of the algorithms with different choices of (ϵ, λ) was also examined through simulations. The size of the random matrices is still 256×512. The success rate corresponding to each ratio k/n was calculated by performing the algorithms on 100 randomly realized examples of recovery problems with accurate measurements of sparse signals. The maximum numbers of iterations for NSIHT and NSHTP are set as 35 and 20 respectively.

Fig. 3.5 demonstrates the results for the algorithms with fixed $\lambda^* = 5$ and several different values of $\epsilon = \epsilon^*, 1.1\epsilon^*, 1.5\epsilon^*, 2\epsilon^*$, where ϵ^* is given by $\epsilon^* = \max\left\{\sigma_1^2 + 1, \lambda^* - \sigma_m^2\right\}$.

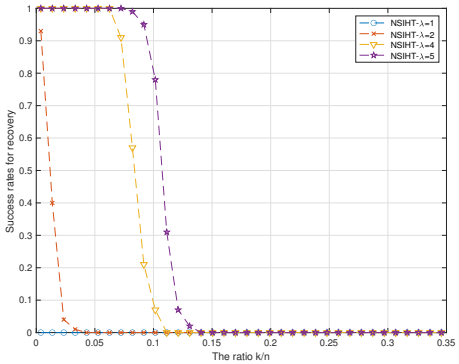


(a) NSIHT with different ϵ

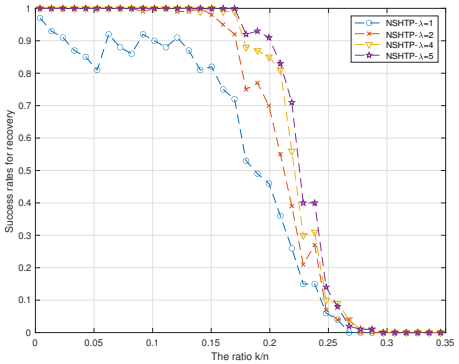


(b) NSHTP with different ϵ

Figure 3.5: Recovery performance of NSIHT and NSHTP with a fixed step length ($\lambda^* = 5$) and different choices of ϵ ($\epsilon_1 = \epsilon^*$, $\epsilon_2 = 1.1\epsilon^*$, $\epsilon_3 = 1.5\epsilon^*$, $\epsilon_4 = 2\epsilon^*$). The size of measurement matrices is 256×512, and measurements are exact $y = Ax^*$. Maximum iterations for NSIHT and NSHTP are 35 and 20, respectively.



(a) NSIHT with $\lambda = 1, 2, 4, 5$



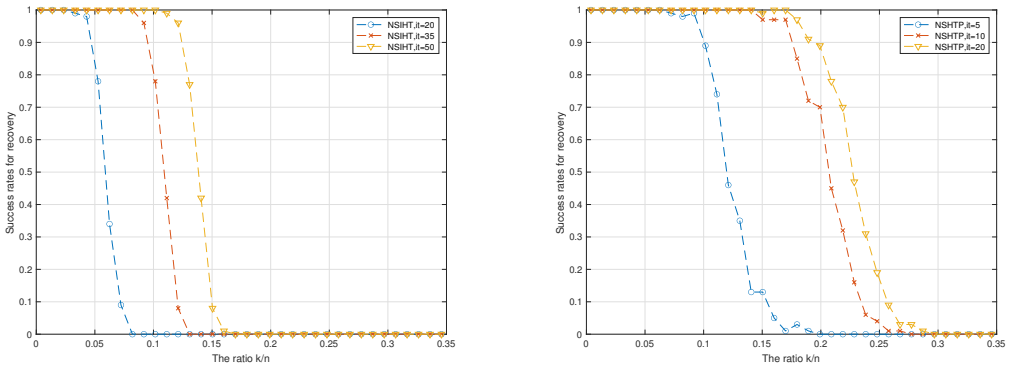
(b) NSHTP with $\lambda = 1, 2, 4, 5$

Figure 3.6: Recovery performance of NSIHT and NSHTP with different choices of stepsizes. The size of measurement matrices is 256×512, and measurements are exact $y = Ax^*$.

Fig. 3.6 demonstrates the results for the algorithms with several choices of the step lengths and ϵ given by $\epsilon = \max\left\{\sigma_1^2 + 1, \lambda - \sigma_m^2\right\}$ accordingly. It can be observed that the overall recovery performance of the algorithms is more sensitive to the change of λ than that of ϵ. Under such a choice of ϵ, we see that NSIHT and NSHTP perform very poorly with $\lambda = 1$.

The influence of the number of iterations

We also investigate the influence of the prescribed maximum number of iterations to the performance of NSIHT and NSHTP algorithms. The results were given in Fig. 3.7 from which it can be seen that using a pursuit step in NSHTP does admit a certain advantage over the NSIHT in the sense that the NSHTP usually requires less number of iterations than NSIHT to recover a signal.



(a) NSIHT with 20, 35, 50 iterations

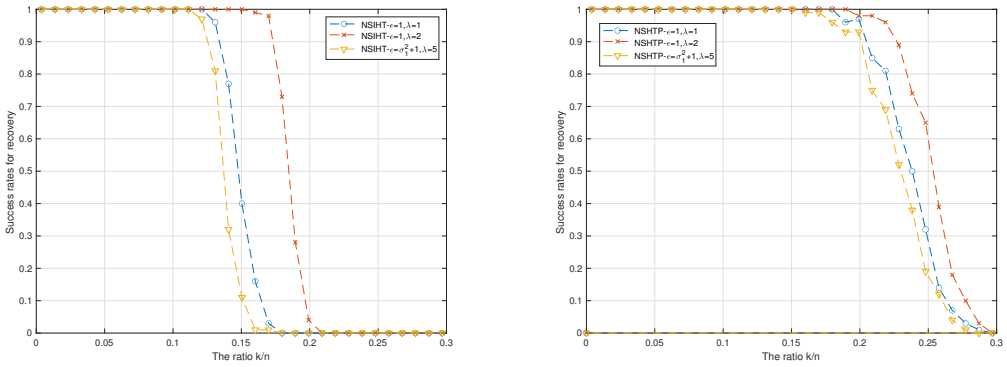
(b) NSHTP with 5, 10, 20 iterations

Figure 3.7: Comparison of NSIHT and NSHTP by performing different numbers of iterations. The size of measurement matrices is 256×512, and measurements are exact $y = Ax^*$.

Different choices of (ϵ, λ)

In addition, we also perform experiments of NSIHT and NSHTP with small values of ϵ, i.e., $\epsilon = 1, 2$ with $\lambda = 1$, and compare them with the algorithmic parameters chosen as (3.34). The theoretical analysis performed for NSIHT and NSHTP in this chapter is based on the condition $\epsilon > \sigma_1^2$ in order to apply the matrix inverse expansion (3.7). However, we

note that for any $\epsilon > 0$, the direction $-\left(\nabla^2 f(x) + \epsilon I\right)^{-1} \nabla f(x)$ is a theoretical descent direction since $\nabla^2 f(x) + \epsilon I$ is positive definite. In this experiment, the size measurement matrix is still set as 256×512 and we set $y = Ax^*$ as the measurements. The results in Figure 3.8 indicate that NSIHT and NSHTP with small ϵ are still able to recover the target signal, although their theoretical convergence requires further studies.



(a) NSIHT: 50 iterations

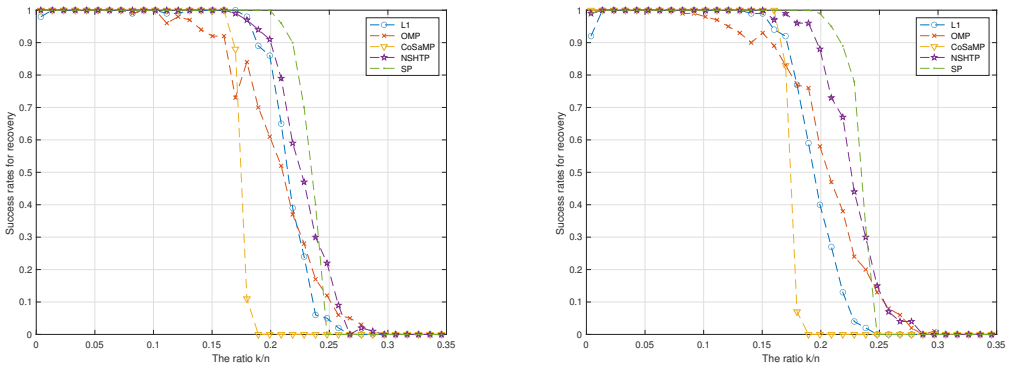
(b) NSHTP: 30 iterations

Figure 3.8: Comparison of NSIHT and NSHTP with different algorithmic parameters (ϵ, λ). The size of measurement matrix is 256×512, and the measurements are exact, i.e., $y = Ax^*$.

Comparisons between existing algorithms

The experiments were also carried out to compare the NSHTP with a few mainstream algorithms such as ℓ_1-minimization, orthogonal matching pursuit (OMP), compressive sampling matching pursuit (CoSaMP) and the subspace pursuit (SP). In our experiments, the size of random matrices is still 256×512 and the signal length is 512. We compare the performance of the algorithms with both exact and inexact measurements. For inexact measurements, we set $y = Ax^* + 0.001\eta$ where η is a standard Gaussian random vector. For each sparsity level, the success rate of an algorithm is calculated based on 100 random samples of the problem. The total number of iterations for CoSaMP, SP and NSHTP algorithms is set to be 30, whereas the OMP is terminated after k iterations which is determined by the sparsity level of x^*, i.e., $k = \|x^*\|_0$. The parameters (ϵ, λ) in NSHTP

are given by (3.34).



(a) Exact measurements: $y = Ax^*$ (b) Inexact measurements: $y = Ax^* + 0.001\eta$

Figure 3.9: Comparison of NSHTP and some existing algorithms, such as ℓ_1-minimization, OMP, CoSaMP and SP. The size of generated Gaussian random matrix is 256×512. The number of iterations for CoSaMP, SP and NSHTP is 30, whereas OMP is terminated after k iterations.

The outcome of the experiment is summarized in Fig. 3.9 which shows that the recovery performance of the NSHTP is comparable to the mainstream algorithms ℓ_1-minimization, CoSaMP, OMP and SP. It can be also seen that the NSHTP is less sensitive to the noises than ℓ_1-minimization. It may outperform ℓ_1-minimization for signal recovery in a noisy environment.

3.4.3 Phase transition

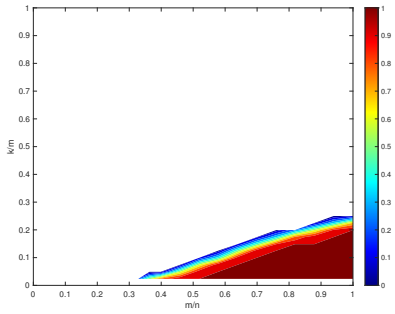
Experiments in this section demonstrate the phase transition in terms of sparse signal recovery and time consumption for both gradient-based and Newton-step-based hard thresholding algorithms. In this experiment, the length n of the target signal vector x^* is fixed to be $n = 512$. At each value of the number of measurements m, we take 40 different sparsity levels uniformly between 1 and m, i.e., $k = \left\lfloor \frac{1}{40}m \right\rfloor, \ldots, \left\lfloor \frac{39}{40}m \right\rfloor, m$. For the number of measurements m, we require $m \geq 40$ to ensure that there exist at least 40 different sparsity levels. A total of 20 different choices of m are taken, where 10 uniformly from $m \in [40, 0.4n]$ and 10 uniformly from $m \in (0.4n, n]$. The measurement matrix A is a

Gaussian random matrix whose entries follow the $\mathcal{N}(0,1)$, and the measurements of x^* are inexact, i.e., $y = Ax^* + 0.001\eta$, where η is a standard Gaussian random vector. For IHT and HTP, we take two choices of stepsize: $\lambda = 0.001$ and $\lambda = 0.01$, respectively. And for NSIHT and NSHTP, we adopt two choices of algorithmic parameters (ϵ, λ) for different purposes, e.g., $\epsilon = 1, \lambda = 2$ is to slightly perturb the singular Hessian and $\epsilon = \sigma_1^2 + 1, \lambda = 5$ is taken based on our theoretical results. The maximum iterations for IHT and NSIHT are 50, while for HTP and NSHTP are 30. The algorithm terminates whether when meeting the criterion (3.35) or when the maximum of iterations is reached. The recovery counts as successful if the criterion (3.35) is satisfied. At a pair of (m, k), success rate is calculated based on 50 trials. In Fig. 3.10 and 3.11, the horizontal axis represents over-sampling rate m/n, while vertical axis represents the under-sampling rate k/m. The range of them both is between 0 and 1, and the area $(k/m) \times (m/n) = [0, 1]^2$ is called the phase transition region.

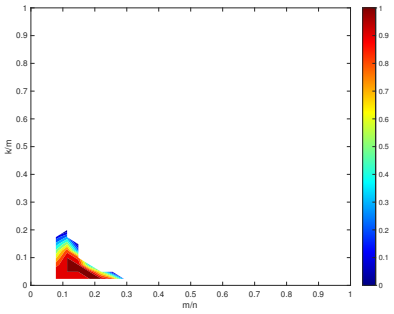
Fig. 3.10 and 3.11 show phase transition of IHT, HTP, NSIHT, and NSHTP in terms of success rate and average time consumption for successful recovery, respectively. The larger the region covered, the better the performance of signal reconstruction. A general trend for phase transition of an algorithm in terms of success rate is that as the ratio m/n increases, the algorithm can recover the signals with relatively a high sparsity level. However, this phenomenon does not hold for IHT and NSIHT with our experimental settings. From these two graphs, we can see that a) the hard thresholding methods with a pursuit step surpass their counterparts without a pursuit step in the success rate of signal recovery, b) Newton-type hard thresholding algorithms are efficient, stable, and robust for sparse signal recovery, and c) NSHTP is more stable than HTP, although the latter takes slightly less time.

3.4.4 Main findings

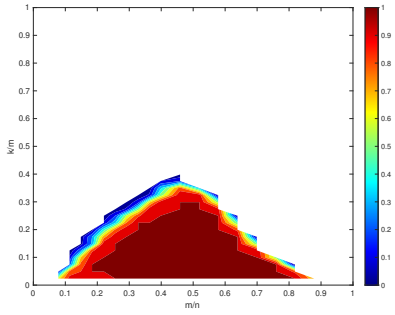
The main findings from empirical results of NSIHT and NSHTP can be summarized as follows.



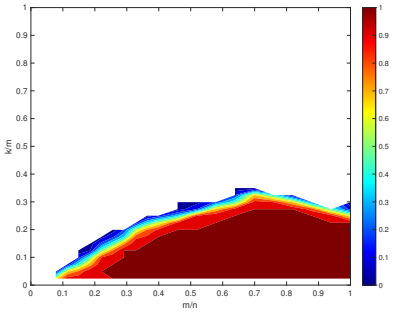
(a) IHT with $\lambda = 0.001$



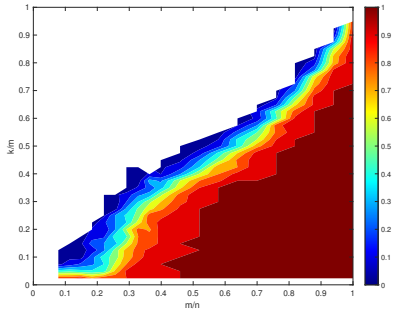
(b) IHT with $\lambda = 0.01$



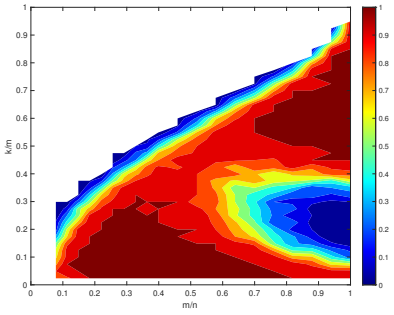
(c) NSIHT with $\epsilon = 1, \lambda = 2$



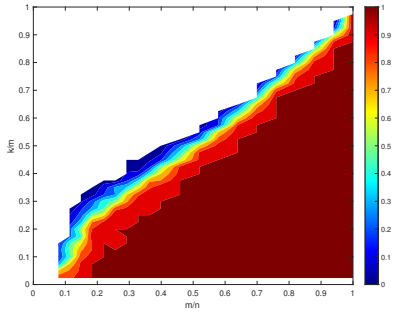
(d) NSIHT with $\epsilon = \sigma_1^2 + 1, \lambda = 5$



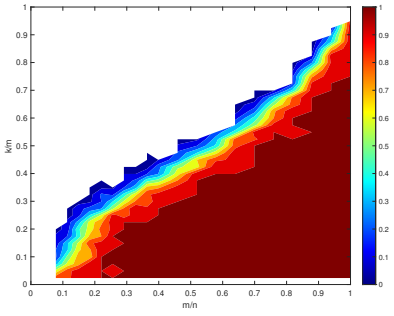
(e) HTP with $\lambda = 0.001$



(f) HTP with $\lambda = 0.01$

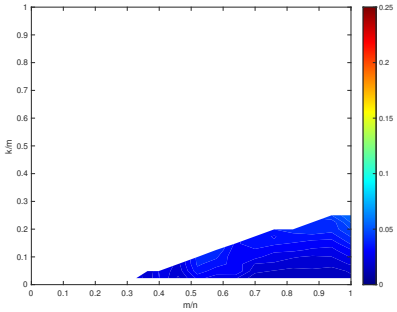


(g) NSHTP with $\epsilon = 1, \lambda = 2$

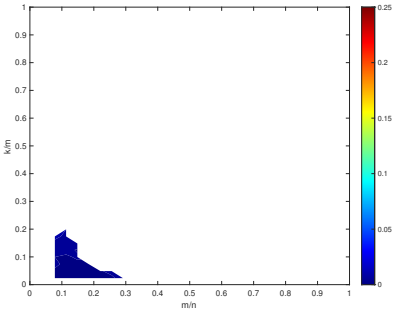


(h) NSHTP with $\epsilon = \sigma_1^2 + 1, \lambda = 5$

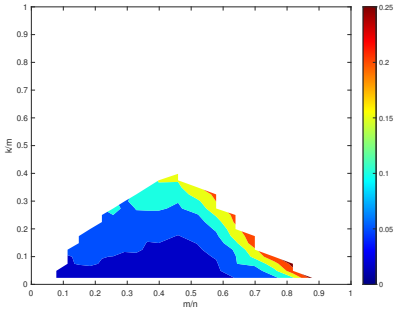
Figure 3.10: Phase transition of IHT, HTP, NSIHT and NSHTP with respect to success rate for sparse signal recovery.



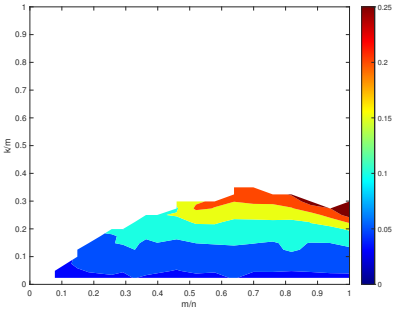
(a) IHT with $\lambda = 0.001$



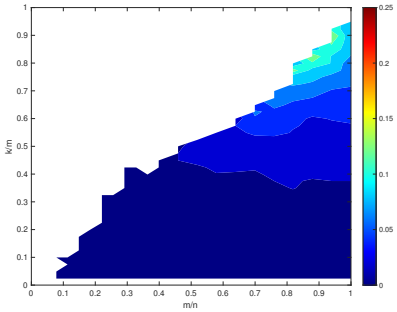
(b) IHT with $\lambda = 0.01$



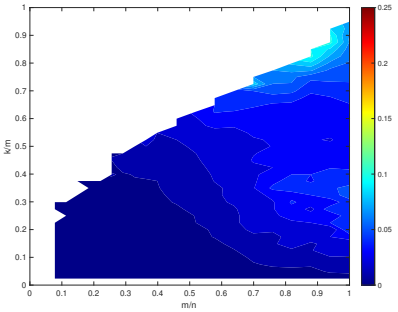
(c) NSIHT with $\epsilon = 1, \lambda = 2$



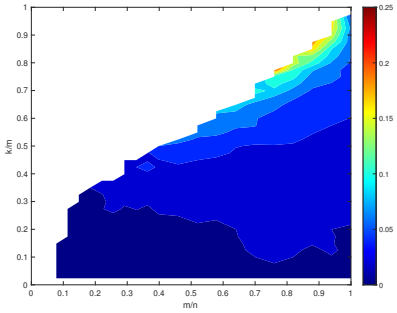
(d) NSIHT with $\epsilon = \sigma_1^2 + 1, \lambda = 5$



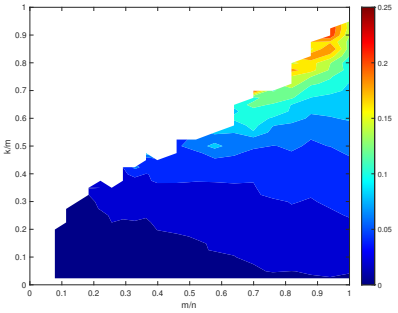
(e) HTP with $\lambda = 0.001$



(f) HTP with $\lambda = 0.01$



(g) NSHTP with $\epsilon = 1, \lambda = 2$



(h) NSHTP with $\epsilon = \sigma_1^2 + 1, \lambda = 5$

Figure 3.11: Phase transition of HTP and NSHTP with respect to time consumption.

- Newton-step-based hard thresholding algorithms are generally more stable than classic hard thresholding methods in residual reduction with suitably chosen parameters (ϵ, λ).

- Relatively speaking, NSIHT and NSHTP are more sensitive to the change in λ than in ϵ, in both residual reduction and sparse signal recovery.

- Our proposed algorithms are able to locate the target sparse vector with different choices of parameters (ϵ, λ), such as $\lambda = 5, \epsilon = \max\left\{\sigma_1^2 + 1, \lambda - \sigma_m^2\right\}$ according to our theoretical analysis and $\lambda = 1, \epsilon = 1$ for slight Hessian perturbation.

- Newton-step-based hard thresholding methods require a little more time than gradient-based hard thresholding methods, but they are usually stabler for sparse signal recovery.

3.5 Conclusion

In this chapter, we introduced a class of Newton-step-based hard thresholding algorithms that adopt a Newton-like search direction. The unique feature of the proposed algorithms is that the traditional steepest descent search direction is replaced by a Newton-like search direction. We have proved that with proper choices of the parameter and stepsize, the proposed algorithms can recover a sparse signal if the measurement matrix satisfies the standard restricted isometry property. The empirical results indicate that the new algorithms can compete with several existing algorithms for signal recovery and stabilize the residual reduction during iterations.

CHAPTER 4

NEWTON-TYPE OPTIMAL K-THRESHOLDING ALGORITHMS

4.1 Introduction

In Chapter 3, we have shown that the combination of a Newton-like method and the hard thresholding is efficient for sparse signal recovery from theoretical and practical perspectives. The hard thresholding approach is easy to implement, but it may dramatically increase the objective function, especially when imposed on a highly non-compressible vector. Zhao [134] proposed a novel thresholding technique called optimal thresholding which overcomes the inherent drawback of hard thresholding approach. Inspired by the remarkable performance of both Newton-type methods and such a novel thresholding technique, we propose the so-called Newton-type optimal k-thresholding in this chapter[1].

We consider the specific problem for sparse signal recovery

$$\min_x \left\{ \frac{1}{2}\|y - Ax\|_2^2 : \|x\|_0 \leq k \right\}. \tag{4.1}$$

The key step of the iterative hard thresholding (IHT) algorithm for solving (4.1) can be formulated as

$$x^{\mathrm{p}+1} = \mathcal{H}_k \left(x^p + \lambda A^T (y - Ax^p) \right), \tag{4.2}$$

[1]The main result in this chapter has been published (jointly with Prof. Yun-Bin Zhao) on the Journal of Operations Research Society of China, pp. 1-23, 2022 [85]

where $\lambda > 0$ is a stepsize, and the hard thresholding operator $\mathcal{H}_k$ retains the largest k magnitudes of a vector and zeroes out the others. Since the IHT with a fixed stepsize is not performed as well as expected, several enhanced versions of (4.2) have been proposed, such as the normalized iterative hard thresholding (NIHT) [19] using an adaptive stepsize, and the hard thresholding pursuit (HTP) [60] including a pursuit step. However, such improvements have not changed the nature of hard thresholding, that it per se is independent of the objective function. As pointed out in [134, 139], the weakness of the hard thresholding operator $\mathcal{H}_k$ is that when applied to a non-sparse iterate generated by the classic gradient method, it may cause an ascending value of the objective of (4.1) at the thresholded vector, compared to the objective value at its unthresholded counterpart. As a result, direct use of the hard thresholding operator to a non-sparse or non-compressible vector in the course of an algorithm may lead to significant numerical oscillation and divergence of the algorithm. To overcome such an inherent drawback, Zhao [134, 139] has proposed the optimal k-thresholding technique which performs thresholding and objective-reduction simultaneously. The key step of optimal k-thresholding (OT) algorithm for solving (4.1) is written as

$$x^{p+1} = \mathcal{Z}_k^{\#}\left(x^p + \lambda A^T\left(y - Ax^p\right)\right), \tag{4.3}$$

where the optimal k-thresholding operator $\mathcal{Z}_k^{\#}(\cdot)$ retains the k entries that minimize the objective function. Empirical results in [134] indicate that this approach has significantly overcome the numerical oscillation that appeared in the hard thresholding pursuit (HTP) algorithm. The search direction in (4.3) is the steepest descent direction, but it can be replaced by another direction provided that it is a descent direction of the objective function. This motivates us to integrate the optimal k-thresholding and a Newton-type search direction to develop an algorithm that may alleviate or eliminate the drawback of hard thresholding and enhance the numerical performance of the Newton-type hard thresholding method.

The main contribution of this chapter can be summarized as follows.

i. We introduce another class of Newton-like methods for sparse signal recovery which incorporates the Newton-type search direction and the novel optimal thresholding technique. Algorithms called Newton-type optimal k-thresholding (NTOT) and its enhanced versions, NTROT and NTROTP, are described in Algorithms 4.1-4.3, respectively.

ii. The guaranteed performance including convergence of the proposed algorithms for sparse signal recovery is shown in terms of appropriate choices of the algorithmic parameters and the restricted isometry property (RIP). The main results for the proposed algorithms are summarized in Theorems 4.3.4, 4.3.8 and 4.3.9, respectively.

iii. We conducted numerical experiments to compare the proposed algorithms with several commonly-used algorithms in terms of sparse signal recovery. Empirical results show that Newton-type optimal thresholding methods are more efficient than Newton-type hard thresholding methods, and they are also comparable to several mainstream algorithms in this field.

4.2 Algorithms

The key iterative procedure of Newton-step-based hard thresholding methods for solving the problem (4.1) is

$$x^{p+1} = \mathcal{H}_k \left(x^p + \lambda \left(A^T A + \epsilon I \right)^{-1} A^T \left(y - A x^p \right) \right), \tag{4.4}$$

where $\lambda > 0$ is the stepsize, $\epsilon > 0$ is a parameter to modify the singular Hessian $A^T A$. To further improve the performance of (4.4), we equip the optimal k-thresholding operator $\mathcal{Z}_k^{\#}(\cdot)$, leading to

$$x^{p+1} = \mathcal{Z}_k^{\#} \left(x^p + \lambda \left(A^T A + \epsilon I \right)^{-1} A^T \left(y - A x^p \right) \right). \tag{4.5}$$

Given a vector $u \in \mathbb{R}^n$, $\mathcal{Z}_k^{\#}(u) = u \otimes w^*$ (the Hadamard product of two vectors), where the vector $w^* \in \mathbb{R}^n$ is the optimal solution to the following quadratic 0-1 optimization problem:

$$w^* := \arg\min_w \left\{ \|y - A(u \otimes w)\|_2^2 : \mathbf{e}^T w = k, w \in \{0,1\}^n \right\}, \tag{4.6}$$

where $\mathbf{e} = (1, \ldots, 1)^T \in \mathbb{R}^n$ is the vector of ones, and $\{0,1\}^n$ denotes the set of n-dimensional 0-1 vectors. Based on (4.5) and the definition of $\mathcal{Z}_k^{\#}$, we can develop the Newton-type optimal k-thresholding (NTOT) algorithm, stated as Algorithm 4.1.

Algorithm 4.1 Newton-Type Optimal k-Thresholding (NTOT)

- Input: measurement matrix A, measurement vector y, sparsity level k, parameter $\epsilon > 0$, stepsize λ, and an initial point x^0.
- Iteration:

$$u^p = x^p + \lambda(A^T A + \epsilon I)^{-1} A^T (y - Ax^p), \tag{NTOT$_1$}$$
$$w^p = \arg\min_w \{ \|y - A(u^p \otimes w)\|_2^2 : \ \mathbf{e}^T w = k, \ w \in \{0,1\}^n \}, \tag{NTOT$_2$}$$
$$x^{p+1} = u^p \otimes w^p. \tag{NTOT$_3$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Since solving the 0-1 problem (4.6) is generally expensive, Zhao [134] suggests solving its tightest convex relaxation, i.e.,

$$\overline{w} := \arg\min_w \left\{ \|y - A(u \otimes w)\|_2^2 : \mathbf{e}^T w = k, 0 \le w \le \mathbf{e} \right\}, \tag{4.7}$$

where $0 \le w \le \mathbf{e}$ is the convex hull of $w \in \{0,1\}^n$. If the solution to (4.7) is k-sparse, then the vector $u \otimes \overline{w}$ is k-sparse. However, the optimal solution $\overline{w}$ to (4.7) may not necessarily be k-sparse, thus we use hard thresholding operator $\mathcal{H}_k$ to truncate the iterate so that it satisfies the constraint of the problem (4.1). Thereby, we have the Newton-type relaxed optimal k-thresholding (NTROT) algorithm, shown in Algorithm 4.2.

To further stabilize the NTROT, a pursuit step can be performed after NTROT$_3$, leading to Algorithm 4.3, which is the main algorithm concerned this chapter.

Algorithm 4.2 Newton-Type Relaxed Optimal k-Thresholding (NTROT)

- Input: measurement matrix A, measurement vector y, sparsity level k, parameter $\epsilon > 0$, stepsize λ, and an initial point x^0.
- Iteration:

$$u^p = x^p + \lambda(A^TA + \epsilon I)^{-1}A^T(y - Ax^p), \tag{NTROT$_1$}$$
$$\overline{w}^p = \arg\min_w\{\|y - A(u^p \otimes w)\|_2^2 : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\}, \tag{NTROT$_2$}$$
$$x^{p+1} = \mathcal{H}_k(u^p \otimes \overline{w}^p). \tag{NTROT$_3$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Algorithm 4.3 Newton-Type Relaxed Optimal k-Thresholding Pursuit (NTROTP)

- Input: measurement matrix A, measurement vector y, sparsity level k, parameter $\epsilon > 0$, stepsize λ, and an initial point x^0.
- Iteration:

$$u^p = x^p + \lambda(A^TA + \epsilon I)^{-1}A^T(y - Ax^p), \tag{NTROTP$_1$}$$
$$\overline{w}^p = \arg\min_w\{\|y - A(u^p \otimes w)\|_2^2 : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\}, \tag{NTROTP$_2$}$$
$$S^{p+1} = \mathcal{L}_k(u^p \otimes \overline{w}^p), \tag{NTROTP$_3$}$$
$$x^{p+1} = \arg\min_w\{\|y - Az\|_2^2 : \mathrm{supp}(z) \subseteq S^{p+1}\}. \tag{NTROTP$_4$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

4.2.1 Extension to the general case

Newton-type optimal thresholding algorithms also can be generalized to an ordinary form for solving the general problem $\min_x \{f(x) : \|x\|_0 \leq k\}$. Following the essential idea of optimal thresholding, the optimal k-thresholding operator with respect to the problem above is defined as $\mathcal{Z}_k^{\#}(u) = u \otimes w^*$, where w^* is the optimal solution of the following problem

$$\min_w \{f(u^p \otimes w) : \ \mathbf{e}^T w = k, \ w \in \{0,1\}^n\}.$$

The relaxation problem is given naturally as

$$\min_w \{f(u^p \otimes w) : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\}.$$

Similar to the GNSHTP algorithm (see Algorithm 3.4), we can write a generalized version of NTROTP, given in Algorithm 4.4.

Algorithm 4.4 Generalized Newton-Type Relaxed Optimal k-Thresholding Pursuit (GNTROT)

- Input: objective function $f(x)$, sparsity level k, parameter $\epsilon \geq 0$, stepsize λ, and an initial point x^0.
- Iteration: If $\nabla^2 f(x^p) \succ 0$, set $\epsilon = 0$; Otherwise, choose a suitable ϵ such that $\nabla^2 f(x^p) + \epsilon I \succ 0$. Then set

$$\begin{aligned} u^p &= x^p - \lambda(\nabla^2 f(x^p) + \epsilon I)^{-1} \nabla f(x^p), \\ \overline{w}^p &= \arg\min_w \{f(u^p \otimes w) : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\}, \\ S^{p+1} &= \mathcal{L}_k(u^p \otimes \overline{w}^p), \\ x^{p+1} &= \arg\min_z \{f(z) : \mathrm{supp}(z) \subseteq S^{p+1}\}. \end{aligned}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

4.3 Theoretical Analysis

We perform theoretical analysis for the guaranteed performance and the convergence of NTOT, NTROT and NTROTP for sparse signal recovery. Sufficient conditions are established by the restricted isometry property (RIP) and algorithmic parameters. Without loss of generality, our analysis is carried out with noisy measurements, i.e., $y = Ax + \eta$, where x can be well approximated by its largest k magnitudes i.e., $x \approx x_S$. Our results include the case where the target signal x is k-sparse and the measurements are exact as a special case.

4.3.1 Analysis of NTOT in noisy scenarios

The following three lemmas are very helpful to show the main result in this section. The first two are recalled from Chapter 3, and the third one is extracted from the proof of relaxed optimal k-thresholding pursuit (ROTP) algorithm in [134].

Lemma 4.3.1 [84]. Let $A \in \mathbb{R}^{m\times n}$ with $m \ll n$ be a measurement matrix satisfying the RIP. Given a vector $u \in \mathbb{R}^n$ and an index set $\Omega \subset [N]$, if (ϵ, λ) is chosen such that $\epsilon > \sigma_1^2$ and $\lambda \leq \epsilon + \sigma_m^2$, where σ_1, σ_m are the largest and smallest singular values of the matrix A, respectively, then one has

$$\left\|\Big((I - \lambda(A^TA + \epsilon I)^{-1}A^TA)u\Big)_\Omega\right\|_2 \leq (\delta_t + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2})\|u\|_2$$

provided that $\mathrm{card}(\Omega \cup \mathrm{supp}(u)) \leq t$, where t is a certain integer number.

Lemma 4.3.2 [84]. Let σ_1 denote the largest singular value of matrix A. For any vector $u \in \mathbb{R}^n$, one has

$$\left\|\Big(A^TA + \epsilon I\Big)^{-1} A^Tu\right\|_2 \leq \frac{\sigma_1}{\epsilon + \sigma_1^2}\|u\|_2 .$$

Lemma 4.3.3 [134]. Let $y = A\hat{x} + \nu$ be the measurements of the k-sparse vector $\hat{x} \in \mathbb{R}^n$, and let $u \in \mathbb{R}^n$ be an arbitrary vector. Let $\mathcal{Z}_k^{\#}(u)$ be the optimal k-thresholding vector of

u. Then for any k-sparse binary vector $\hat{w} \in \{0,1\}^n$ satisfying $\mathrm{supp}(\hat{x}) \subseteq \mathrm{supp}(\hat{w})$, one has

$$\left\|\mathcal{Z}_k^{\#}(u) - \hat{x}\right\|_2 \leq \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\|(\hat{x}-u)\otimes\hat{w}\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}}\|\nu\|_2. \tag{4.8}$$

The following proof was shown by Zhao in the proof of Theorem 4.8 in [134]. We include it here for completeness.

Proof. Let $y = A\hat{x} + \nu$. By the definition of restricted isometry constant, we can write

$$\begin{aligned}\left\|y - A\mathcal{Z}_k^{\#}(u)\right\|_2 &= \left\|A(\hat{x} - \mathcal{Z}_k^{\#}(u)) + \nu\right\|_2 \\ &\geq \left\|A(\hat{x} - \mathcal{Z}_k^{\#}(u))\right\|_2 - \|\nu\|_2 \\ &\geq \sqrt{1-\delta_{2k}}\left\|\hat{x} - \mathcal{Z}_k^{\#}(u)\right\|_2 - \|\nu\|_2, \end{aligned}\tag{4.9}$$

where the last inequality holds since the vector $\hat{x} - \mathcal{Z}_k^{\#}(u)$ is $2k$-sparse. By the definition of optimal thresholding operator, we can write

$$w' = \arg\min_{w}\left\{\|y - A(u\otimes w)\|_2^2 : \mathbf{e}^T w = k,\ w \in \{0,1\}^n\right\}.$$

The vector $\hat{w}$ satisfies $\mathrm{supp}(\hat{x}) \subseteq \mathrm{supp}(\hat{w})$, which implies $\hat{x} = \hat{x}\otimes\hat{w}$. Since $\hat{w}$ is a k-sparse vector, thus we have

$$\begin{aligned}\left\|y - A\mathcal{Z}_k^{\#}(u)\right\|_2 &= \|y - A(u\otimes w')\|_2 \\ &\leq \|y - A(u\otimes\hat{w})\|_2 \\ &= \|A\left((\hat{x}-u)\otimes\hat{w}\right)\|_2 + \|\nu\|_2 \\ &\leq \sqrt{1+\delta_k}\|(\hat{x}-u)\otimes\hat{w}\|_2 + \|\nu\|_2. \end{aligned}\tag{4.10}$$

Combining (4.9) and (4.10) and rearranging terms leads to (4.8). ■

We now state and show the sufficient condition for the guaranteed performance of NTOT in noisy settings.

Theorem 4.3.4. *Let $y = Ax + \eta$ be the measurements of the signal $x \in \mathbb{R}^n$ with measurement error η. Let $S = \mathcal{L}_k(x)$ and σ_1 and σ_m be, respectively, the largest and smallest singular values of the matrix $A \in \mathbb{R}^{m\times n}$. Suppose that the restricted isometry constant of A satisfies*

$$\delta_{2k} < 0.5436,$$

and that ϵ is a given parameter satisfying

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\sqrt{\frac{1-\delta_{2k}}{1+\delta_k}} - \delta_{2k}} - 1\right)\sigma_1^2\right\}. \tag{4.11}$$

If the stepsize λ in NTOT is chosen such that

$$\epsilon + \sigma_1^2 + \left(\delta_{2k} - \sqrt{\frac{1-\delta_{2k}}{1+\delta_k}}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \lambda \leq \epsilon + \sigma_m^2, \tag{4.12}$$

then the sequence $\{x^p\}$ generated by the NTOT satisfies

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho \left\|x^p - x_S\right\|_2 + \tau \left\|Ax_{\overline{S}} + \eta\right\|_2, \tag{4.13}$$

where

$$\rho = \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\left(\delta_{2k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)$$

and

$$\tau = \frac{1}{\sqrt{1-\delta_{2k}}}\left(\frac{\lambda\sigma_1^2\sqrt{1+\delta_k}}{\epsilon + \sigma_1^2} + 2\right).$$

In particular, when x is k-sparse and $\eta = 0$, the sequence $\{x^p\}$ converges to x.

Proof. Let $\hat{w}$ be a k-sparse binary vector such that $S \subseteq \text{supp}(\hat{w})$, where $S = \mathcal{L}_k(x)$. We know that $x_S = x_S \otimes \hat{w}$. From the structure of NTOT algorithm, we know that $\mathcal{Z}_k^{\#}(u^p) = u^p \otimes w^p$, where w^p is the optimal solution to the binary optimization problem in (NTOT_2). Note that $y = Ax + \eta = Ax_S + \eta'$, where $\eta' = Ax_{\overline{S}} + \eta$. By Lemma 4.3.3,

we immediately have

$$\|u^p \otimes w^p - x_S\|_2 \leq \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\|(u^p - x_S) \otimes \hat{w}\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}}\|\eta'\|_2. \tag{4.14}$$

By the definition of u^p in (NTOT_1), we can write

$$\begin{aligned} u^p - x_S &= x^p - x_S + \lambda\left(A^TA + \epsilon I\right)^{-1} A^T\left(y - Ax^p\right) \\ &= \left(I - \lambda\left(A^TA + \epsilon I\right)^{-1} A^TA\right)(x^p - x_S) + \lambda\left(A^TA + \epsilon I\right)^{-1} A^T\eta'. \end{aligned} \tag{4.15}$$

According to the relation (4.15), we can write

$$\begin{aligned} \|(u^p - x_S) \otimes \hat{w}\|_2 =& \left\|(u^p - x_S)_{\text{supp}(\hat{w})}\right\|_2 \\ \leq& \left\|\left((I - \lambda(A^TA + \epsilon I)^{-1}A^TA)(x^p - x_S)\right)_{\text{supp}(\hat{w})}\right\|_2 \\ &+ \left\|\left(\lambda\left(A^TA + \epsilon I\right)^{-1} A^T\eta'\right)_{\text{supp}(\hat{w})}\right\|_2, \end{aligned} \tag{4.16}$$

where the last relation follows from the triangle inequality. Since $\text{card}(\text{supp}(\hat{w}) \cup \text{supp}(x^p - x_S)) \leq \text{card}(\text{supp}(\hat{w}) \cup S \cup S^p) \leq 2k$ due to the fact $S \subseteq \text{supp}(\hat{w})$, applying Lemma 4.3.1 leads to

$$\left\|\left((I - \lambda(A^TA + \epsilon I)^{-1}A^TA)(x^p - x_S)\right)_{\text{supp}(\hat{w})}\right\|_2 \leq \left(\delta_{2k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x_S - x^p\|_2, \tag{4.17}$$

with the choices of (ϵ, λ): $\lambda \leq \epsilon + \sigma_m^2$ and $\epsilon > \sigma_1^2$. By Lemma 4.3.2, the second term of (4.16) on the right-hand side can be bounded as

$$\left\|\left(\lambda\left(A^TA + \epsilon I\right)^{-1} A^T\eta'\right)_{\text{supp}(\hat{w})}\right\|_2 \leq \lambda\left\|\left(A^TA + \epsilon I\right)^{-1} A^T\eta'\right\|_2 \leq \frac{\lambda\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2. \tag{4.18}$$

Combining (4.14) and (4.16)-(4.18) leads to

$$\left\|x^{p+1} - x_S\right\|_2 = \|u^p \otimes w^p - x_S\|_2 \leq \rho\left\|x^p - x_S\right\|_2 + \tau\left\|\eta'\right\|_2, \tag{4.19}$$

where

$$\rho = \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}(\delta_{2k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon+\sigma_1^2}) \text{ and } \tau = \frac{1}{\sqrt{1-\delta_{2k}}}\left(\frac{\lambda\sigma_1\sqrt{1+\delta_k}}{\epsilon+\sigma_1^2} + 2\right).$$

From (4.19), to guarantee the recovery of x_S by the NTOT, it is sufficient to ensure that $\rho < 1$, which is equivalent to

$$\lambda > \epsilon + \sigma_1^2 + \left(\delta_{2k} - \sqrt{\frac{1-\delta_{2k}}{1+\delta_k}}\right)\frac{\epsilon+\sigma_1^2}{\sigma_1^2}.$$

Together with the condition $\lambda \leq \epsilon + \sigma_m^2$, the range of λ is given as

$$\epsilon + \sigma_1^2 + \left(\delta_{2k} - \sqrt{\frac{1-\delta_{2k}}{1+\delta_k}}\right)\frac{\epsilon+\sigma_1^2}{\sigma_1^2} < \lambda \leq \epsilon + \sigma_m^2,$$

which is the relation (4.12) in the theorem. The remaining proof is to show that the range in (4.12) exists. In fact, if the following two conditions are satisfied, the existence of such a range is guaranteed when

$$\delta_{2k} - \sqrt{\frac{1-\delta_{2k}}{1+\delta_k}} < 0 \tag{4.20}$$

and

$$\epsilon + \sigma_1^2 + \left(\delta_{2k} - \sqrt{\frac{1-\delta_{2k}}{1+\delta_k}}\right)\frac{\epsilon+\sigma_1^2}{\sigma_1^2} < \epsilon + \sigma_m^2. \tag{4.21}$$

By noting that $\delta_k \leq \delta_{2k}$, it is straightforward to verify that the inequality (4.20) is guaranteed under the condition $\delta_{2k} < 0.5436$ which is the positive real root of the cubic equation $\gamma^3 + \gamma^2 + \gamma = 1$. The inequality (4.21) can be written as

$$\epsilon > \left(\frac{\sigma_1^2 - \sigma_m^2}{\sqrt{\frac{1-\delta_{2k}}{1+\delta_k}} - \delta_{2k}} - 1\right)\sigma_1^2,$$

which is also guaranteed under the choice of ϵ given in (4.12). Thus the desired result follows.

In particular, if $\eta = 0$ and x is k-sparse, the relation (4.13) is reduced to

$$\left\|x^{p+1} - x\right\|_2 \leq \rho \left\|x^p - x\right\|_2 \leq \rho^{p+1} \left\|x^0 - x\right\|_2,$$

which implies that $\{x^p\}$ converges to x as $p \to \infty$. ■

4.3.2 Analysis of NSROT in noisy scenarios

We denote by $S = \mathcal{L}_k(x)$, the index set of the largest k magnitudes of x. The measurements are given as $y = Ax + \eta$, where η is a noise vector. We first recall some useful technical results which have been shown in [134] and [139]. Lemma 4.3.5 concerns a property of the hard thresholding operator $\mathcal{H}_k$.

Lemma 4.3.5 [134]. Let $z \in \mathbb{R}^n$ be a given vector and $v \in \mathbb{R}^n$ be a k-sparse vector with $\Phi = \text{supp}(v)$. Denote $\Omega = \mathcal{L}_k(z)$. Then one has

$$\left\|v - \mathcal{H}_k(z)\right\|_2 \leq \left\|(v-z)_{\Omega\cup\Phi}\right\|_2 + \left\|(v-z)_{\Omega\setminus\Phi}\right\|_2.$$

Next lemma demonstrates a property of the solution derived from the relaxed optimization problem (NSROT$_2$).

Lemma 4.3.6 [134, 139]. Let $\Lambda \subseteq \{1, \ldots, n\}$ be any given index set, and let $w \in \mathbb{R}^n$ be any given vector satisfying $\mathbf{e}^T w = \gamma k$ and $0 \leq w \leq \mathbf{e}$, where γ is a positive integer number. Decompose the vector w_Λ as

$$w_\Lambda = w_{\Lambda_1} + \cdots + w_{\Lambda_{h-1}} + w_{\Lambda_h},$$

where $h > 0$ is an integer number such that $\text{card}(\Lambda) = (h-1)k + \alpha$ where $0 \leq \alpha < k$, w_{Λ_1} is the first k largest magnitudes in $\{w_i : i \in \Lambda\}$, w_{Λ_2} is the second k largest magnitudes in

$\{w_i : i \in \Lambda\}$, and so on. Then one has

$$\|w_{\Lambda_1}\|_\infty + \cdots + \left\|w_{\Lambda_{h-1}}\right\|_\infty + \|w_{\Lambda_h}\|_\infty \leq \gamma + 1.$$

The next result is actually implied from the proof of Theorem 4.8 in [134].

Lemma 4.3.7 [134]. Let $y = Ax + \eta$ be the measurements of x and $\hat{w} \in \{0,1\}^n$ be a k-sparse binary vector such that $S = \mathcal{L}_k(x) \subseteq \text{supp}(\hat{w})$. Let $S^{p+1} = \text{supp}(x^{p+1})$, u^p and w^p be defined in NTROT. One has

$$\begin{aligned}
&(i)\ \|(x_S - u^p \otimes w^p)_{S\cup S^{p+1}}\|_2 \leq \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\,\|(x_S - u^p)\otimes \hat{w}\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}}\,\|\eta'\|_2 \\
&\qquad\qquad + \frac{1}{\sqrt{1-\delta_{2k}}}\left\|A\left[(x_S - u^p \otimes w^p)_{\overline{S\cup S^{p+1}}}\right]\right\|_2, \\
&(ii)\ \left\|A\left[(x_S - u^p \otimes w^p)_{\overline{S\cup S^{p+1}}}\right]\right\|_2 \leq 2\sqrt{1+\delta_k}\|\mathcal{H}_k(u^p - x_S)\|_2.
\end{aligned}$$

Theorem 4.3.8. *Let $y = Ax + \eta$ be the measurements of $x \in \mathbb{R}^n$ with measurement error η. Let $S = \mathcal{L}_k(x)$ and let σ_1 and σ_m denote, respectively, the largest and smallest singular values of the matrix $A \in \mathbb{R}^{m\times n}$. Suppose that the restricted isometry constant of A satisfies that*

$$\delta_{3k} < 0.2119.$$

Let ϵ be a given parameter satisfying

$$\epsilon > \max\left\{\sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}}+1} - \delta_{3k}} - 1\right)\sigma_1^2\right\}. \tag{4.22}$$

If the stepsize λ in NTROT satisfies

$$\epsilon + \sigma_1^2 + \left(\delta_{3k} - \frac{1}{3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}}+1}\right)\frac{\epsilon+\sigma_1^2}{\sigma_1^2} \ < \ \lambda \ \leq \ \epsilon + \sigma_m^2, \tag{4.23}$$

then the sequence $\{x^p\}$ *generated by the NTROT satisfies that*

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho \left\|x^p - x_S\right\|_2 + \tau \left\|Ax_{\overline{S}} + \eta\right\|_2, \tag{4.24}$$

where

$$\rho = \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\left(\delta_{2k} + 2\delta_{3k} + 3\sigma_1^2 - 3\frac{\lambda\sigma_1^2}{\epsilon+\sigma_1^2}\right) + \delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon+\sigma_1^2} \tag{4.25}$$

and

$$\tau = \frac{1}{\sqrt{1-\delta_{2k}}}\left(\frac{3\lambda\sigma_1\sqrt{1+\delta_k}}{\epsilon+\sigma_1^2} + 2\right) + \frac{\lambda\sigma_1}{\epsilon+\sigma_1^2}. \tag{4.26}$$

In particular, when x *is* k*-sparse and* $\eta = 0$*, then the sequence* $\{x^p\}$ *converges to* x*.*

Proof. Let S, σ_1, σ_m be defined as in the theorem. Note that $y := Ax + \eta = Ax_S + \eta'$ where $\eta' = Ax_{\overline{S}} + \eta$. Denote by $S^{p+1} = \mathrm{supp}(x^{p+1})$. Applying Lemma 4.3.5, we immediately have

$$\begin{aligned}\left\|x_S - x^{p+1}\right\|_2 &= \left\|x_S - \mathcal{H}_k(u^p \otimes \overline{w}^p)\right\|_2 \\ &\leq \left\|(u^p \otimes \overline{w}^p - x_S)_{S^{p+1}\cup S}\right\|_2 + \left\|(u^p \otimes \overline{w}^p - x_S)_{S^{p+1}\setminus S}\right\|_2. \end{aligned}\tag{4.27}$$

In what follows, we bound each of the terms on the right-hand side of the above inequality. By the definition of u^p in NTROT, we have

$$u^p - x_S = \left(I - \lambda\left(A^TA + \epsilon I\right)^{-1}A^TA\right)(x^p - x_S) + \lambda(A^TA + \epsilon I)^{-1}A^T\eta'. \tag{4.28}$$

Noting that $(x_S)_{S^{p+1}\setminus S} = (x_S \otimes w^p)_{S^{p+1}\setminus S} = 0$. Since $0 \leq \overline{w}^p \leq \mathbf{e}$, we have

$$\left\|(u^p \otimes \overline{w}^p - x_S)_{S^{p+1}\setminus S}\right\|_2 = \left\|((u^p - x_S) \otimes \overline{w}^p)_{S^{p+1}\setminus S}\right\|_2 \leq \left\|(u^p - x_S)_{S^{p+1}\setminus S}\right\|_2. \tag{4.29}$$

Substituting (4.28) into (4.29), we can write

$$\begin{aligned}\left\|(u^p \otimes \overline{w}^p - x_S)_{S^{p+1}\backslash S}\right\|_2 \leq & \left\|\left[\left(I - \lambda\left(A^TA + \epsilon I\right)^{-1}A^TA\right)(x^p - x_S)\right]_{S^{p+1}\backslash S}\right\|_2 \\ & + \lambda\left\|\left[(A^TA + \epsilon I)^{-1}A^T\eta'\right]_{S^{p+1}\backslash S}\right\|_2 \\ \leq & \left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x^p - x_S\|_2 + \frac{\lambda\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2, \end{aligned} \tag{4.30}$$

where the last inequality follows from Lemma 4.3.1 (with the fact $\text{card}(\text{supp}(x^p - x_S) \cup (S^{p+1}\backslash S)) \leq 3k$) and Lemma 4.3.2. We now provide an upper bound for the first term of the right-hand side of (4.27). Let $\hat{w}$ be a k-sparse binary vector satisfying $S \subseteq \text{supp}(\hat{w})$. By Lemma 4.3.7, we have

$$\begin{aligned}\|(x_S - u^p \otimes w^p)_{S\cup S^{p+1}}\|_2 \leq & \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\,\|(x_S - u^p)\otimes\hat{w}\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}}\,\|\eta'\|_2 \\ & + \frac{1}{\sqrt{1-\delta_{2k}}}\left\|A\left[(x_S - u^p \otimes w^p)_{\overline{S\cup S^{p+1}}}\right]\right\|_2 \end{aligned} \tag{4.31}$$

and

$$\left\|A\left[(x_S - u^p \otimes w^p)_{\overline{S\cup S^{p+1}}}\right]\right\|_2 \leq 2\sqrt{1+\delta_k}\|\mathcal{H}_k(u^p - x_S)\|_2. \tag{4.32}$$

Applying Lemma 4.3.1 (with $\epsilon > \sigma_1^2$, $\lambda \leq \epsilon + \sigma_m^2$ and $\text{card}(\text{supp}(x^p - x_S) \cup \text{supp}(\hat{w})) \leq 2k$) and (4.3.2), we have

$$\begin{aligned}\|(x_S - u^p)\otimes\hat{w}\|_2 \leq & \left\|\left[(I - \lambda(A^TA + \epsilon I)^{-1}A^TA)(x^p - x_S)\right]\otimes\hat{w}\right\|_2 \\ & + \left\|(\lambda(A^TA + \epsilon I)^{-1}A^T\eta')\otimes\hat{w}\right\|_2 \\ \leq & \left(\delta_{2k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x_S - x^p\|_2 + \frac{\lambda\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2. \end{aligned} \tag{4.33}$$

Let $\Phi = \mathcal{L}_k(x_S - u^p)$. As $\text{card}(\text{supp}(x^p - x_S) \cup \Phi) \leq 3k$, by a proof similar to (4.33), we

also have

$$\|\mathcal{H}_k(x_S - u^p)\|_2 = \|(x_S - u^p)_\Phi\|_2 \leq \left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x^p - x_S\|_2 + \frac{\lambda\sigma_1}{\epsilon + \sigma_1^2}\|\eta'\|_2. \tag{4.34}$$

Combining (4.31)-(4.34), we have

$$\begin{aligned}\|(x_S - u^p \otimes w^p)_{S\cup S^{p+1}}\|_2 \leq \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\left(\delta_{2k} + 2\delta_{3k} + 3\sigma_1^2 - \frac{3\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right)\|x_S - x^p\|_2 \\ + \frac{1}{\sqrt{1-\delta_{2k}}}\left(\frac{3\lambda\sigma_1\sqrt{1+\delta_k}}{\epsilon + \sigma_1^2} + 2\right)\|\eta'\|_2 .\end{aligned} \tag{4.35}$$

Substituting (4.30) and (4.35) into (4.27) leads to

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho\|x^p - x_S\|_2 + \tau\|Ax_{\overline{S}} + \eta\|_2 , \tag{4.36}$$

where ρ and τ are given as (4.25) and (4.26), respectively. Due to the fact $\delta_k \leq \delta_{2k} \leq \delta_{3k}$, we see from (4.36) that

$$\rho \leq \left(3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} + 1\right)\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right).$$

Thus to ensure $\rho < 1$, it is sufficient to require that

$$\left(3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} + 1\right)\left(\delta_{3k} + \sigma_1^2 - \frac{\lambda\sigma_1^2}{\epsilon + \sigma_1^2}\right) < 1,$$

which is equivalent to

$$\lambda > \epsilon + \sigma_1^2 + \left(\delta_{3k} - \frac{1}{3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} + 1}\right)\frac{\epsilon + \sigma_1^2}{\sigma_1^2}.$$

This together with $\lambda \leq \epsilon + \sigma_m^2$ implies that if the range of λ is given as (4.23), then it guarantees that $\rho < 1$. To ensure the existence of the interval in (4.23), the following two

conditions have to be guaranteed:

$$\delta_{3k} - \frac{1}{3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}}+1} < 0, \quad \epsilon > \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{3\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}}+1} - \delta_{3k}} - 1 \right) \sigma_1^2. \tag{4.37}$$

The first condition in (4.37) is ensured by $\delta_{3k} < 0.2119$, and the second condition in (4.37) is ensured by the choice of ϵ given in (4.22). The proof of the theorem is complete.

In particular, if $\eta = 0$ and x is k-sparse, the relation (4.24) is reduced to

$$\left\| x^{p+1} - x \right\|_2 \leq \rho \left\| x^p - x \right\|_2 \leq \rho^p \left\| x^0 - x \right\|_2,$$

which implies that $\{x^p\}$ converges to x as $p \to \infty$. ■

4.3.3 Analysis of NTROTP in noisy scenarios

NTROTP is a combination of NTROT and a pursuit step, thus we can utilize Lemma 3.3.8. The main result is shown below.

Theorem 4.3.9. *Let $y = Ax + \eta$ be the measurements of the signal $x \in \mathbb{R}^n$ with measurement error η. Let $S = \mathcal{L}_k(x)$ and σ_1 and σ_m denote, respectively, the largest and smallest singular values of the matrix $A \in \mathbb{R}^{m \times n}$. Suppose that the restricted isometry constant of A satisfies that*

$$\delta_{3k} < 0.2,$$

and ϵ is a given parameter satisfying

$$\epsilon > \max \left\{ \sigma_1^2, \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}}} - \delta_{3k}} - 1 \right) \sigma_1^2 \right\}. \tag{4.38}$$

If the stepsize λ in NTROTP satisfies

$$\epsilon + \sigma_1^2 + \left(\delta_{3k} - \frac{1}{\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}}} \right) \frac{\epsilon + \sigma_1^2}{\sigma_1^2} \;<\; \lambda \;\leq\; \epsilon + \sigma_m^2, \tag{4.39}$$

then the sequence $\{x^p\}$ generated by the NTROTP satisfies that

$$\left\| x^{p+1} - x_S \right\|_2 \leq \rho^* \left\| x^p - x_S \right\|_2 + \tau^* \left\| A x_{\overline{S}} + \eta \right\|_2 \tag{4.40}$$

where

$$\rho^* := \left(\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}} \right) \left(\delta_{3k} + \sigma_1^2 - \frac{\lambda \sigma_1^2}{\epsilon + \sigma_1^2} \right) \tag{4.41}$$

and

$$\tau^* := \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}} + \frac{1}{(1-\delta_{2k})\sqrt{1+\delta_{2k}}} \left(\frac{3\lambda\sigma_1\sqrt{1+\delta_k}}{\epsilon + \sigma_1^2} + 2 \right) + \frac{\lambda \sigma_1}{(\epsilon + \sigma_1^2)\sqrt{1-(\delta_{2k})^2}}. \tag{4.42}$$

In particular, when x is k-sparse and $\eta = 0$, then the sequence $\{x^p\}$ converges to x.

Proof. NTROTP comprises of NTROT and a pursuit step. Hence we can treat $\mathcal{H}_k(u^p \otimes \overline{w}^p)$ as an intermediate vector generated by NTROT in NTROTP algorithm. From the proof of Theorem 4.3.9, we see that

$$\left\| x_S - \mathcal{H}_k(u^p \otimes \overline{w}^p) \right\|_2 \leq \rho \left\| x^p - x_S \right\|_2 + \tau \left\| \eta' \right\|_2 \tag{4.43}$$

where the constants ρ and τ are given by (4.25) and (4.26) respectively. From the step NTROTP$_4$, x^{p+1} is the solution to the pursuit step. By Lemma 3.3.8, we have

$$\left\| x_S - x^{p+1} \right\|_2 \leq \frac{1}{\sqrt{1-(\delta_{2k})^2}} \left\| x_S - \mathcal{H}_k(u^p \otimes \overline{w}^p) \right\|_2 + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}} \|\eta'\|_2, \tag{4.44}$$

where $\eta' = A x_{\overline{S}} + \eta$. Using $\delta_k \leq \delta_{2k} \leq \delta_{3k}$ and combining (4.43) and (4.44) leads to

$$\left\| x_S - x^{p+1} \right\|_2 \leq \rho^* \left\| x^p - x_S \right\|_2 + \tau^* \left\| \eta' \right\|_2,$$

where ρ^* and τ^* are defined as (4.41) and (4.42) respectively. Note that $\rho^* < 1$ is equivalent to

$$\lambda > \epsilon + \sigma_1^2 + \left(\delta_{3k} - \frac{1}{\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}}} \right) \frac{\epsilon + \sigma_1^2}{\sigma_1^2},$$

which is ensured by the choice of λ given in (4.39). This means the choice of λ in (4.39) ensures that $\rho^* < 1$. To guarantee the existence of the range (4.39), it is sufficient to require that

$$\epsilon + \sigma_1^2 + \left(\delta_{3k} - \frac{1}{\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}}} \right) \frac{\epsilon + \sigma_1^2}{\sigma_1^2} < \epsilon + \sigma_m^2,$$

which is equivalent to

$$\delta_{3k} - \frac{1}{\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}}} < 0, \quad \epsilon > \left(\frac{\sigma_1^2 - \sigma_m^2}{\frac{1}{\frac{3}{1-\delta_{3k}} + \frac{1}{\sqrt{1-(\delta_{3k})^2}}} - \delta_{3k}} - 1 \right) \sigma_1^2. \tag{4.45}$$

Note that $\dfrac{1}{\sqrt{1-(\delta_{3k})^2}} < \dfrac{1}{1-\delta_{3k}}$ since $0 < \delta_{3k} \leq 1$. The first condition in (4.45) is satisfied when $\delta_{3k} < 0.2$. The second condition (4.45) is also satisfied provided ϵ is chosen large enough, i.e., satisfying (4.38).

In particular, if $\eta = 0$ and x is k-sparse, the relation (4.40) is reduced to

$$\left\| x^{p+1} - x \right\|_2 \leq \rho^* \left\| x^p - x \right\|_2 \leq (\rho^*)^{p+1} \left\| x^0 - x \right\|_2,$$

which implies that $\{x^p\}$ converges to x as $p \to \infty$. ■

Before going to the numerical section, we discuss the improved theoretical analysis of the proposed algorithms, the choice of algorithmic parameters (ϵ, λ), and the complexity of the main algorithms in this chapter.

Remark 4.3.1

The bound for δ_{3k} in Theorems 4.3.8 and 4.3.9 can be replaced by the one for δ_{2k} either through a more subtle analysis or through the relation $\delta_{ck} \leq c\delta_{2k}$, where c is a positive integer number. Such a relation bridging RICs with different orders is shown in Proposition

6.6 in [63]. In fact, from $\delta_{3k} < 3\delta_{2k}$, it is evident that $\delta_{2k} < 0.0706$ implies $\delta_{3k} < 0.2119$, and that $\delta_{2k} < 0.0666$ implies $\delta_{3k} < 0.2$. Therefore, we may use the bound $\delta_{2k} < 0.0706$ in Theorem 4.3.8 and $\delta_{2k} < 0.0666$ in Theorem 4.3.9 without damaging the results in these theorems.

Remark 4.3.2

Similar to Newton-step-based hard thresholding, we only need to compute the matrix $\left(A^TA + \epsilon I\right)^{-1}$ once in Newton-type optimal thresholding algorithms. This inverse can be obtained by using singular value decomposition (SVD) of matrix A (see Remark 3.3.1 for detail), which is also used for the choice of algorithmic parameter (ϵ, λ). As pointed out in Remark 3.3.2, it is difficult to determine the choice of (ϵ, λ) directly according to the condition (4.11) and (4.12) in Theorem 4.3.4, condition (4.22) and (4.23) in Theorem 4.3.8 and condition (4.38) and (4.39) in Theorem 4.3.9. However, we use a simple rule implied from these conditions, that is $\epsilon > \sigma_1^2$ and $\lambda \leq \epsilon + \sigma_m^2$, to choose (ϵ, λ) in numerical experiment or practical application scenarios.

Remark 4.3.3

The computation of matrix A^TA and the LU decomposition of $A^TA + \epsilon I$ require $O(mn^2)$ and $O(n^3)$, respectively [84]. Using the interior-point method in [113] to solve the relaxation optimization problem requires $O(n^{3.5}L)$ [139], where L is the size of the problem data encoding in the binary space. The hard thresholding operator $\mathcal{H}_k$ requires about $O(n \log k)$, where n and k are the vector length and the sparsity level [139]. Therefore, each iteration of NTROT requires about $O(mn^2 + n^3 + n^{3.5}L)$ flops. The pursuit step involves $O(k^3)$ basic operations [84] by solving the linear system $[A_\Omega]^T y = [A_\Omega]^T [A_\Omega] z$[1], or requires about $O\left(\frac{4}{3}m^3\right)$ flops [139] by using Cholesky decomposition of the matrix $[A_\Omega]^T[A_\Omega]$, where $\Omega = \text{supp}(u^p \otimes \overline{w}^p)$. Thus each iteration of NTROTP involves either $O(mn^2+n^3+n^{3.5}L+k^3)$ or $O(mn^2 + n^3 + n^{3.5}L + m^3)$ flops depending on which strategy is applied to the pursuit step.

[1] $[A_\Omega] \in \mathbb{R}^{m\times k}$ denotes the submatrix of A which retains the columns indexed by Ω and deletes others.

4.4 Numerical Performance

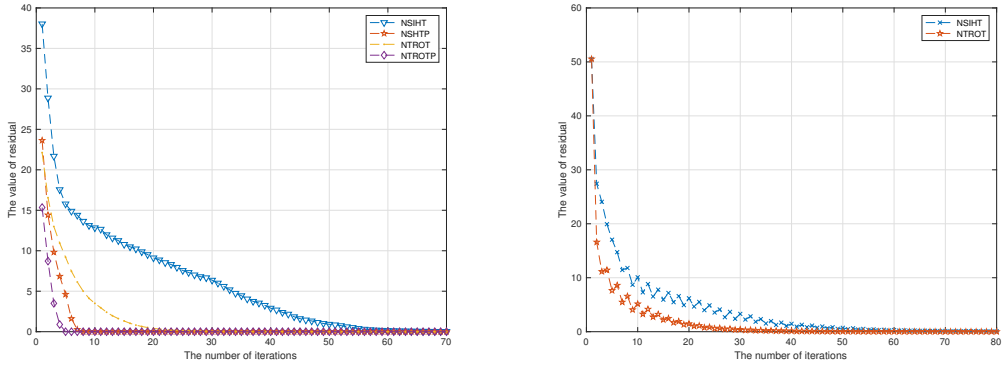
Simulations were performed to test the performance of the proposed algorithms with respect to residual reduction, the average number of iterations, success frequency for signal recovery and phase transition in terms of success rates. Without a specified statement, the measurement matrices generated for experiments are still Gaussian random matrices, whose entries are independent and identically distributed and follow the standard normal distribution $\mathcal{N}(0,1)$. Nonzero entries of realized sparse signals also follow the $\mathcal{N}(0,1)$, and their position follows a uniform distribution. The algorithms considered in this section are ℓ_1-minimization [38, 31], orthogonal matching pursuit (OMP) [112, 90], subspace pursuit (SP), Newton-step-based hard thresholding [84], quasi-Newton iterative projection (QNIP) [73] and quasi-Newton projection pursuit (QNPP) [117]. The optimization problems in our algorithms were solved by the CVX which is developed by Grant and Boyd [66] with solver 'Mosek' [3].

4.4.1 Residual reduction

The experiment was carried out to compare the residual-reduction performance of the algorithms with given (ϵ, λ). In this experiment, we set $A \in \mathbb{R}^{256\times 512}, y = Ax^*$, $k = 70$ and $x^0 = 0$. The stepsize λ and parameter ϵ are set, respectively, as

$$\lambda = 5, \epsilon = \max\left\{\sigma_1^2 + 1, \lambda - \sigma_m^2\right\},$$

which guarantees that $\epsilon > \sigma_1^2$ and $\lambda \leq \epsilon + \sigma_m^2$, where σ_1 and σ_m denote the largest and the smallest singular value of the generated matrix. Fig. 4.1 demonstrates the change of the residual value, i.e., $\|y - Ax\|_2$, in the course of iterations of the algorithms. From Fig. 4.1 (a), it can be seen that the NTROTP is more powerful than other algorithms in residual reduction. In the same experiment environment, we also compare the residual change in the course of NSIHT and NTROT which use different thresholding operators. Fig. 4.1 (b)



(a) (b)

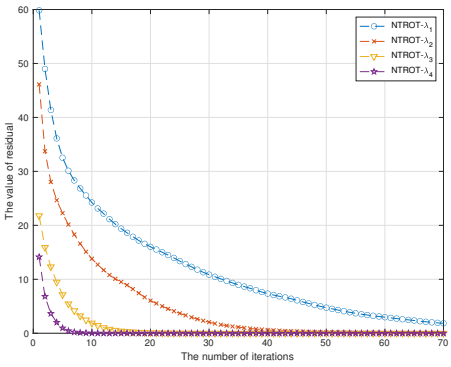
Figure 4.1: (a) Comparison of residual-reduction performances of several algorithms; (b) Residual change in the course of iterations using different thresholding operators. Experimental settings: $A \in \mathbb{R}^{256\times 512}, y = Ax^*$ and $k = 70$.

shows that the algorithm with optimal thresholding operator can reduce the residual more efficiently than the one with hard thresholding operator.

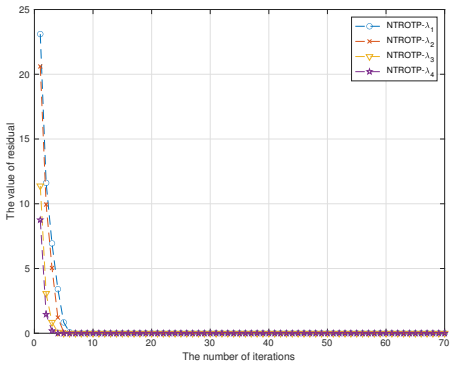
The performance of NTROT and NTROTP are related to the choice of (ϵ, λ). Thus we test the residual-reduction performance of the proposed algorithms in terms of different values of parameter ϵ and stepsize λ. The results are shown in Fig. 4.2 and Fig. 4.3, respectively. In Fig. 4.2, the stepsize λ is fixed as $\lambda = 10$, and $\epsilon = \epsilon^*, 1.1\epsilon^*, 1.5\epsilon^*$ and $2\epsilon^*$, where $\epsilon^* = \sigma_1^2 + 1$. In Fig. 4.3, the parameter ϵ is fixed as $\epsilon = \sigma_1^2 + 1$, and stepsize λ is taken as $\lambda = 1, 2, 5, 10$ respectively. Such choices of (ϵ, λ) satisfy that $\epsilon > \sigma_1^2$ and $\lambda \leq \epsilon + \sigma_m^2$. It can be seen that the NTROT is more sensitive to the change of ϵ and λ than the NTROTP which is generally insensitive to the change of (ϵ, λ). This also indicates that NTROTP is a stable algorithm for reducing the value of objective value of problem (4.1).

4.4.2 The number of iterations

The simulations were also performed to examine the impact of sparsity levels and measurement levels on the average number of iterations needed for signal recovery and objective minimization via Newton-type iterative algorithms. The stopping criterion for signal recov-

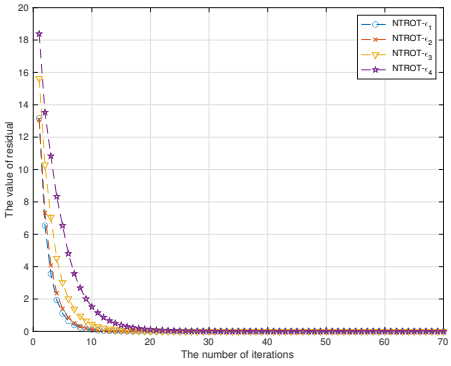


(a) NTROT with different λ

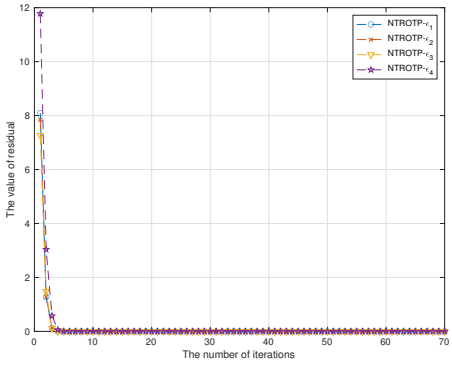


(b) NTROTP with different λ

Figure 4.2: Residual reduction by NTROT and NTROTP with $\epsilon = \sigma_1^2 + 1$ and different stepsize $\lambda = 1, 2, 5, 10$. Experimental settings: $A \in \mathbb{R}^{256\times512}, y = Ax^*$ and $k = 70$.

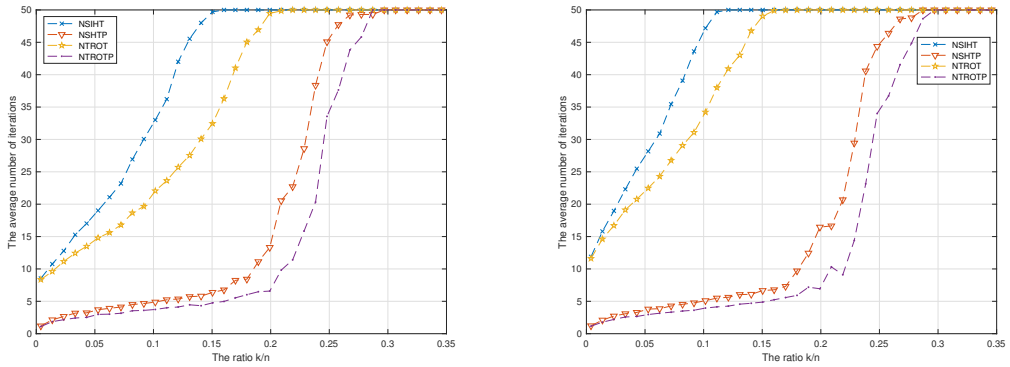


(a) NTROT with different ϵ



(b) NTROTP with different ϵ

Figure 4.3: Residual reduction by NTROT and NTROTP with $\lambda = 10$ and different parameter ϵ: $\epsilon_1 = \epsilon^*$, $\epsilon_2 = 1.1\epsilon^*$, $\epsilon_3 = 1.5\epsilon^*$, $\epsilon_4 = 2\epsilon^*$. Experimental settings: $A \in \mathbb{R}^{256\times512}, y = Ax^*$ and $k = 70$.



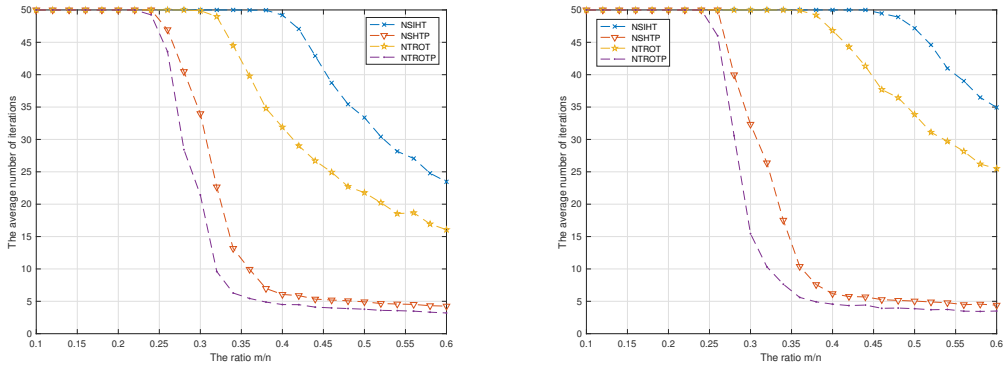
(a) Sparse signal recovery (b) Objective function minimization

Figure 4.4: Comparison of the average number of iterations required by different algorithms in terms of different sparsity levels. Experimental settings: the size of A is 256×512, $y = Ax^*$, and the ratio k/n varies from 0.01 to 0.35.

ery is $\|x^* - x\|_2/\|x^*\|_2 \leq 10^{-3}$, while the one for objective minimization is $\|y - Ax\|_2 \leq 10^{-3}$. All algorithms start from $x^0 = 0$ and terminate either when the corresponding stopping criterion is met or when the maximum number of iterations (i.e., 50 iterations) is reached. The results with respect to the change of sparsity levels and number of measurements are shown in Fig. 4.4 and Fig. 4.5, respectively.

Fig. 4.4 demonstrates the influence of sparsity levels on the number of iterations needed by NSIHT, NSHTP, NTROT and NTROTP. In this experiment, the size of measurement matrices is still 256×512, and the ratio k/n varies from 0.01 to 0.35. The average number of iterations is calculated based on 50 random examples for each sparsity level k/n. A common feature of these algorithms is that with the increase of the sparsity levels, the required iterations to reconstruct signals or minimize the objective function also increase. We also observe that both the optimal thresholding and pursuit step help reduce the required number of iterations of algorithms to reconstruct a signal. Furthermore, compared with NSIHT and NSHTP, the required iterations of NTROT and NTROTP will not change too much dealing with signal reconstruction or solving the problem.

Fig. 4.5 compares the average number of iterations required by several algorithms applying to different measurement levels. The target signal is fixed as $x^* \in \mathbb{R}^{500}$ with $k = 50$, and the length of the observed vector $y = Ax^*$, i.e., the number of measurements,



(a) Sparse signal recovery

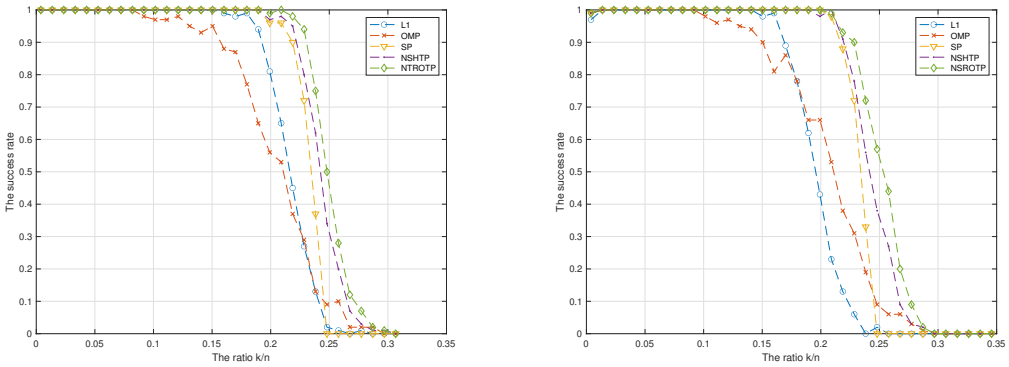
(b) Objective function minimization

Figure 4.5: Comparison of the average number of iterations required by different algorithms in terms of different number of measurements. Experimental settings: $n = 500$, $y = Ax^*$, and the number of measurements m varies from 50 to 300.

varies from 50 to 300. At each different number of measurements, 50 random trials are generated. When $m/n < 0.25$, we see that no algorithm could achieve 'success' for both signal recovery and objective minimization due to the fact that the measurement levels are too low. A general trend shown in Fig. 4.5 is that the more measurements obtained for the target signal x^*, the less number of iterations needed for meeting the stopping criterion. We see that NSHTP and NTROTP could recover the signal or solve the sparse optimization problem by using a relatively small number of iterations when the ratio $m/n \geq 0.35$, and the NTROTP needs fewer iterations than NSIHT, NSHTP, and NTROT.

4.4.3 The performance of signal recovery

Simulations were carried out to compare the signal reconstruction performance of the NTROTP algorithm and several existing methods with both exact and inexact measurements. In this experiment, the size of matrices is still 256×512. All iterative algorithms start at $x^0 = 0$. The ratio k/n increases from 0.01 to 0.35. We first compare our algorithm with several heuristic algorithms including ℓ_1-minimization, OMP, SP, NSHTP and NTROTP. The results were given in Fig. 4.6. Then we compare our algorithm with two existing Newton-Type methods including QNIP and QNPP, for which the results



(a) Under exact measurements (b) Under inexact measurements

Figure 4.6: Comparison of success frequencies of signal recovery via different algorithms

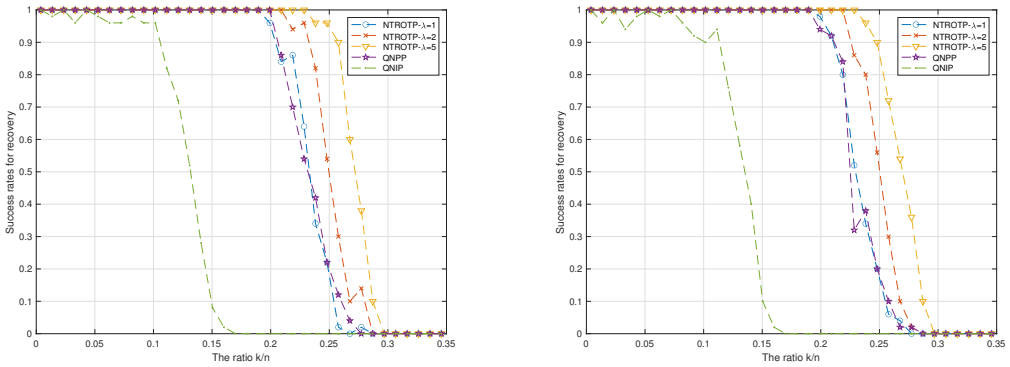
were summarized in Fig. 4.7. The vertical axes in Fig. 4.6 and Fig. 4.7 represent the reconstruction success rates which is calculated based on 50 random examples.

First experiment

In the experiments producing the results in Fig. 4.6, the measurements of x^* are set as $y = Ax^* + 0.001\eta$, where $\eta \in \mathbb{R}^{256}$ is a standard Gaussian random vector. The iterative algorithms terminate after 20 iterations except for the OMP which stops after k iterations owing to the its structure, where $k = \|x^*\|_0$. The choice of (ϵ, λ) is the same as (4.36). The condition $\|x^p - x^*\|_2 / \|x^*\|_2 \leq 10^{-3}$ is set as the recovery criterion. Fig. 4.6 indicate that the NTROTP is stable and robust for sparse signal recovery compared with other algorithms used in these experiments.

Second experiment

The entries of Gaussian random matrices used in the experiment that generates Fig. 4.7 follow the normal distribution $\mathcal{N}(0, 1/m)$, where m denotes the number of measurements, i.e., $m = 256$. In such an experiment, the noisy measurements of x^* are set as $y = Ax^* + 10^{-5}\eta$, where $\eta \in \mathbb{R}^{256}$ is a standard Gaussian random vector; all algorithms



(a) Under exact measurements

(b) Under inexact measurements

Figure 4.7: Comparison of success frequencies of signal recovery via different Newton-type algorithms

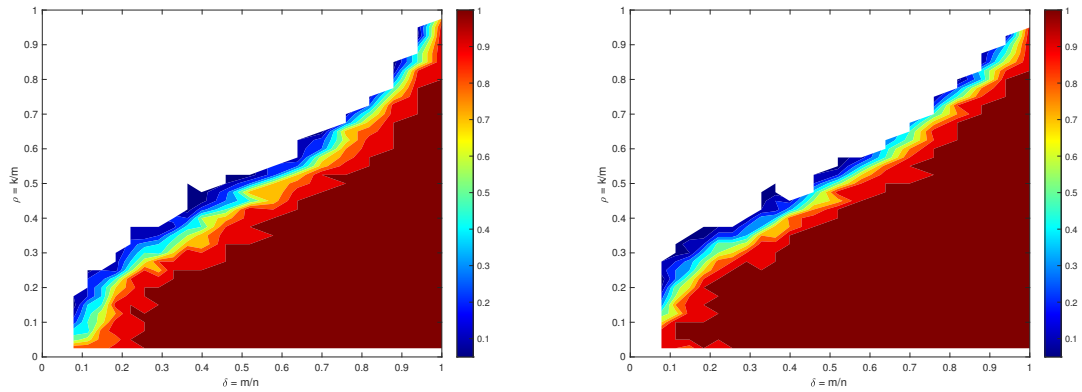
terminate when either the criterion

$$\|x^p - x^*\|_2 / \|x^*\|_2 \leq 10^{-3}$$

is met or a total of 50 iterations are performed for QNIP and NTROTP and 300 iterations are performed for QNIP (which works slowly and thus we allow this algorithm to perform many more iterations than the other two). For the NTROTP algorithm, we set $\epsilon = 10^{-5}$ to slightly perturb the singular Hessian $A^T A$ such that $A^T A + \epsilon I$ is positive definite. The parameter λ is set as $\lambda = 1, 2$ and 5, respectively. As indicated by Fig. 4.7, QNPP and NTROTP algorithms are comparable, and they outperform QNIP even if we allow the QNIP to perform many more iterations than the other two algorithms. Fig. 4.7 also indicates that the NTROTP with a small value of ϵ is also efficient and robust for signal recovery, although this has not been shown rigorously from a theoretical viewpoint.

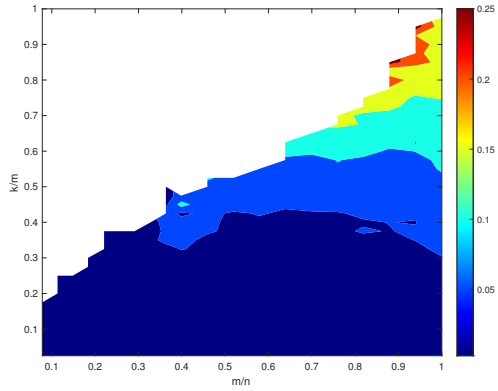
4.4.4 Phase transition

We also performed experiments to compare the phase transition phenomena of NSHTP and NTROTP algorithms. The vector length n is fixed as 512. A total of 40 sparsity levels are taken uniformly from 1 to m, i.e., $k = \left\lfloor \frac{1}{40}m \right\rfloor, \ldots, \left\lfloor \frac{39}{40}m \right\rfloor, m$. The minimal number

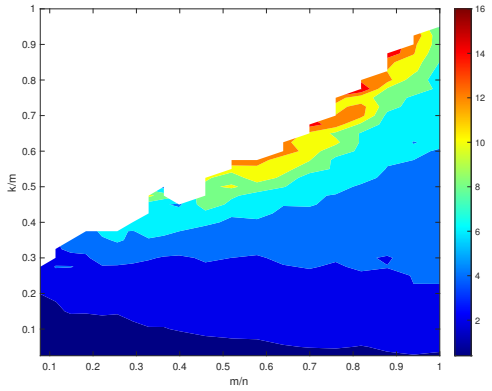


(a) NSHTP-success rate

(b) NTROTP-success rate



(c) NSHTP-time consumption

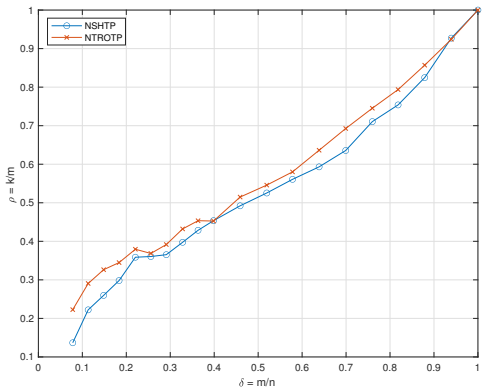


(d) NTROTP-time consumption

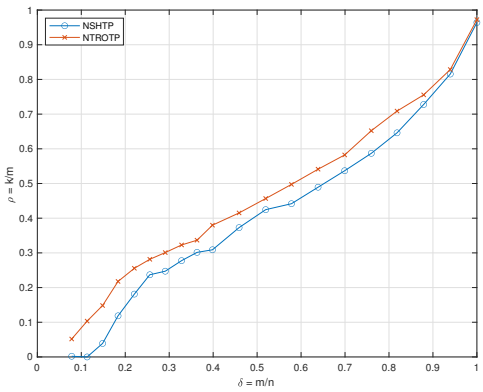
Figure 4.8: Phase transition of NSHTP and NTROTP with respect to success rate and time consumption under inexact measurements

of measurements is 40 which guarantees that there exist at least 40 different sparsity levels. We take 10 values of m uniformly when $m \in [40, 0.4n]$ and $m \in (0.4n, n]$, respectively. The measurements are inexact, i.e., $y = Ax + 0.001\eta$, where the entries of matrix A and vector η follow the standard normal distribution $\mathcal{N}(0, 1)$. The choice of algorithmic parameters (ϵ, λ) of these two algorithms is $\lambda = 5, \epsilon = \max\left\{\sigma_1^2 + 1, \lambda - \sigma_m^2\right\}$. The iterative algorithms terminate either when the criterion $\|x^p - x^*\|_2 / \|x^*\|_2 \leq 10^{-3}$ is met or a maximum of 30 iterations are reached. At each pair of (m, k), the success rate is calculated based on 20 trials. The vertical axis represents the over-sampling rate $\rho = k/m$ and the horizontal axis is the under-sampling rate $\delta = m/n$. Clearly, the range of ρ and δ is $(0, 1]$.

Fig. 4.8 shows the success rate change and time consumption for recovery success of



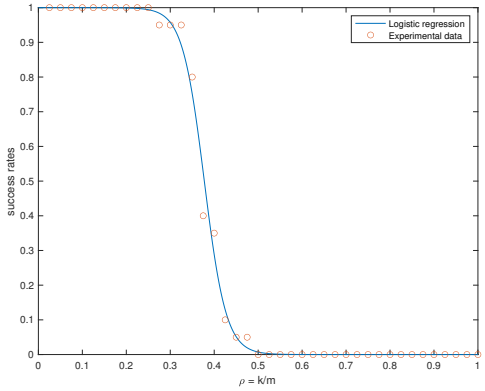
(a) success rate = 0.1



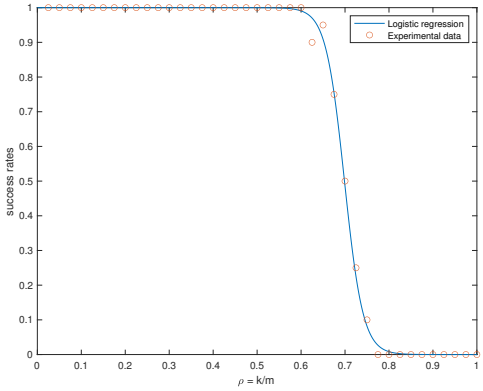
(b) success rate = 0.9

Figure 4.9: Phase transition curve of NSHTP and NTROTP under inexact measurements in terms of success rates

NSHTP and NTROTP over the phase transition region $\delta \times \rho = [0, 1]^2$. The shaded area indicates the algorithm can recover the signal with a certain level of success rate. We can see from Fig. 4.8 (a) and (b) that both NSHTP and NTROTP are efficient for sparse signal recovery. Fig. 4.8 (c) and (d) indicate that NTROTP requires more time than NSHTP due to solving the convex relaxation optimization problem. However, comparing their phase transition curves, we found it is worthwhile to apply the Newton-type optimal thresholding methods.



(a) $\delta = 0.3281$



(b) $\delta = 0.7598$

Figure 4.10: Examples of using logistic model to estimate the success rate for NTROTP at a certain value of δ

Based on the data obtained from Fig. 4.8, for every given δ, we can compute the logistic regression model to determine the specific success rate corresponding to ρ. The phase transition curve for different success rates can be derived, where 4.9 (a) and (b) show the results for success rates 0.1 and 0.9, respectively. In particular, two examples in Fig. 4.10 show that logistic regression models can well estimate the success rate corresponding to ρ. The higher the phase transition curve, the better signal recovery capability. In both Fig. 4.9 (a) and (b), it can be seen that NTROTP performs better than NSHTP.

4.5 Conclusion

A class of Newton-type optimal k-thresholding algorithms is proposed in this chapter. Under the restricted isometry property (RIP), we have proved that the NTOT, NTROT and NTROTP algorithms are guaranteed to reconstruct sparse signals with a proper choice of the algorithmic parameter and iterative stepsize. Simulations indicate that the proposed algorithms, especially the NTROTP, are stable, robust, and comparable to some popular compressed sensing algorithms for signal reconstruction.

CHAPTER 5

PARTIAL GRADIENT OPTIMAL K-THRESHOLDING ALGORITHMS

5.1 Introduction

A typical type of thresholding algorithms, including the traditional hard thresholding methods and the algorithms proposed in previous chapters, adopts the full steepest descent direction or a Newton-like direction as a search direction to generate an iterate on which a certain thresholding is performed. Traditional hard thresholding methods discard a large part of a vector when the vector is dense, thus a large amount of important information has been lost in such a thresholding process. Recent study [134] shows that the hard thresholding should be applied to a compressible vector instead of a dense vector to avoid a big loss of information. On the other hand, the optimal k-thresholding as a novel thresholding technique may overcome the intrinsic drawback of the hard thresholding, and performs thresholding and objective function minimization simultaneously. This motivates us to propose the so-called partial gradient optimal thresholding method studied in this chapter[1], which is an integration of the partial gradient method and the optimal k-thresholding technique.

Consider the sparse optimization problem (3.4) as pointed out in Chapter 3. The key

[1]The main result in this chapter has been published (jointly with Yun-Bin Zhao, Michal Kočvara and Zhongfeng Sun) has been published on the Journal of Global Optimization, pp. 1-21, 2022 [86]

step of thresholding-type methods for solving (3.4) can be formulated as

$$x^{p+1} = \mathcal{T}_k \left(x^p + \lambda \mathbf{d} \right), \tag{5.1}$$

where $\mathcal{T}_k(\cdot)$ represents a thresholding operator to produce k-sparse vectors, λ is the stepsize and $\mathbf{d}$ represents a search direction. The formula (5.1) can be reduced to the well-known iterative hard thresholding (IHT) method by setting $\mathcal{T}_k = \mathcal{H}_k$ and $\mathbf{d} = -\nabla f(x)$, i.e.,

$$x^{p+1} = \mathcal{H}_k \left(x^p - \lambda \nabla f \left(x^p \right) \right),$$

where $\nabla f(x^p)$ denotes the gradient of f at x^p. The vector $x^p - \lambda \nabla f(x^p)$ is generally not sparse since the direction $-\nabla f\left(x^p\right)$ is not necessarily sparse. To meet the sparsity constraint in (3.4), the hard thresholding operator $\mathcal{H}_k$ reduces the number of nonzero entries to k from n, which causes a sudden change of the sparsity level and usually increases the objective value. Inspired by the fact that the hard thresholding operator should be imposed on a compressible vector to avoid such a significant loss of information [134, 139], we adopt the partial gradient instead of the full gradient as a search direction to generate the following sparse or compressible vector:

$$u^p := x^p - \lambda \mathcal{H}_q \left(\nabla f \left(x^p \right) \right) \tag{5.2}$$

on which a thresholding operator is then performed to generate the next sparse iterate x^{p+1}. The integer number $q > 0$ is the number of elements extracted from the full gradient in the above formula. In other words, we only use q significant entries of the gradient as our search direction. Clearly, the vector generated by (5.2) may be not k-sparse, thus we may use the hard thresholding of u^p to produce an iterate satisfying the constraint of the problem (3.4). The optimal k-thresholding is more powerful and more efficient than the

hard thresholding. This suggests the following iterative scheme:

$$x^{p+1} = \mathcal{Z}_k^{\#}\left(x^p - \lambda \mathcal{H}_q\left(\nabla f\left(x^p\right)\right)\right),$$

where $\mathcal{Z}_k^{\#}$ denotes the optimal k-thresholding operator, $\lambda > 0$ denotes a stepsize, and $q > 0$ is an integer.

The main contributions of this chapter can be summarized as follows.

i. We introduce a class of partial gradient optimal thresholding algorithms, including partial gradient optimal k-thresholding (PGOT), relaxed partial gradient optimal k-thresholding (RPGOT) and relaxed partial gradient optimal k-thresholding pursuit (PGROTP), described as Algorithms 5.1, 5.2 and 5.3, respectively. In particular, the PGROTP is considered the recommended version of the proposed algorithm used for sparse signal recovery.

ii. Under the restricted isometry property (RIP) assumption, we have shown the sufficient conditions and convergence of our algorithms with q in the range $q \in [2k, n]$. The main results for the proposed algorithms are summarized in Theorems 5.3.3, 5.3.5 and 5.3.7, respectively.

iii. We have conducted a series of numerical experiments to investigate the performance of the proposed algorithms and make comparisons between them and some other existing algorithms. Simulations from synthetic data indicate that PGROTP algorithm is robust and comparable to several existing methods for sparse signal recovery. Empirical results are given in Section 5.4.

5.2 Algorithms

We still focus on the specific problem

$$\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2}\|y - Ax\|_2^2 : \|x\|_0 \leq k \right\}, \tag{5.3}$$

where $A \in \mathbb{R}^{m\times n}$ with $m \ll n$ is the measurement matrix and $y \in \mathbb{R}^m$ denote the linear measurements. The gradient of the objective function in (5.3) is given as

$$\nabla f(x) = -A^T(y - Ax).$$

By using the major part of this specific gradient, we define u^p as follows:

$$u^p = x^p + \lambda \mathcal{H}_q \left(A^T \left(y - Ax^p \right) \right),$$

where $0 < q \leq n$ is an integer number. The optimal k-thresholding of u^p with respect to the objective function in (5.3) is given by $\mathcal{Z}_k^{\#}\left(u^p\right) = u^p \otimes w^*$, where

$$w^* = \arg\min_w \left\{ \|y - A(u \otimes w)\|_2^2 : \mathbf{e}^T w = k, w \in \{0,1\}^n \right\}.$$

Thus, the partial gradient optimal k-thresholding algorithm for solving (5.3) can be stated as

$$x^{p+1} = \mathcal{Z}_k^{\#} \left(x^p + \lambda \mathcal{H}_q \left(A^T \left(y - Ax^p \right) \right) \right).$$

By the definition of the operator $\mathcal{Z}_k^{\#}$, the so-called partial gradient optimal k-thresholding algorithm can be described explicitly as Algorithm 5.1. For algorithmic description and analysis simplicity, we set $\lambda = 1$ throughout the rest of the chapter. It should be stressed that in practical applications, suitable stepsize should be used in order to speed up the convergence of the algorithms. Moreover, q needs to satisfy $q \geq k$ to guarantee enough partial direction. Rather than working on the binary problem directly, one can solve its relaxation leading to the so-called partial gradient relaxed optimal thresholding (PGROT) algorithm, given in Algorithm 5.2. The PGROT can be further enhanced by including a pursuit step (i.e., an orthogonal projection step) to find a possibly better iterate than the point generated by PGROT. This consideration leads to the so-called PGROTP algorithm which is described in Algorithm 5.3.

In the PGROT algorithm, the solution $\overline{w}^p$ to (PGROT$_2$) is not k-sparse in general, thus

Algorithm 5.1 Partial Gradient Optimal k-Thresholding (PGOT)

- Input: matrix A, vector y, sparsity level k, integer number $q \geq k$, and initial point $x^0 = 0$.
- Iteration:

$$u^p = x^p + \mathcal{H}_q\left(A^T(y - Ax^p)\right) \tag{PGOT$_1$}$$
$$w^p = \arg\min_w \{\|y - A(w \otimes u^p)\|_2^2 : \ \mathbf{e}^T w = k, \ w \in \{0,1\}^n\} \tag{PGOT$_2$}$$
$$x^{p+1} = w^p \otimes u^p \tag{PGOT$_3$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Algorithm 5.2 Partial Gradient Relaxed Optimal k-Thresholding (PGROT)

- Input: matrix A, vector y, sparsity level k, integer number $q \geq k$, and initial point $x^0 = 0$.
- Iteration:

$$u^p = x^p + \mathcal{H}_q\left(A^T(y - Ax^p)\right) \tag{PGROT$_1$}$$
$$\overline{w}^p = \arg\min_w \{\|y - A(w \otimes u^p)\|_2^2 : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\} \tag{PGROT$_2$}$$
$$x^{p+1} = \mathcal{H}_k(\overline{w}^p \otimes u^p) \tag{PGROT$_3$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Algorithm 5.3 Partial Gradient Relaxed Optimal k-thresholding Pursuit (PGROTP)

- Input: matrix A, vector y, sparsity level k, integer number $q \geq k$, and initial point $x^0 = 0$.
- Iteration:

$$u^p = x^p + \mathcal{H}_q\left(A^T(y - Ax^p)\right) \tag{PGROTP$_1$}$$
$$\overline{w}^p = \arg\min_w \{\|y - A(w \otimes u^p)\|_2^2 : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\} \tag{PGROTP$_2$}$$
$$S^{p+1} = \mathcal{L}_k(\overline{w}^p \otimes u^p) \tag{PGROTP$_3$}$$
$$x^{p+1} = \arg\min_z \{\|y - Az\|_2 : \ \ \mathrm{supp}(z) \subseteq S^{p+1}\} \tag{PGROTP$_4$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

---

the final step in PGROT is to restore the sparsity of an iterate. It should be noted that the use of hard thresholding operator here is different from the one in the traditional IHT algorithm, as a result of the vector u^p generated by the partial gradient is $(k+q)$-sparse instead of being a usually dense vector in IHT. Moreover, compared with the optimal k-thresholding adopting a full descent direction, the partial gradient is beneficial for reducing the complexity of optimization problems, e.g., the problem (PGOT$_2$) and (PGROT$_2$), in the sense that the search region has been significantly reduced.

5.2.1 Extension to general problem

We can also extend the above algorithms for solving the problem (3.4). But for conciseness, we only state the generalized version for PGROTP, given in Algorithm 5.4.

Algorithm 5.4 Generalized Partial Gradient Relaxed Optimal k-thresholding Pursuit (GRPGOTP)

- Input: matrix A, vector y, sparsity level k, integer number $q \geq k$, and initial point $x^0 = 0$.
- Iteration:

$$\begin{aligned} u^p &= x^p - \mathcal{H}_q\left(\nabla f(x^p)\right) \\ w^p &= \arg\min_w \{f(w \otimes u^p) : \ \mathbf{e}^T w = k, \ 0 \leq w \leq \mathbf{e}\} \\ S^{p+1} &= \mathcal{L}_k(w^p \otimes u^p) \\ x^{p+1} &= \arg\min_z \{f(z) : \ \ \mathrm{supp}(z) \subseteq S^{p+1}\} \end{aligned}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

5.3 Theoretical Analysis

In this section, we perform a theoretical analysis including convergence of the proposed algorithms in terms of the restricted isometry property (RIP). Our analysis is carried out for the case of inexact measurements, i.e., $y = Ax + \eta$.

5.3.1 PGOT in noisy scenarios

Recall the following technical results concerning the properties of RIP matrix, and hard thresholding operator $\mathcal{H}_k$, respectively.

Lemma 5.3.1 [31, 89, 60]. Given a vector $u \in \mathbb{R}^n$ and a set $\Omega \subseteq [N]$, one has

$$\begin{aligned} &\left\| \left(\left(I - A^T A \right) v \right)_\Omega \right\|_2 \leq \delta_t \|u\|_2 \qquad \text{if} \quad \text{card}(\Omega \cup \text{supp}(v)) \leq t, \\ &\left\| (A^T u)_\Omega \right\|_2 \leq \sqrt{1+\delta_t} \|u\|_2 \qquad \text{if} \quad \text{card}(\Omega) \leq t. \end{aligned}$$

Lemma 5.3.2 [142]. For any integer number $\alpha \geq \beta$, any vector $z \in \mathbb{R}^n$ and for any β-sparse vector $u \in \mathbb{R}^n$ (i.e., $\|u\|_0 \leq \beta$), one has

$$\|u - \mathcal{H}_\alpha(z)\|_2 \leq \frac{\sqrt{5}+1}{2} \|(u-z)_{\Lambda \cup \Omega}\|_2,$$

where $\Lambda = \text{supp}(u)$ and $\Omega = \text{supp}\left(\mathcal{H}_\alpha(z)\right)$.

When $\beta \leq \alpha$, a β-sparse vector is also α-sparse. Thus Lemma 5.3.2 above follows exactly from Lemma 2.2 in [142]. The notation $\lceil \cdot \rceil$ throughout this chapter denotes the ceiling function. We are ready to perform the theoretical analysis of the algorithm PGOT. The main result of PGOT is summarized as follows.

Theorem 5.3.3. *Let $y = Ax + \eta$ be the measurements of the signal $x \in \mathbb{R}$ with measurement error η, and $q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of A satisfies*

$$\delta_{3k} < \alpha^*, \tag{5.4}$$

where $\alpha^* \in (0,1)$ is the unique real root of the univariate equation

$$\left(\left\lceil \frac{q}{k} \right\rceil + 1 \right)^2 \alpha^3 + \left(\left\lceil \frac{q}{k} \right\rceil + 1 \right)^2 \alpha^2 + \frac{2}{3+\sqrt{5}} \alpha - \frac{2}{3+\sqrt{5}} = 0. \tag{5.5}$$

Then the sequence $\{x^p\}$ *generated by PGOT satisfies that*

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho\|x^p - x_S\|_2 + \tau\|Ax_{\overline{S}} + \eta\|_2, \tag{5.6}$$

where

$$\rho := \frac{\sqrt{5}+1}{2}\left(\left\lceil\frac{q}{k}\right\rceil + 1\right)\delta_{3k}\sqrt{\frac{1+\delta_k}{1-\delta_{2k}}} < 1 \tag{5.7}$$

is guaranteed under the condition (5.4)*, and the constant* τ *is given as*

$$\tau = \frac{(\sqrt{5}+1)\left(\left\lceil\frac{q}{k}\right\rceil + 1\right)(1+\delta_{2k}) + 4}{2\sqrt{1-\delta_{2k}}}. \tag{5.8}$$

In particular, if $\eta = 0$*, the sequence* $\{x^p\}$ *generated by PGOT converges to* x*.*

Proof. Denote $y = Ax + \eta = Ax_S + \eta'$, where $S := \mathcal{L}_k(x)$ is the support set of the largest k magnitudes of x and $\eta' = Ax_{\overline{S}} + \eta$. Let $\hat{w} \in \{0,1\}^n$ be a k-sparse vector, such that $S \subseteq \mathrm{supp}(\hat{w})$. Applying Lemma 4.3.3, we immediately have

$$\left\|x_S - x^{p+1}\right\|_2 = \|x_S - \overline{w}^p \otimes u^p\|_2 \leq \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}}\,\|(x_S - u^p)\otimes\hat{w}\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}}\,\|\eta'\|_2\,. \tag{5.9}$$

By the structure of PGOT, we know $u^p = x^p + \mathcal{H}_q\left(A^T(y - Ax^p)\right)$. Denote $\Omega := \mathcal{L}_q(A^T(y - Ax^p))$. Applying Lemma 5.3.2, we have

$$\begin{aligned}\|x_S - u^p\|_2 &= \left\|x_S - x^p - \mathcal{H}_q\left(A^T(y - Ax^p)\right)\right\|_2 \\ &\leq \frac{\sqrt{5}+1}{2}\left\|\left(x_S - x^p - A^T(y - Ax^p)\right)_{\Omega\cup(S\cup S^p)}\right\|_2.\end{aligned} \tag{5.10}$$

We also have that

$$\begin{aligned}
&\left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{\Omega\cup S\cup S^p}\right\|_2 \\
&= \left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{\Omega\setminus(S\cup S^p)} + \left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{S\cup S^p}\right\|_2 \\
&\leq \left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{\Omega\setminus(S\cup S^p)}\right\|_2 + \left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{S\cup S^p}\right\|_2 \\
&\leq \left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{\Omega}\right\|_2 + \left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{S\cup S^p}\right\|_2 . \qquad (5.11)
\end{aligned}$$

Noting that $y = Ax_S + \eta'$, the second term of (5.11) on the right-hand side can be bounded by Lemma 5.3.1, such as

$$\begin{aligned}
\left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{S\cup S^p}\right\|_2 &\leq \left\|\left((I - A^TA)(x_S - x^p)\right)_{S\cup S^p}\right\|_2 + \left\|(A^T\eta')_{S\cup S^p}\right\|_2 \\
&\leq \delta_{2k}\left\|x_S - x^p\right\|_2 + \sqrt{1+\delta_{2k}}\left\|\eta'\right\|_2 , \qquad (5.12)
\end{aligned}$$

since $\text{card}(\text{supp}(x_S - x^p) \cup (S \cup S^p)) \leq 2k$. Separate the set Ω into t sets, such as $\Omega = T_1 \cup T_2 \ldots, T_t$, where $t = \left\lceil \frac{q}{k} \right\rceil$ and $\text{card}(T_i) \leq k$ for $i = 1, \ldots, t$. Thus, we can write

$$\begin{aligned}
\left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{\Omega}\right\|_2 &\leq \sum_{i=1}^{t}\left\|\left(x_S - x^p - A^T\left(y - Ax^p\right)\right)_{T_i}\right\|_2 \\
&\leq \sum_{i=1}^{t}\left\|\left((I - A^TA)(x_S - x^p)\right)_{T_i}\right\|_2 + \sum_{i=1}^{t}\left\|\left(A^T\eta'\right)_{T_i}\right\|_2 \\
&\leq t\delta_{3k}\left\|x_S - x^p\right\|_2 + t\sqrt{1+\delta_k}\left\|\eta'\right\|_2 , \qquad (5.13)
\end{aligned}$$

where the last inequality follows from Lemma 5.3.1 since $\text{card}(T_i \cup \text{supp}(x_S - x^p)) \leq 3k$. Due to the fact $\delta_k \leq \delta_{2k}$, and combining (5.10)-(5.13), we have

$$\begin{aligned}
\left\|(x_S - u^p) \otimes \hat{w}\right\|_2 &\leq \left\|x_S - u^p\right\|_2 \qquad \text{(since } 0 \leq \hat{w} \leq \mathbf{e}\text{)} \\
&\leq \frac{\sqrt{5}+1}{2}(t+1)\delta_{3k}\left\|x_S - x^p\right\|_2 + \frac{\sqrt{5}+1}{2}(t+1)\sqrt{1+\delta_{2k}}\left\|Ax_{\overline{S}} + \eta\right\|_2 . \qquad (5.14)
\end{aligned}$$

Substituting (5.14) into (5.9) yields

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho \left\|x^p - x_S\right\|_2 + \tau \left\|\eta'\right\|_2,$$

where

$$\rho = \frac{\sqrt{5}+1}{2}\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)\delta_{3k}\sqrt{\frac{1+\delta_k}{1-\delta_{2k}}} \tag{5.15}$$

and

$$\tau = \frac{(\sqrt{5}+1)\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)(1+\delta_{2k}) + 4}{2\sqrt{1-\delta_{2k}}}.$$

For the guaranteed performance of sequence $\{x^p\}$ generated by PGOT, what required is that $\rho < 1$ which implies

$$\frac{\sqrt{5}+1}{2}(t+1)\delta_{3k}\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} < 1,$$

since $\delta_k \leq \delta_{2k} \leq \delta_{3k}$. It is equivalent to

$$(t+1)^2\delta_{3k}^3 + (t+1)^2\delta_{3k}^2 + \frac{2}{3+\sqrt{5}}\delta_{3k} - \frac{2}{3+\sqrt{5}} < 0. \tag{5.16}$$

The left-hand side of the above inequality is clearly a cubic function in terms of δ_{3k}. Denote

$$f(\delta_{3k}) = (t+1)^2\delta_{3k}^3 + (t+1)^2\delta_{3k}^2 + \frac{2}{3+\sqrt{5}}\delta_{3k} - \frac{2}{3+\sqrt{5}}.$$

The derivative of f is given as

$$f'(\delta_{3k}) = 3(t+1)^2\delta_{3k}^2 + 2(t+1)^2\delta_{3k} + \frac{2}{3+\sqrt{5}},$$

which is always positive as a result of $\delta_{3k} \in (0, 1]$. Therefore, the function $f(\delta_{3k})$ is strictly monotone increasing over the interval $(0, 1]$. It is clear to see that

$$f(0) = -\frac{2}{3+\sqrt{5}} < 0, \quad f(1) = 2(t+1)^2 > 0,$$

by which we see that there always exists a unique real root $\alpha^* \in (0,1)$ satisfying $f(\alpha^*) = 0$. The inequality (5.16) holds under $\delta_{3k} < \alpha$, thus $\rho < 1$ can be guaranteed. The desired results follow.

In particular, if the target vector x is k-sparse and the measurements are exact, i.e., $y = Ax$, the sequence $\{x^p\}$ generated by PGOT converges to x as $p \to \infty$ with

$$\left\|x - x^{p+1}\right\|_2 \leq \rho \left\|x - x^p\right\|_2 \leq \rho^{p+1} \left\|x - x^0\right\|_2,$$

where ρ is the same as (5.15). ■

The bound of δ_{3k} in Theorem 5.3.3 is implicit, but we can provide an explicit bound of δ_{3k} for PGOT. As a result of

$$\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} < \frac{1+\delta_{3k}}{1-\delta_{3k}},$$

a sufficient condition for (5.7) is

$$\frac{\sqrt{5}+1}{2}(t+1)\delta_{3k}\frac{1+\delta_{3k}}{1-\delta_{3k}} < 1.$$

The relation above amounts to $h(\delta_{3k}) < 0$, where $h(\delta_{3k})$ is a quadratic function defined as

$$h\left(\delta_{3k}\right) = \phi\delta_{3k}^2 + (\phi+1)\delta_{3k} - 1, \quad \phi = \frac{\sqrt{5}+1}{2}(t+1).$$

The discriminant of $h(\delta_{3k})$ is $\Delta = (\phi+1)^2 + 4\phi$ which is always positive. Thus the range of δ_{3k} is

$$\frac{-(\phi+1) - \sqrt{\phi^2+6\phi+1}}{2\phi} < \delta_{3k} < \frac{-(\phi+1) + \sqrt{\phi^2+6\phi+1}}{2\phi}. \tag{5.17}$$

Clearly, the left hand side of (5.17) is negative, and it is easy to verify the right hand side is between 0 and 1. Together with $\delta_{3k} \in (0,1]$, a sufficient condition for $h(\delta_{3k}) < 0$ is

$$\delta_{3k} < \frac{-(\phi+1) + \sqrt{\phi^2+6\phi+1}}{2\phi}.$$

Rearranging terms leads to the following corollary.

Corollary 5.3.4. *Let $y = Ax + \eta$ be the measurements of the signal x and $q \geq 2k$ be an integer number. Suppose the restricted isometry constant of A satisfy*

$$\delta_{3k} < \frac{2}{\sqrt{\phi^2 + 6\phi + 1} + \phi + 1},$$

where

$$\phi = \frac{\sqrt{5}+1}{2}\left(\left\lceil \frac{q}{k} \right\rceil + 1\right).$$

Then the sequence $\{x^p\}$ generated by PGOT satisfies that

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho \left\|x^p - x_S\right\|_2 + \tau \|Ax_{\overline{S}} + \eta\|_2,$$

where ρ and τ are given by (5.7) *and* (5.8)*, respectively.*

5.3.2 PGROT in noisy scenarios

Theorem 5.3.5. *Let $y = Ax + \eta$ be the measurements of the signal x with measurement error η, and $q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of matrix A satisfies*

$$\delta_{3k} < \beta^*, \tag{5.18}$$

where β^* is the unique real root of the equation

$$9\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)^2 \beta^3 + 9\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)^2 \beta^2 + \frac{2}{7+3\sqrt{5}}\beta - \frac{2}{7+3\sqrt{5}} = 0 \tag{5.19}$$

in $(0, 1)$. Then the sequence $\{x^p\}$ generated by PGROT satisfies

$$\left\|x^{p+1} - x_S\right\|_2 \leq \overline{\rho}\|x^p - x_S\|_2 + \overline{\tau}\|Ax_{\overline{S}} + \eta\|_2, \tag{5.20}$$

where

$$\overline{\rho} = 3\left(\frac{\sqrt{5}+1}{2}\right)^2 \left(\left\lceil \frac{q}{k} \right\rceil + 1\right) \delta_{3k} \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}} < 1 \tag{5.21}$$

is ensured under (5.18)*, and the constant* $\overline{\tau}$ *is given as*

$$\overline{\tau} = \left(\frac{\sqrt{5}+1}{2}\right)^2 \frac{3\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)(1+\delta_k)}{\sqrt{1-\delta_{2k}}} + \frac{\sqrt{5}+1}{\sqrt{1-\delta_{2k}}}. \tag{5.22}$$

In particular, if x *is* k*-sparse and* $\eta = 0$*, the sequence* $\{x^p\}$ *generated by PGROT converges to* x*.*

Proof. Denote $S^{p+1} = \mathrm{supp}(x^{p+1})$. Applying Lemma 5.3.2, we have

$$\left\|x_S - x^{p+1}\right\|_2 = \left\|x_S - \mathcal{H}_k\left(w^p \otimes u^p\right)\right\|_2 \leq \frac{\sqrt{5}+1}{2} \|(x_S - w^p \otimes u^p)_{S \cup S^{p+1}}\|_2. \tag{5.23}$$

Note that $\hat{w}$ is a k-sparse binary vector satisfying $S \subseteq \mathrm{supp}(\hat{w})$. Let $y = Ax + \eta = Ax_S + \eta'$. Applying Lemma 4.3.7 leads to

$$\begin{aligned}
&\|(x_S - u^p \otimes w^p)_{S \cup S^{p+1}}\|_2 \\
&\leq \sqrt{\frac{1+\delta_k}{1-\delta_{2k}}} \left(\|(x_S - u^p) \otimes \hat{w}\|_2 + 2\left\|\mathcal{H}_k\left(u^p - x_S\right)\right\|_2\right) + \frac{2}{\sqrt{1-\delta_{2k}}} \|\eta'\|_2 \\
&\leq 3\sqrt{\frac{1+\delta_k}{1-\delta_{2k}}} \|x_S - u^p\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}} \|\eta'\|_2 ,
\end{aligned} \tag{5.24}$$

where the last inequality holds since $\|(x_S - u^p) \otimes \hat{w}\|_2 \leq \left\|\mathcal{H}_k\left(u^p - x_S\right)\right\|_2 \leq \|x_S - u^p\|_2$. By (5.23) and (5.24), we have

$$\left\|x_S - x^{p+1}\right\|_2 \leq \frac{\sqrt{5}+1}{2} \left(3\sqrt{\frac{1+\delta_k}{1-\delta_{2k}}} \|x_S - u^p\|_2 + \frac{2}{\sqrt{1-\delta_{2k}}} \|\eta'\|_2\right).$$

Let $t = \left\lceil \frac{q}{k} \right\rceil$. Based on (5.14), we have

$$\|x_S - u^p\|_2 \leq \frac{\sqrt{5}+1}{2}(t+1)\delta_{3k} \|x_S - x^p\|_2 + \frac{\sqrt{5}+1}{2}(t+1)\sqrt{1+\delta_k} \|\eta'\|_2 .$$

Combining the above two relations yields

$$\left\|x_S - x^{p+1}\right\|_2 \leq \overline{\rho}\left\|x_S - x^p\right\|_2 + \overline{\tau}\|\eta'\|_2 \tag{5.25}$$

where $\overline{\rho}$ and $\overline{\tau}$ are given as (5.22) and (5.22) respectively. Due to the fact $\delta_k \leq \delta_{2k} \leq \delta_{3k}$, the requirement for guaranteed performance $\overline{\rho} < 1$ implies

$$3\left(\frac{\sqrt{5}+1}{2}\right)^2 (t+1)\delta_{3k}\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} < 1, \tag{5.26}$$

which is equivalent to

$$9(t+1)^2\delta_{3k}^3 + 9(t+1)^2\delta_{3k}^2 + \frac{2}{7+3\sqrt{5}}\delta_{3k} - \frac{2}{7+3\sqrt{5}} < 0.$$

Denote

$$f\left(\delta_{3k}\right) = 9(t+1)^2\delta_{3k}^3 + 9(t+1)^2\delta_{3k}^2 + \frac{2}{7+3\sqrt{5}}\delta_{3k} - \frac{2}{7+3\sqrt{5}},$$

whose derivative is given as

$$f'\left(\delta_{3k}\right) = 27(t+1)^2\delta_{3k}^2 + 18(t+1)^2\delta_{3k} + \frac{2}{7+3\sqrt{5}}$$

which is always positive over the interval $(0, 1]$. Hence, there always exists a unique real positive root $\beta^* \in (0, 1)$ satisfying $f(\beta^*) = 0$ because

$$f(0) = -\frac{2}{7+3\sqrt{5}} < 0,\ f(1) = 18(t+1)^2 > 0.$$

Therefore, we see that once $\delta_{3k} < \beta^*$, the relation (5.26) can be guaranteed, and the desired results follow.

Especially, when the target vector x is k-sparse and $\eta = 0$, the relation (5.25) will be reduced to

$$\left\|x - x^{p+1}\right\|_2 \leq \overline{\rho}\left\|x - x^p\right\|_2 \leq \overline{\rho}^{p+1}\left\|x - x^0\right\|_2.$$

Therefore, the sequence $\{x^p\}$ generated by RPGOT converges to x when $p \to \infty$. ■

We can also provide an explicit bound of δ_{3k} for PGROT by using the relation

$$\sqrt{\frac{1+\delta_{3k}}{1-\delta_{3k}}} < \frac{1+\delta_{3k}}{1-\delta_{3k}}.$$

A sufficient condition for $\overline{\rho} < 1$ is

$$3\left(\frac{\sqrt{5}+1}{2}\right)^2 \left(\left\lceil \frac{q}{k} \right\rceil + 1\right) \delta_{3k} \frac{1+\delta_{3k}}{1-\delta_{3k}} < 1,$$

which implies

$$\psi\delta_{3k}^2 + (\psi+1)\delta_{3k} - 1 < 0, \text{ where } \psi = 3\left(\frac{\sqrt{5}+1}{2}\right)^2 \left(\left\lceil \frac{q}{k} \right\rceil + 1\right).$$

By the same analysis as in Section 5.3.1, we immediately have the following corollary.

Corollary 5.3.6. *Let $y = Ax + \eta$ be the measurements of the signal x with measurement error η, and $q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of matrix A satisfies*

$$\delta_{3k} < \frac{2}{\sqrt{\psi^2 + 6\psi + 1} + \psi + 1},$$

where

$$\psi = 3\left(\frac{\sqrt{5}+1}{2}\right)^2 \left(\left\lceil \frac{q}{k} \right\rceil + 1\right).$$

Then the sequence $\{x^p\}$ generated by PGROT satisfies

$$\left\| x^{p+1} - x_S \right\|_2 \leq \overline{\rho} \|x^p - x_S\|_2 + \overline{\tau} \|Ax_{\overline{S}} + \eta\|_2,$$

where $\overline{\rho}$ and $\overline{\tau}$ are given as (5.21) *and* (5.22)*, respectively.*

5.3.3 PGROTP in noisy scenarios

The PGROTP is a combination of PGROT and a pursuit step. Thus we can still utilize Lemma 3.3.8 to complete our analysis.

Theorem 5.3.7. *Let $y = Ax + \eta$ be the noisy measurements of the signal x with noise η, and $q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of matrix A satisfies*

$$\delta_{3k} < \frac{1}{3\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\left\lceil\frac{q}{k}\right\rceil+1\right)+1}. \tag{5.27}$$

Then the sequence $\{x^p\}$ generated by PGROTP satisfies

$$\left\|x^{p+1}-x_S\right\|_2 \leq \rho^*\|x^p-x_S\|_2+\tau^*\|Ax_{\overline{S}}+\eta\|_2, \tag{5.28}$$

where

$$\rho^* = \left(\frac{\sqrt{5}+1}{2}\right)^2\frac{3\left(\left\lceil\frac{q}{k}\right\rceil+1\right)\delta_{3k}}{1-\delta_{3k}} < 1 \tag{5.29}$$

is guaranteed by (5.27)*, and the constant τ^* is given as*

$$\tau^* = \frac{3\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\left\lceil\frac{q}{k}\right\rceil+1\right)(1+\delta_k)+\sqrt{5}+1}{(1-\delta_{2k})\sqrt{1+\delta_{2k}}}+\frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}. \tag{5.30}$$

In particular, when x is k-sparse and $\eta = 0$, the sequence $\{x^p\}$ generated by PGROTP converges to x.

Proof. The PGROTP can be regarded as a combination of PGROT with a pursuit step, thus we could analyze these two steps separately. Denote $\overline{x}^{p+1}$ as the intermediate vector generated by PGROT. Based on the analysis of PGROT algorithm, we have

$$\|\overline{x}^{p+1}-x_S\|_2 \leq \overline{\rho}\,\|x_S-x^p\|_2+\overline{\tau}\|\eta'\|_2$$

where $\overline{\rho}$ and $\overline{\tau}$ are the same as (5.21) and (5.22) respectively. Applying Lemma 3.3.8, we

immediately have the relation for pursuit step, such as

$$\left\|x_S - x^{p+1}\right\|_2 \leq \frac{1}{\sqrt{1-(\delta_{2k})^2}}\|\overline{x}^{p+1} - x_S\|_2 + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}\|\eta'\|_2.$$

Using the relationship of RICs, i.e., $\delta_k \leq \delta_{2k} \leq \delta_{3k}$, and combining above two inequalities, we have

$$\left\|x_S - x^{p+1}\right\|_2 \leq \rho^* \left\|x_S - x^p\right\|_2 + \tau^*\|\eta'\|_2,$$

where ρ^* and τ^* are given in (5.29) and (5.30) respectively. To guarantee the recovery of x_S by PGROTP, it is sufficient to require $\rho^* < 1$ which is equivalent to

$$\delta_{3k} < \frac{1}{3\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\left\lceil\frac{q}{k}\right\rceil+1\right)+1}$$

Thus the desired results follow. ■

Remark 5.3.1

We list a few RIC conditions in terms of δ_{3k} for the proposed algorithms with different q, i.e., $q = 2k, 3k, 4k$. The results shown in Table 5.3.1 are derived based on (5.5), (5.19) and (5.27) for the q as given above, respectively. It is worth mentioning that, when $q = n$, the partial gradient $\mathcal{H}_q(\nabla f(x))$ becomes the full gradient $\nabla f(x)$, and the partial gradient optimal thresholding algorithms are reduced to the optimal k-thresholding (OT), relaxed optimal k-thresholding (ROT) and relaxed optimal k-thresholding pursuit (ROTP) algorithms, respectively. The sufficient conditions for the convergence of these algorithms are studied in [134] and [139].

The value of q	PGOT	RPGOT	RPGOTP
$q = 2k$	$\delta_{3k} < 0.1729$	$\delta_{3k} < 0.0407$	$\delta_{3k} < 0.0407$
$2k < q \leq 3k$	$\delta_{3k} < 0.1348$	$\delta_{3k} < 0.0308$	$\delta_{3k} < 0.0308$
$3k < q \leq 4k$	$\delta_{3k} < 0.1106$	$\delta_{3k} < 0.0248$	$\delta_{3k} < 0.0248$

Table 5.1: The upper bounds of δ_{3k} for several different q

Remark 5.3.2

The quadratic optimization problem in the proposed algorithms, i.e., Algorithms 5.2 and 5.3, can be written as

$$\min_{w} \left\{ \|y - [A_\Lambda]\,(w \otimes [u^p]_\Lambda)\|_2^2 : \mathbf{e}^T w = k, w \in \{0,1\}^{(k+q)} \right\},$$

where $\Lambda = \mathrm{supp}(u^p)$ and $u^p = x^p + \lambda \mathcal{H}_q\left(A^T(y - Ax^p)\right)$. In this case, the complexity of solving the problem becomes $O((q+k)^{3.5}L)$ instead of $O(n^{3.5}L)$, where L is the size of the problem data encoding in binary. And according to Remark 4.3.3, it is easy to calculate that the complexity of each iteration of PGROT is $O(mn + (q+k)^{3.5}L)$, and each iteration of PGROTP requires about $O(mn + (q+k)^{3.5}L + k^3)$ operations when applying least square method to the pursuit step or about $O(mn + (q+k)^{3.5}L + m^3)$ flops when the Cholesky decomposition strategy is introduced to work on the matrix $[A_\Omega]^T [A_\Omega]$. The primary cost of the relaxed-optimal-hard-thresholding-based algorithms generally lies in solving quadratic optimization problems. The partial-gradient-based algorithm in this paper is quite different from the strategy in [134]. To improve the chance for the solution of (ROT) to be binary or nearly binary, two or more relaxed optimal k-thresholding steps are employed by the algorithms in [134]. However, the iterate generated by partial gradient methods is at most $(k+q)$-sparse, based on whether the vector derived from the relaxed k-optimal thresholding is more compressible than the one generated by the algorithms in [134].

5.3.4 Solving sparse optimization problem

Our results can be generalized to the case of solving the sparse optimization problem (5.3), which provides an estimation of the distance between the iterate x^p generated by the proposed algorithms and the global solution x^* of the problem (5.3), given in the following Theorems, respectively.

Theorem 5.3.8. *Let x^* be the solution to the problem* (5.3)*, and $\eta^* := y - Ax^*$. Let*

$q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of matrix A satisfies one of the following conditions:

(i) $\delta_{3k} < \alpha^*$, where $\alpha \in (0,1)$ is the unique real root of

$$\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)^2 \alpha^3 + \left(\left\lceil \frac{q}{k} \right\rceil + 1\right)^2 \alpha^2 + \frac{2}{3+\sqrt{5}}\alpha - \frac{2}{3+\sqrt{5}} = 0,$$

(ii) δ_{3k} satisfies

$$\delta_{3k} < \frac{2}{\sqrt{\phi^2 + 6\phi + 1} + \phi + 1}, \text{ where } \phi = \frac{\sqrt{5}+1}{2}\left(\left\lceil \frac{q}{k} \right\rceil + 1\right).$$

The sequence generated by PGOT satisfies the relation (5.6).

Theorem 5.3.9. *Let x^* be the solution to the problem* (5.3)*, and $\eta^* := y - Ax^*$. Let $q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of matrix A satisfies one of the following conditions:*

(i) $\delta_{3k} < \beta^*$, where $\beta^* \in (0,1)$ is the real root of

$$9\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)^2 \beta^3 + 9\left(\left\lceil \frac{q}{k} \right\rceil + 1\right)^2 \beta^2 + \frac{2}{7+3\sqrt{5}}\beta - \frac{2}{7+3\sqrt{5}} = 0,$$

(ii) δ_{3k} satisfies

$$\delta_{3k} < \frac{2}{\sqrt{\psi^2 + 6\psi + 1} + \psi + 1}, \text{ where } \psi = 3\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\left\lceil \frac{q}{k} \right\rceil + 1\right).$$

Then the sequence generated by RPGOT satisfies the relation (5.20).

Theorem 5.3.10. *Let x^* be the solution to the problem* (5.3)*, and $\eta^* := y - Ax^*$. Let $q \geq 2k$ be a positive integer number. Suppose the restricted isometry constant of matrix A satisfies the following condition:*

$$\delta_{3k} < \frac{1}{3\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\left\lceil \frac{q}{k} \right\rceil + 1\right) + 1}.$$

Then the sequence generated by RPGOTP satisfies the relation (5.28).

These results demonstrate how far the iterate x^{p+1} generated by the algorithm is from the optimal solution x^* of the problem (5.3) and show that the bound of $\|x^{p+1} - x^*\|$ depends on the value $\|\eta^*\|_2 = \|y - Ax^*\|_2$.

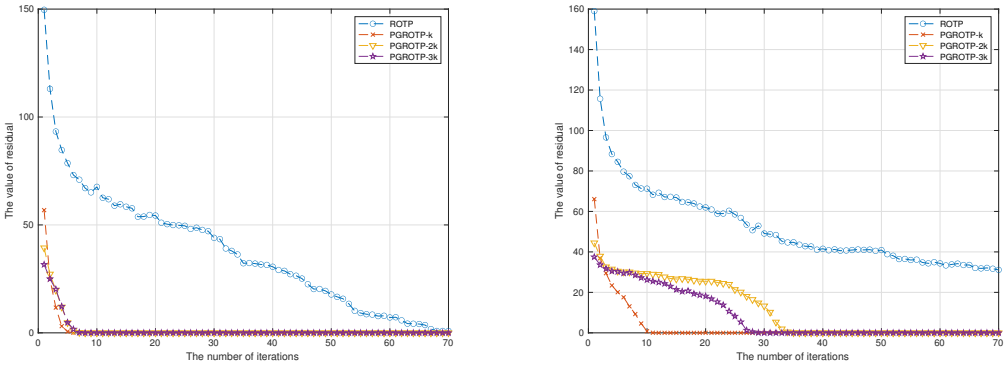
5.4 Numerical Experiments

Simulations via synthetic data are carried out to demonstrate the numerical performance of the PGROTP which is the main implementable algorithm proposed in this chapter. We test the algorithm on four aspects: objective reduction, the average number of iterations required for solving the problem (5.3), success frequency in vector reconstruction and phase transition in terms of success rate. Unless otherwise stated, the measurement matrices used in experiments are Gaussian random matrices whose entries follow standard normal distribution $\mathcal{N}(0, 1)$. For sparse vectors, their entries also follow the $\mathcal{N}(0, 1)$ and the position of nonzero entries of the vector follows the uniform distribution. All the convex optimization problems arising were solved by CVX developed by Grant and Boyd [66] with solver 'Mosek' [3].

5.4.1 Residual reduction

This experiment is used to investigate the objective-reduction performance of the PGROTP with different q, including $q = k, 2k, 3k$ and n. We set $A \in \mathbb{R}^{500\times1000}$ and $y = Ax^*$, where x^* is a generated k -sparse vector. Thus x^* is a global solution of the problem (7). Fig. 1 records the changes of the objective value $\|y - Ax\|_2$ in the course of the algorithm up to 70 iterations. Fig. 1 (a) and Fig. 1 (b) include the results for the sparsity level $\|x^*\|_0 = 162$ and 197 , respectively. It can be seen that PGROTP is able to reduce the objective value during iterations. Moreover, this experiment also indicates that the optimal k -thresholding methods with partial gradients often perform better in objective reduction

than the ones using full gradients.



(a) Sparsity level $k = 162$ (b) Sparsity level $k = 197$

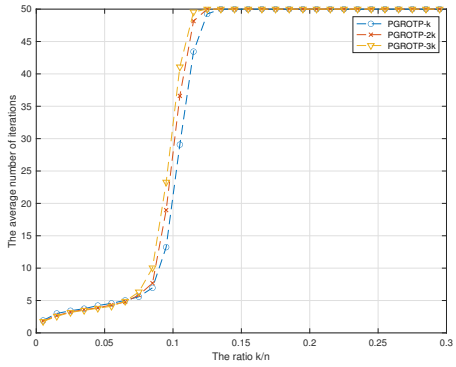
Figure 5.1: Residual change in the course of iterations for PGROTP with different q, i.e., $q = k, 2k, 3k, n$. The size of measurement matrix is 500×1000, and the measurements are exact, i.e., $y = Ax^*$.

5.4.2 Number of iterations

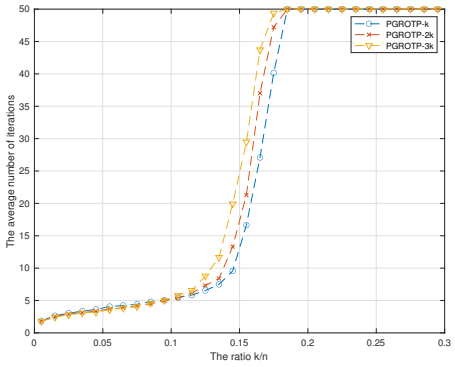
Experiments were also performed to demonstrate the average number of iterations needed for PGROTP to reconstruct the sparse signal. The signal length is fixed to be 1000, and the size of the measurement matrix is $m \times 1000$, where m takes the following a few different values: $m = 300, 400, 500, 600$. The measurements $y = Ax^*$ are accurate. In this experiment, if the iterate x^p satisfies the criterion

$$\frac{\|x - x^*\|_2}{\|x^*\|_2} \leq 10^{-3}, \tag{5.31}$$

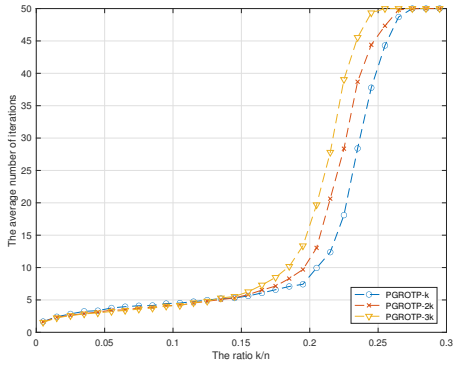
the algorithm terminates and the number of iterations p is recorded. If the algorithm cannot meet the criterion (5.31) within 50 iterations, the algorithm stops, and the number of iterations performed is recorded as 50. For each given sparsity level, the average number of iterations is obtained by attempting 50 trials. The outcomes are shown in Fig. 5.2 which indicate that the required iterations of PGROTP for vector reconstructions are usually low when the sparsity level of x^* is low, and that the number of iterations which is required for solving the problem increases as the sparsity level increases. This figure also shows



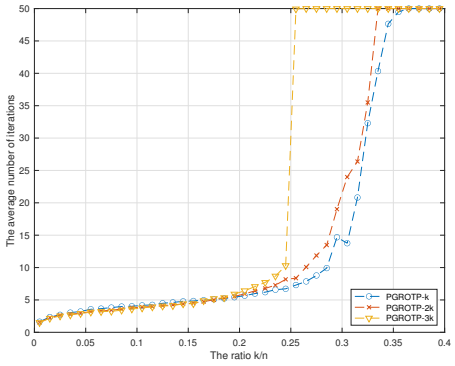
(a) $m = 300$



(b) $m = 400$



(c) $m = 500$



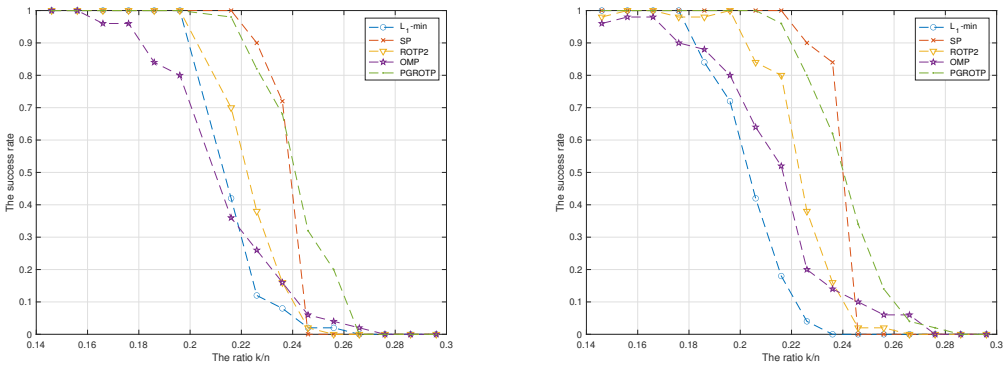
(d) $m = 600$

Figure 5.2: Comparison of the average number of iterations required by different q. The length of the target signal is 1000, and measurements are exact, i.e., $y = Ax^*$.

that which is for a given sparsity level, the more measurements are required, the lower the average number of iterations are needed by the PGROTP to meet the reconstruction criterion (5.31).

5.4.3 Sparse signal recovery

Simulations were also carried out to compare the success rates of the PGROTP algorithm in sparse vector reconstruction with several existing algorithms, such as ℓ_1-minimization, SP, OMP and ROTP2 (see [134] and [139]). The size of the measurement matrix is still 500×1000. For every given ratio of the sparsity level k and n, the success rate of the algorithm is obtained by 50 random attempts. In this experiment, SP, ROTP2 and PGROTP perform a total of 50 iterations, whereas OMP is performed k iterations. After performing the required iterations, the algorithm is counted as success if the condition (5.31) is satisfied. The success rates for accurate and inaccurate measurements are summarized in Fig. 3 (a) and (b), respectively. The inaccurate measurements are given as $y = Ax^* + 0.001\eta$, where η is a standard Gaussian vector. Compared with several existing algorithms, it can be seen that the PGROTP is robust and efficient for the sparse vector reconstruction in both noise and noiseless environment.



(a) Exact measurements

(b) Inexact measurements

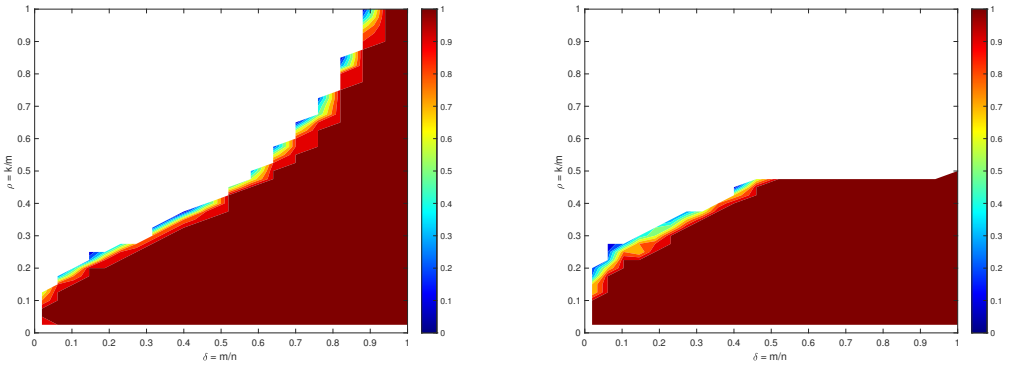
Figure 5.3: Comparisons of success rate of sparse signal reconstruction between several algorithms.

5.4.4 Phase transition

We also conducted experiments to demonstrate the phase transition of PGROTP($q = k$) and compare with ℓ_1-minimization, SP and ROTP2 algorithms. In this experiment, the length of the target vector x^* is fixed as $n = 2048$. We take 40 different sparsity levels uniformly from 1 to m, i.e., $k = \left\lfloor \frac{1}{40}m \right\rfloor, \ldots, \left\lfloor \frac{39}{40}m \right\rfloor, n$. For the number of measurements m, we require $m \geq 40$ to ensure that there are 40 different sparsity levels. We take a total of 20 different values for the measurements, 10 uniformly from $[40, 0.4n]$ and 10 from $(0.4n, n]$. The matrix A is a Gaussian random matrix whose entries follow the normal distribution $\mathcal{N}(0, 1/\sqrt{m})$. The measurements of x^* are inexact, i.e., $y = Ax^* + 10^{-5}\eta$, where η is a standard Gaussian vector. The iterative algorithms terminate either when the criterion (5.31) is met, or the maximum of 30 iterations are reached. The algorithm is counted as success when (5.31) is satisfied. For each pair of (m, k), we use 10 trials to calculate the success rate and the average time consumption for success recoveries. In Fig. 5.4-5.7, the vertical axis represents over-sampling rate $\rho = k/m$, while the horizontal axis represents under-sampling rate $\delta = m/n$.

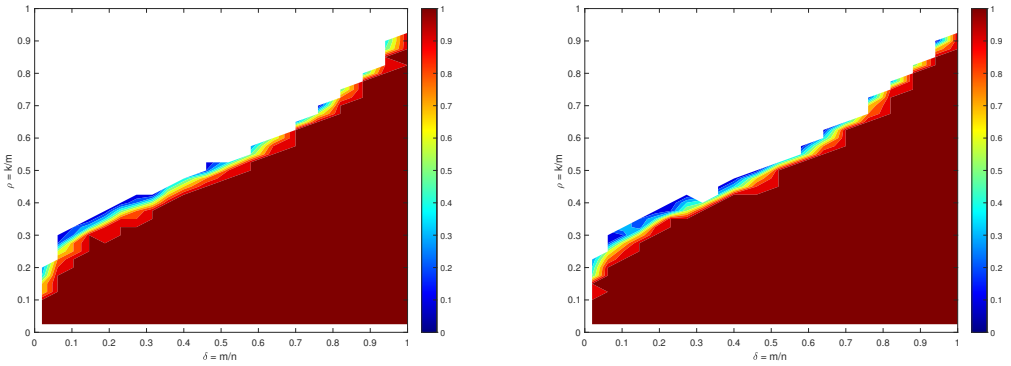
Fig. 5.4 shows how the success rate changes with δ and ρ for these four algorithms. It can be seen that all four algorithms are efficient for sparse signal recovery. What surprised us is the performance of the SP algorithm, which fails to recover the signal when $\rho > 0.5$ even if the number of measurements increases, except for when the measurement matrix is square. In Fig. 5.5, we plot their phase transition curves to make an intuitional comparison of their recovery capability of sparse signal recovery.

For every given δ, the logistic regression model is used to determine the specific success rate corresponding to ρ. The phase transition curves can be obtained in terms of different success rates, where Fig. 5.5 (a)-(c) demonstrate the curves for the success rates 0.9, 0.5 and 0.1, respectively. Such curves delineate the regions where the algorithms can recover the target signal with a specific success rate and the region where they cannot. In general, the higher the curve, the better the performance of the algorithm. Fig. 5.5 indicates that PGROTP performs comparably to the other three algorithms. We also see that



(a) ℓ_1-minimization

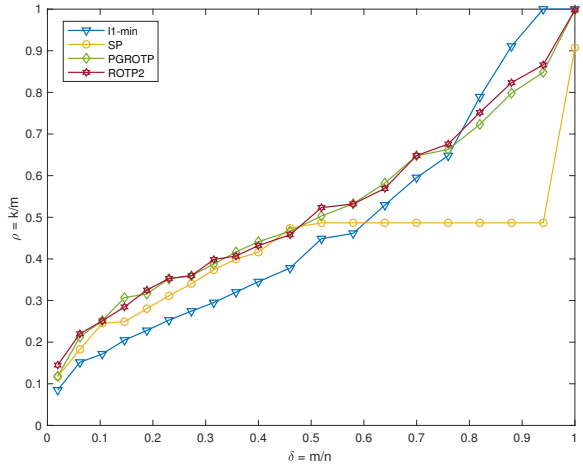
(b) SP



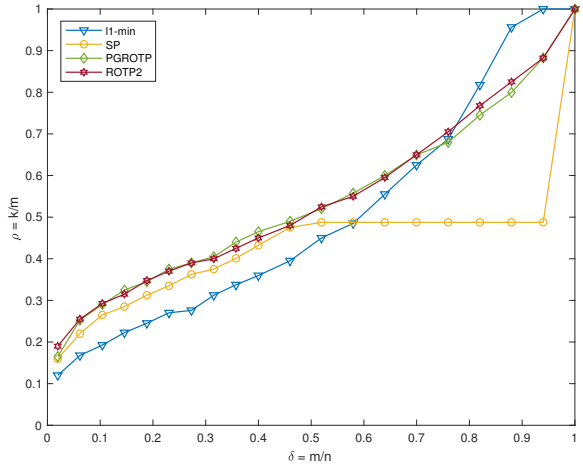
(c) PGROTP

(d) ROTP2

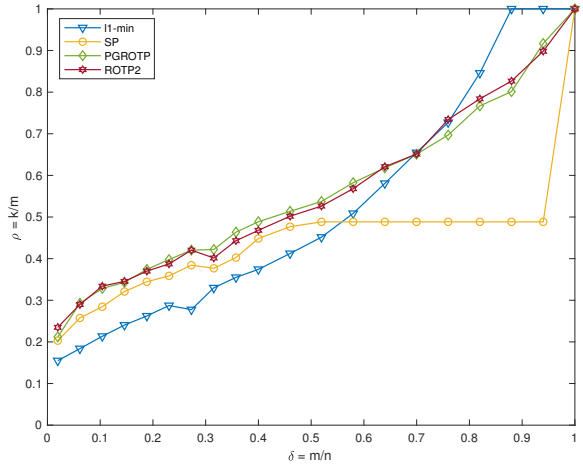
Figure 5.4: Phase transition of ℓ_1-minimization, SP, PGROTP and ROTP2 algorithms in terms of success rates



(a) success rate = 0.9

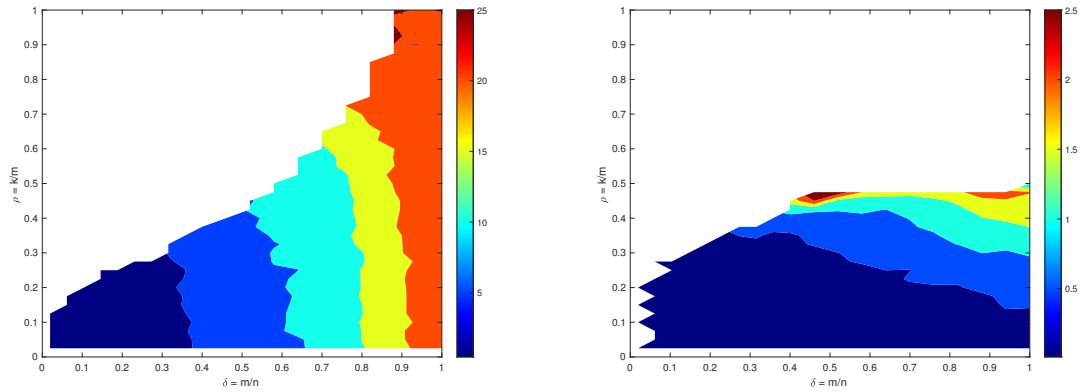


(b) success rate = 0.5



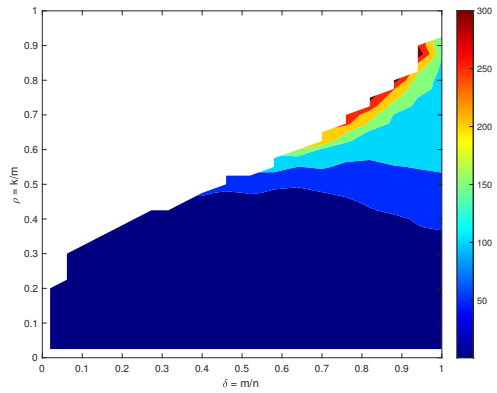
(c) success rate = 0.1

Figure 5.5: Phase transition curve of ℓ_1-minimization, SP, PGROTP and ROTP2 under inexact measurements in terms of success rates

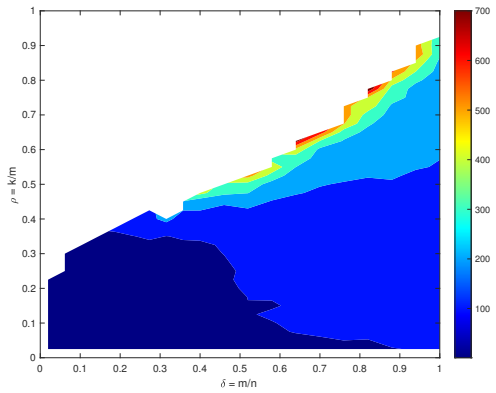


(a) ℓ_1-minimization

(b) SP



(c) PGROTP



(d) ROTP2

Figure 5.6: Time consumption of ℓ_1-minimization, SP, PGROTP and ROTP2 algorithms for successful recovery

although PGROTP adopts a partial gradient direction, it maintains a similar performance to ROTP2 in sparse signal recovery.

We now shed some light on the comparison between the CPU time for successful recovery of these four algorithms, as shown in Fig. 5.6. For iterative-type algorithms, it can be seen that when more measurements are used to recover a denser signal, more time is required for successful reconstruction. The graph for ℓ_1-minimization indicates that its time consumption is highly related to the size of the measurement matrix rather than the sparsity level, which is quite different from the case of iterative methods. And we also see that the optimal-thresholding-type algorithms require much more time than the others due to solving the quadratic optimization problem.

Finally, we compare the performance of PGROTP and ROTP2 via time consumption ratio. Fig. 5.7 shows the time ratio $r = \dfrac{\text{time}(\text{ROTP2}(\delta, \rho))}{\text{time}(\text{PGROTP}(\delta, \rho))}$ over the phase transition region. The difference in time consumption between the two algorithms is more pronounced when more measurements are used to recover a sparse signal. Fig. 5.7 clearly shows that when dealing with signals with low sparsity levels, the PGROTP algorithm can recover a wide range of signals with less time than ROTP2. This does show that using the partial gradient information may help reduce the runtime of the optimal k-thresholding method.

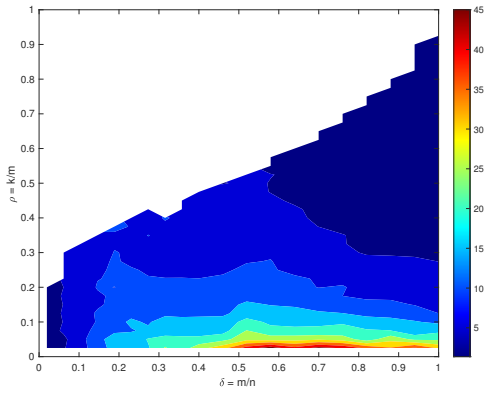


Figure 5.7: Time consumption ratio of PGROTP and ROTP2

5.4.5 Main findings

The main findings from numerical performances can be summarized as follows. In reducing the objective value, optimal k-thresholding techniques using partial gradients perform better than those using full gradients. When the sparsity level is relatively low relative to its length, PGROTP requires a small number of iterations to recover a sparse signal. The PGROTP is robust and efficient for sparse signal recovery, and it is also comparable to ℓ_1-minimization, SP and ROTP2 in terms of the success rates. The optimal-thresholding-type algorithms usually require more computational time to solve the sparse optimization problem than traditional hard thresholding and existing heuristic algorithms. The primary computational task is to solve quadratic optimization problems. Nevertheless, optimal-thresholding algorithms have the advantage of being more stable, robust and capable of stronger reconstruction than some existing methods.

5.5 Conclusion

Motivated by the recently proposed optimal k-thresholding technique, we developed the partial gradient-based optimal k-thresholding method. Under the restricted isometry property (RIP), we established sufficient conditions including the convergence of the algorithms when $q \in [2k, n)$ for sparse signal recovery. Numerical experiments demonstrate that the

PGROTP algorithm is efficient for sparse signal reconstruction, and it is comparable to some other popular compressed sensing algorithms. The framework of gradient optimal thresholding can be extended to solve the general sparse optimization problem (3.4). We leave this as future work.

CHAPTER 6

OPTIMAL PARTIAL GRADIENT HARD THRESHOLDING ALGORITHMS

6.1 Introduction

In Chapter 5, we presented methods that combine the partial gradient method and optimal k-thresholding technique, and demonstrated that they are efficient for sparse signal reconstruction. For unconstrained optimization problems, using a partial gradient may slow down the decrease of the objective value compared to the full gradient. However, for sparsity-constrained optimization problems, it is helpful to reduce a sudden increase of the objective value during the course of iterations. In this chapter, we adopt the so-called optimal partial gradient direction to develop a new hard-thresholding-type approach.

We still consider the problem of the form:

$$\min_x \{f(x) : \|x\|_0 \leq k\}. \tag{6.1}$$

The ideal search direction of an iterative algorithm for solving (6.1) is the one that could simultaneously keep the sparsity of iterates and reduce the objective value, i.e.,

$$f(x^{p+1}) = f(x^p + \mathbf{d}) < f(x^p),$$

where both the new iterate x^{p+1} and the current iterate x^p are k-sparse. However,

such a direction is usually difficult to find. Thresholding methods have been proposed as compromises by generating an iterate using traditional descent directions and then restoring sparsity by using specific thresholding strategies. However, as pointed out in [134], imposing the hard thresholding operator on a dense vector may increase the objective value, thus we adopt the partial gradient to generate an iterate sparser than the one generated by a full gradient. In Chapter 5, we use the major part of the gradient, i.e., $\mathcal{H}_q(\nabla f(x))$ where $q > 0$ is an integer. In this chapter, we adopt the so-called optimal partial gradient that minimizes the objective function, leading to

$$u^p = x^p + \lambda\ \mathbf{d}_{\text{opt-par}}(x^p) = x^p - \lambda \nabla f(x^p) \otimes w^*, \tag{6.2}$$

where $\lambda > 0$ denotes the stepsize, and the vector w^* is given as

$$w^* = \arg\min_w \left\{ f\left(x^p - \lambda \nabla f(x^p) \otimes w\right) : \mathbf{e}^T w = q, w \in \{0,1\}^n \right\}.$$

The integer $q > 0$ is used to determine how many elements are extracted from the full gradient. Since the iterate u^p generated by (6.2) may not be k-sparse, we use the hard thresholding operator $\mathcal{H}_k$ to generate the iterate satisfying the constraint of (6.1), leading to

$$x^{p+1} = \mathcal{H}_k\left(x^p + \lambda\ \mathbf{d}_{\text{opt-par}}(x^p)\right),$$

which is the key step for the proposed algorithms in this chapter. The main contribution can be highlighted as follows.

i. We introduce a class of optimal-partial-gradient-based hard thresholding methods, including optimal partial gradient hard thresholding (OPGHT), relaxed optimal partial gradient hard thresholding (ROPGHT) and relaxed optimal partial gradient hard thresholding pursuit (ROPGHTP), described in Algorithms 6.1-6.3, respectively.

ii. Theoretical analysis of the proposed algorithms for sparse signal recovery was carried out using the restricted isometry property (RIP). The main results are given in

Theorems 6.3.1, 6.3.3, and 6.3.4, respectively.

iii. Numerical experiments indicate that the proposed algorithms are comparable to the hard threshold pursuit (HTP) algorithm, and the ROPGHTP is less sensitive to the choice of stepsize than HTP.

6.2 Algorithms

We focus on the specific problem

$$\min_x \left\{ \frac{1}{2}\|y - Ax\|_2^2 : \|x\|_0 \leq k \right\}, \tag{6.3}$$

for which the key step for optimal gradient hard thresholding can be written as

$$x^{p+1} = \mathcal{H}_k \left(x^p - \lambda \nabla f(x^p) \otimes w^p \right),$$

where $\nabla f(x) = -A^T(y - Ax)$ and w^p is the minimizer of the following problem:

$$w^p = \arg\min_w \left\{ \|y - A\left(x^p - \lambda \nabla f(x^p) \otimes w\right)\|_2^2 : \mathbf{e}^T w = q, w \in \{0,1\}^n \right\}. \tag{6.4}$$

The optimal partial gradient hard thresholding (OPGHT) is given as Algorithm 6.1 with stepsize $\lambda \equiv 1$. The problem in (6.4) is a binary optimization problem. To avoid possible expensive computation, we can solve its relaxation version as follows:

$$\overline{w}^p = \arg\min_w \left\{ \|y - A\left(x^p - \lambda \nabla f(x^p) \otimes w\right)\|_2^2 : \mathbf{e}^T w = q, 0 \leq w \leq \mathbf{e} \right\},$$

based on which the relaxed version of OPGHT is stated as Algorithm 6.2. Moreover, it can be further accelerated by using a pursuit step, leading eventually to the so-called relaxed optimal partial gradient hard thresholding pursuit (ROPGHTP), which is described as Algorithm 6.3.

Algorithm 6.1 Optimal Partial Gradient Hard Thresholding (OPGHT)

- Input: measurement matrix A, measurements y, sparsity level k, integer $q > 0$ and an initial point x^0.
- Iteration:

$$d^p = A^T(y - Ax^p) \tag{OPGHT$_1$}$$
$$w^p = \arg\min_w \{\|y - A\left(x^p + w \otimes d^p\right)\|_2^2 : \mathbf{e}^T w = q, w \in \{0,1\}^n\} \tag{OPGHT$_2$}$$
$$x^{p+1} = \mathcal{H}_k(x^p + w^p \otimes d^p) \tag{OPGHT$_3$}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Algorithm 6.2 Relaxed Optimal Partial Gradient Hard Thresholding (ROPGHT)

- Input: measurement matrix A, measurements y, sparsity level k, integer $q > 0$ and an initial point x^0.
- Iteration:

$$\begin{aligned} d^p &= A^T(y - Ax^p) \\ \overline{w}^p &= \arg\min_w \{\|y - A\left(x^p + w \otimes d^p\right)\|_2^2 : \mathbf{e}^T w = q, 0 \leq w \leq \mathbf{e}\} \\ v^p &= \mathcal{H}_q(\overline{w}^p \otimes d^p) \\ x^{p+1} &= \mathcal{H}_k(x^p + v^p) \end{aligned}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

Algorithm 6.3 Relaxed Optimal Partial Gradient Hard Thresholding Pursuit (ROPGHTP)

- Input: measurement matrix A, measurements y, sparsity level k, integer $q > 0$ and an initial point x^0.
- Iteration:

$$\begin{aligned} d^p &= A^T(y - Ax^p) \\ \overline{w}^p &= \arg\min_w \{\|y - A\left(x^p + w \otimes d^p\right)\|_2^2 : \mathbf{e}^T w = q, 0 \leq w \leq \mathbf{e}\} \\ S^{p+1} &= \mathcal{L}_k(x^p + \mathcal{H}_q(\overline{w}^p \otimes d^p)) \\ x^{p+1} &= \arg\min_z \{\|y - Az\|_2^2 : \text{supp}(z) \subseteq S^{p+1}\} \end{aligned}$$

- Output: k-sparse vector $x^* := x^{p+1}$.

6.3 Theoretical Analysis

In this section, under the restricted isometry property (RIP), we establish the guaranteed performance of the proposed algorithms with $q = 2k$.

6.3.1 Analysis of OPGHT in noisy scenarios

The main result for OPGHT is summarized as follows.

Theorem 6.3.1. *Let $y = Ax + \eta$ be the noisy measurements of the signal x, and $q = 2k$. Suppose the restricted isometry constant δ_{4k} of the measurement matrix A satisfies that*

$$\delta_{4k} < 0.4031.$$

The sequence generated by the OPGHT satisfies

$$\left\|x^{p+1} - x_S\right\|_2 \leq \frac{\sqrt{5}+1}{2}\delta_{2k}\sqrt{\frac{1+\delta_{2k}}{1-\delta_{4k}}}\|x^p - x_S\|_2 + \frac{(\sqrt{5}+1)(3+\delta_{2k})}{2\sqrt{1-\delta_{4k}}}\|Ax_{\overline{S}} + \eta\|_2.$$

In particular, if x is k-sparse and $\eta = 0$, the sequence generated by the OPGHT converges to x.

Proof. Let $S^{p+1} = \mathcal{L}_k\left(x^p + w^p \otimes u^p\right)$ and $S = \mathcal{L}_k(x)$. Applying Lemma 5.3.2 we immediately have that

$$\begin{aligned}\left\|x^{p+1} - x_S\right\|_2 &= \left\|\mathcal{H}_k\left(x^p + w^p \otimes u^p\right) - x_S\right\|_2 \\ &\leq \frac{\sqrt{5}+1}{2}\left\|\left(x^p + w^p \otimes d^p - x_S\right)_{S\cup S^{p+1}}\right\|_2 \\ &\leq \frac{\sqrt{5}+1}{2}\left\|x^p - x_S + w^p \otimes d^p\right\|_2. \end{aligned} \tag{6.5}$$

Denote $y = Ax + \eta = Ax_S + \eta'$ with $\eta' = Ax_{\overline{S}} + \eta$, where $\overline{S}$ denotes the complementary

set of S. Due to the fact $\text{card}(S \cup S^p \cup \text{supp}(w^p)) \leq 4k$, we can write

$$\begin{aligned}\|y - A(x^p + w^p \otimes d^p)\|_2 &= \|A(x_S - x^p - w^p \otimes d^p) + \eta'\|_2 \\ &\geq \|A(x_S - x^p - w^p \otimes d^p)\|_2 - \|\eta'\|_2 \\ &\geq \sqrt{1-\delta_{4k}}\|x_S - x^p - w^p \otimes d^p\|_2 - \|\eta'\|_2, \end{aligned} \tag{6.6}$$

where the last inequality follows from the RIP since $\text{card}(x_S - x^p - w^p \otimes d^p) \leq 4k$. By an analysis similar to the proof of Theorem 4.3 in [134], we may bound the left-hand side of (6.6). In fact, there exists a $2k$-sparse vector $\tilde{w} \in \{0,1\}^n$ such that $\tilde{w} \otimes (x_S - x^p) = x_S - x^p$. By the structure of algorithm, the vector w^p is the optimal solution to problem (OPGHT_2), therefore we can write

$$\begin{aligned}\|y - A(x^p + w^p \otimes d^p)\|_2 &\leq \|y - A(x^p + \tilde{w} \otimes d^p)\|_2 \\ &= \|A\left(x_S - x^p - \tilde{w} \otimes d^p\right) + \eta'\|_2 \\ &= \|A((x_S - x^p - d^p) \otimes \tilde{w})) + \eta'\|_2 \\ &\leq \|A((x_S - x^p - d^p) \otimes \tilde{w}))\|_2 + \|\eta'\|_2 \\ &\leq \sqrt{1+\delta_{2k}}\|(x_S - x^p - d^p) \otimes \tilde{w}\|_2 + \|\eta'\|_2, \end{aligned} \tag{6.7}$$

where the last inequality holds since $\text{card}(\text{supp}(\tilde{w})) \leq 2k$. Combining (6.6) and (6.7) leads to

$$\|x_S - x^p - w^p \otimes d^p\|_2 \leq \sqrt{\frac{1+\delta_{2k}}{1-\delta_{4k}}}\|(x_S - x^p - d^p) \otimes \tilde{w}\|_2 + \frac{2}{\sqrt{1-\delta_{4k}}}\|\eta'\|_2. \tag{6.8}$$

By the definition of d^p and $y = Ax_S + \eta'$, we have

$$x_S - x^p - d^p = (I - A^T A)(x_S - x^p) - A^T \eta'. \tag{6.9}$$

Since $\text{card}(\text{supp}(\tilde{w}) \leq 2k)$ and $\text{card}(\text{supp}(x_S - x^p) \cup \text{supp}(\tilde{w})) \leq 2k$, substituting (6.9)

into (6.8) and applying Lemma 3.3.1 yields

$$\begin{aligned}\|(x_S - x^p - d^p) \otimes \tilde{w})\|_2 &= \left\|\Big((I - A^TA)(x_S - x^p)\Big) \otimes \tilde{w} - (A^T\eta') \otimes \tilde{w}\right\|_2 \\ &\leq \left\|\Big((I - A^TA)(x_S - x^p)\Big)_{\mathrm{supp}(\tilde{w})}\right\|_2 + \left\|(A^T\eta')_{\mathrm{supp}(\tilde{w})}\right\|_2 \\ &\leq \delta_{2k}\|x_S - x^p\|_2 + \sqrt{1+\delta_{2k}}\|\eta'\|_2. \end{aligned} \tag{6.10}$$

Combining (6.8) and (6.10), we have

$$\|x_S - x^p - w^p \otimes d^p\|_2 \leq \delta_{2k}\sqrt{\frac{1+\delta_{2k}}{1-\delta_{4k}}}\|x_S - x^p\|_2 + \frac{3+\delta_{2k}}{\sqrt{1-\delta_{4k}}}\|\eta'\|_2. \tag{6.11}$$

Combining (6.11) and (6.5) yields

$$\left\|x^{p+1} - x_S\right\|_2 \leq \frac{\sqrt{5}+1}{2}\delta_{2k}\sqrt{\frac{1+\delta_{2k}}{1-\delta_{4k}}}\|x_S - x^p\|_2 + \frac{(\sqrt{5}+1)(3+\delta_{2k})}{2\sqrt{1-\delta_{4k}}}\|\eta'\|_2.$$

To guarantee that x^{p+1} is better than x^p to approximate x_S, it is sufficient to require

$$\rho := \frac{\sqrt{5}+1}{2}\delta_{2k}\sqrt{\frac{1+\delta_{2k}}{1-\delta_{4k}}} < 1, \tag{6.12}$$

which indicates

$$\left(\frac{\sqrt{5}+1}{2}\right)^2\delta_{4k}^3 + \left(\frac{\sqrt{5}+1}{2}\right)^2\delta_{4k}^2 + \delta_{4k} - 1 < 0,$$

as a result of $\delta_{2k} \leq \delta_{4k}$. Thus sufficient condition for (6.12) is $\delta_{4k} < 0.4031$, as desired.

In particular, if x is k-sparse and $\eta = 0$, the sequence generated by OPGHT converges to x with

$$\|x^{p+1} - x\|_2 \leq \rho\|x^p - x\|_2 \leq \rho^{p+1}\|x^0 - x\|_2,$$

where ρ is given in (6.12). ■

6.3.2 Analysis of ROPGHT in noisy scenarios

The following lemma is useful to our analysis.

Lemma 6.3.2 (Lemma 2.5 in [142]). Let $\Lambda \subseteq \{1, \ldots, n\}$ be an index set satisfying that $\Lambda \cap \mathrm{supp}(u) = \emptyset$ and $\mathrm{card}(\Lambda \cup \mathrm{supp}(u)) \leq t$. One has

$$\left\| \left(A^T A u \right)_\Lambda \right\|_2 \leq \delta_t \|u\|_2.$$

The main result of ROPGHT is summarized as follows.

Theorem 6.3.3. *Let $y = Ax + \eta$ be the noisy measurements of the signal x, and $q = 2k$. Suppose the restricted isometry constant of the measurement matrix A satisfies*

$$\delta_{4k} < 0.1033.$$

The sequence generated by the ROPGHT algorithm satisfies

$$\left\| x^{p+1} - x_S \right\|_2 \leq \rho \|x^p - x_S\|_2 + \tau \|A x_{\overline{S}} + \eta\|_2,$$

where

$$\rho = \left(\frac{\sqrt{5}+1}{2} \right)^2 \left(\delta_{2k} + \delta_{4k} + \frac{\delta_{2k}\sqrt{1+\delta_{2k}} + 3\delta_{3k}\sqrt{1+\delta_k}}{\sqrt{1-\delta_{2k}}} \right) \tag{6.13}$$

and

$$\tau = \left(\frac{\sqrt{5}+1}{2} \right)^2 \left(\frac{\delta_{2k} + 3\delta_k + 6}{\sqrt{1-\delta_{2k}}} + 2\sqrt{1+\delta_{2k}} \right). \tag{6.14}$$

Proof. Still we denote $S^{p+1} = \mathrm{supp}(x^{p+1})$ and $S^p = \mathrm{supp}(x^p)$. Applying Lemma 5.3.2, we have

$$\begin{aligned}
\left\| x^{p+1} - x_S \right\|_2 &= \left\| \mathcal{H}_k \left(x^p + v^p \right) - x_S \right\|_2 \\
&\leq \frac{\sqrt{5}+1}{2} \left\| \left(x^p + v^p - x_S \right)_{S \cup S^{p+1}} \right\|_2 \\
&\leq \frac{\sqrt{5}+1}{2} \left\| x^p + v^p - x_S \right\|_2 \\
&= \frac{\sqrt{5}+1}{2} \| x^p - x_S + \mathcal{H}_{2k}(w^p \otimes d^p) \|_2 \\
&\leq \left(\frac{\sqrt{5}+1}{2} \right)^2 \left\| \left(x^p - x_S + w^p \otimes d^p \right)_{S \cup S^p \cup \mathrm{supp}(v^p)} \right\|_2 .
\end{aligned} \tag{6.15}$$

Since $S \cup S^p \cup \mathrm{supp}(v^p) = (S \cup S^p) \cup (\mathrm{supp}(v^p) \backslash (S \cup S^p))$, we then can write

$$\begin{aligned}
&\left\| (x^p + d^p \otimes w^p - x_S)_{S \cup S^p \cup \mathrm{supp}(v^p)} \right\|_2 \\
&= \left\| x^p + (d^p \otimes w^p)_{S^p \cup S} - x_S + (d^p \otimes w^p)_{\mathrm{supp}(v^p) \backslash (S \cup S^p)} \right\|_2 \\
&\leq \left\| x^p + (d^p \otimes w^p)_{S^p \cup S} - x_S \right\|_2 + \left\| (d^p \otimes w^p)_{\mathrm{supp}(v^p) \backslash (S \cup S^p)} \right\|_2 , \qquad (6.16)
\end{aligned}$$

where the last inequality follows from the triangle inequality. The first term of (6.16) on the right-hand side can be written as

$$\begin{aligned}
&\left\| x^p + (d^p \otimes w^p)_{S^p \cup S} - x_S \right\|_2 \\
&= \left\| (x^p - x_S) \otimes w^p + (d^p \otimes w^p)_{S^p \cup S} + (x^p - x_S) \otimes (\mathbf{e} - w^p) \right\|_2 \\
&= \left\| (x^p - x_S + d^p)_{S^p \cup S} \otimes w^p + (x^p - x_S) \otimes (\mathbf{e} - w^p) \right\|_2 \\
&\leq \left\| (x^p - x_S + d^p)_{S^p \cup S} \otimes w^p \right\|_2 + \left\| (x^p - x_S) \otimes (\mathbf{e} - w^p) \right\|_2 . \qquad (6.17)
\end{aligned}$$

Combining (6.16) and (6.17) leads to

$$\begin{aligned}
\left\| (x^p + d^p \otimes w^p - x_S)_{S \cup S^p \cup \mathrm{supp}(v^p)} \right\|_2 \leq{}& \left\| (d^p \otimes w^p)_{\mathrm{supp}(v^p) \backslash (S \cup S^p)} \right\|_2 + \left\| (x^p - x_S) \otimes (\mathbf{e} - w^p) \right\|_2 \\
&+ \left\| (x^p - x_S + d^p)_{S^p \cup S} \otimes w^p \right\|_2 . \qquad (6.18)
\end{aligned}$$

We now bound the first term of (6.18) on the right-hand side. Since $0 \leq w^p \leq \mathbf{e}$, one has

$$\left\| (d^p \otimes w^p)_{\mathrm{supp}(v^p) \backslash (S \cup S^p)} \right\|_2 \leq \left\| (d^p)_{\mathrm{supp}(v^p) \backslash (S \cup S^p)} \right\|_2 . \qquad (6.19)$$

Note that $d^p = A^T(y - Ax^p)$, and $y = Ax + \eta = x_S + \eta'$ with $\eta' = Ax_{\overline{S}} + \eta$. Thus (6.19)

can be bounded by

$$\begin{aligned}\left\|(d^p)_{\text{supp}(v^p)\backslash(S\cup S^p)}\right\|_2 &= \left\|\left(A^T(y-Ax^p)\right)_{\text{supp}(v^p)\backslash(S\cup S^p)}\right\|_2 \\ &= \left\|\left(A^T(Ax_S+\eta'-Ax^p)\right)_{\text{supp}(v^p)\backslash(S\cup S^p)}\right\|_2 \\ &\leq \left\|\left(A^TA(x_S-x^p)\right)_{\text{supp}(v^p)\backslash(S\cup S^p)}\right\|_2 + \left\|(A^T\eta')_{\text{supp}(v^p)\backslash(S\cup S^p)}\right\|_2 \\ &\leq \delta_{4k}\left\|x_S-x^p\right\|_2 + \sqrt{1+\delta_{2k}}\|\eta'\|_2, \end{aligned} \tag{6.20}$$

where the last inequality follows from Lemma 6.3.2 due to the fact $\text{card}((\text{supp}(v^p)\backslash(S\cup S^p))\cup\text{supp}(x_S-x^p))\leq 4k$ and Lemma 5.3.1 since $\text{card}(\text{supp}(v^p)\backslash(S\cup S^p))\leq 2k$. Combining (6.19) and (6.20) leads to

$$\left\|(d^p\otimes w^p)_{\text{supp}(v^p)\backslash(S\cup S^p)}\right\|_2 \leq \delta_{4k}\left\|x_S-x^p\right\|_2 + \sqrt{1+\delta_{2k}}\|\eta'\|_2. \tag{6.21}$$

The second term on the right-hand side of (6.18) can be bounded as

$$\begin{aligned}\left\|(x^p-x_S+d^p)_{S^p\cup S}\otimes w^p\right\|_2 &\leq \left\|(x^p-x_S+d^p)_{S^p\cup S}\right\|_2 \\ &\leq \left\|\left[\left(I-A^TA\right)(x^p-x_S)\right]_{S^p\cup S}\right\|_2 + \left\|\left(A^T\eta'\right)_{S^p\cup S}\right\|_2 \\ &\leq \delta_{2k}\left\|x_S-x^p\right\|_2 + \sqrt{1+\delta_{2k}}\left\|\eta'\right\|_2, \end{aligned} \tag{6.22}$$

where the last inequality follows from Lemma 5.3.1 since $\text{card}(\text{supp}(x^p-x_S)\cup(S^p\cup S))\leq 2k$ and $\text{card}(S^p\cup S)\leq 2k$. We now show that the third term of (6.18) can be bounded as

$$\left\|(x_S-x^p)\otimes(e-w^p)\right\|_2 \leq \frac{\delta_{2k}\sqrt{1+\delta_{2k}}+3\delta_{3k}\sqrt{1+\delta_k}}{\sqrt{1-\delta_{2k}}}\left\|x_S-x^p\right\|_2 + \frac{\delta_{2k}+3\delta_k+6}{\sqrt{1-\delta_{2k}}}\|\eta'\|_2. \tag{6.23}$$

Since $y=Ax_S+\eta'$, we can write

$$\begin{aligned}\left\|y-A(x^p+d^p\otimes w^p)\right\|_2 &= \left\|Ax_S-A(x^p+d^p\otimes w^p)+\eta'\right\|_2 \\ &\geq \left\|A(x_S-x^p-d^p\otimes w^p)\right\|_2 - \left\|\eta'\right\|_2. \end{aligned} \tag{6.24}$$

There exists a $2k$-sparse binary vector $\tilde{w}$ such that $\text{supp}\,(x_S - x^p) \subseteq \text{supp}(\tilde{w})$. By the algorithm structure, the vector w^p is the optimal solution to the relaxed optimization problem, thus one has

$$\begin{aligned}\|y - A\,(x^p + d^p \otimes w^p)\|_2 &\leq \|y - A\,(x^p + d^p \otimes \tilde{w})\|_2 \\ &\leq \|A\,(x_S - x^p - d^p \otimes \tilde{w})\|_2 + \|\eta'\|_2 \\ &= \|A\,((x_S - x^p - d^p) \otimes \tilde{w})\|_2 + \|\eta'\|_2\,. \end{aligned} \tag{6.25}$$

Let $x_S - x^p = (x_S - x^p) \otimes w^p + (x_S - x^p) \otimes (\mathbf{e} - w^p)$. We can write

$$\begin{aligned}\|A\,(x_S - x^p - d^p \otimes w^p)\|_2 &= \|A\,((x_S - x^p) \otimes (\mathbf{e} - w^p) + (x_S - x^p) \otimes w^p - d^p \otimes w^p)\|_2 \\ &\geq \|A\,((x_S - x^p) \otimes (\mathbf{e} - w^p))\|_2 - \|A\,((x_S - x^p - d^p) \otimes w^p)\|_2\,. \end{aligned} \tag{6.26}$$

The vector $(x_S - x^p) \otimes (\mathbf{e} - w^p)$ is $2k$-sparse, thus, by the RIP, one has

$$\|A\,((x_S - x^p) \otimes (\mathbf{e} - w^p))\|_2 \geq \sqrt{1 - \delta_{2k}}\,\|(x_S - x^p) \otimes (\mathbf{e} - w^p)\|_2\,. \tag{6.27}$$

Combining (6.24)-(6.27) leads to

$$\begin{aligned}\|(x_S - x^p) \otimes (\mathbf{e} - w^p)\|_2 \leq &\frac{1}{\sqrt{1 - \delta_{2k}}}\,\|A\,((x_S - x^p - d^p) \otimes \tilde{w})\|_2 + \frac{2}{\sqrt{1 - \delta_{2k}}}\,\|\eta'\|_2 \\ &\frac{1}{\sqrt{1 - \delta_{2k}}}\,\|A\,((x_S - x^p - d^p) \otimes w^p)\|_2\,. \end{aligned} \tag{6.28}$$

Note that vector $\tilde{w}$ is a $2k$-sparse binary vector such that $\text{supp}(x_S - x^p) \subseteq \text{supp}(\tilde{w})$. Since

$d^p = A^T(y - Ax^p)$, we can write

$$\begin{aligned}
\|A\left((x_S - x^p - d^p) \otimes \tilde{w}\right)\|_2 &\leq \sqrt{1+\delta_{2k}}\, \|(x_S - x^p - d^p) \otimes \tilde{w}\|_2 \\
&= \sqrt{1+\delta_{2k}} \left(\left\| \left(\left(I - A^T A \right) (x_S - x^p) \right) \otimes \tilde{w} + \left(A^T \eta' \right) \otimes \tilde{w} \right\|_2 \right) \\
&\leq \sqrt{1+\delta_{2k}} \left(\left\| \left(\left(I - A^T A \right) (x_S - x^p) \right)_{\mathrm{supp}(\tilde{w})} \right\|_2 + \left\| \left(A^T \eta' \right)_{\mathrm{supp}(\tilde{w})} \right\|_2 \right) \\
&\leq \delta_{2k} \sqrt{1+\delta_{2k}}\, \|x_S - x^p\|_2 + (1+\delta_{2k})\, \|\eta'\|_2 , \qquad (6.29)
\end{aligned}$$

where the last inequality follows from Lemma 5.3.1 and the fact $\mathrm{card}(\mathrm{supp}(x_S - x^p) \cup \mathrm{supp}(\tilde{w})) \leq 2k$ and $\mathrm{card}(\mathrm{supp}(\tilde{w})) \leq 2k$. Decompose the vector w^p as the way described in Lemma 4.3.6, i.e.,

$$w^p = (w^p)_{\Lambda_1} + \cdots + (w^p)_{\Lambda_{h-1}} + (w^p)_{\Lambda_h} ,$$

where Λ_1 denotes the support set of the first largest k magnitudes of w^p, Λ_2 denotes the support set of the second largest k magnitudes of w^p, and so on. Denote $v^{(i)} = (x_S - x^p - d^p) \otimes (w^p)_{\Lambda_i}$. The last term in (6.28) can be decomposed as

$$\|A\left((x_S - x^p - d^p) \otimes w^p\right)\|_2 = \left\| A \sum_{i=1}^{h} v^{(i)} \right\|_2 \leq \sum_{i=1}^{h} \left\| A v^{(i)} \right\|_2 \leq \sqrt{1+\delta_k} \sum_{i=1}^{h} \left\| v^{(i)} \right\|_2 , \quad (6.30)$$

where the last inequality follows the triangle inequality. The term $\|v^{(i)}\|_2$ can be bounded as

$$\begin{aligned}
\left\| v^{(i)} \right\|_2 &= \left\| \left((x_S - x^p - d^p) \otimes w^p \right)_{\Lambda_i} \right\|_2 \\
&= \left\| \left(\left(\left(I - A^T A \right) (x_S - x^p) \right) \otimes w^p + \left(A^T \eta' \right) \otimes w^p \right)_{\Lambda_i} \right\|_2 \\
&\leq \left\| \left(\left(\left(I - A^T A \right) (x_S - x^p) \right) \otimes w^p \right)_{\Lambda_i} \right\|_2 + \left\| \left(\left(A^T \eta' \right) \otimes w^p \right)_{\Lambda_i} \right\|_2 \\
&\leq \left\| (w^p)_{\Lambda_i} \right\|_\infty \left\| \left(\left(I - A^T A \right) \left((x_S - x^p) \right) \right)_{\Lambda_i} \right\|_2 + \left\| (w^p)_{\Lambda_i} \right\|_\infty \left\| \left(A^T \eta' \right)_{\Lambda_i} \right\|_2 \\
&\leq \left\| (w^p)_{\Lambda_i} \right\|_\infty \delta_{3k} \|x_S - x^p\|_2 + \left\| (w^p)_{\Lambda_i} \right\|_\infty \sqrt{1+\delta_k}\, \|\eta'\|_2 , \qquad (6.31)
\end{aligned}$$

where the last inequality follows from Lemma 5.3.1 as a result of $\text{card}(\text{supp}(x_S - x^p) \cup \Lambda_i) \leq 3k$ and $\text{card}(\Lambda_i) \leq k$. Since $\mathbf{e}^T w^p = 2k$, applying Lemma 4.3.6, we immediately have

$$\|w_{\Lambda_1}\|_\infty + \cdots + \left\|w_{\Lambda_{q-1}}\right\|_\infty + \left\|w_{\Lambda_q}\right\|_\infty \leq 3. \tag{6.32}$$

Combining (6.30)-(6.32) leads to

$$\|A\left((x_S - x^p - d^p) \otimes w^p\right)\|_2 \leq 3\delta_{3k}\sqrt{1+\delta_k}\,\|x_S - x^p\|_2 + 3\left(1+\delta_k\right)\|\eta'\|_2 . \tag{6.33}$$

Substituting (6.29) and (6.33) into (6.28) leads to the relation (6.23). Combining (6.15), (6.18), (6.21)-(6.23), we have that

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho\,\|x_S - x^p\|_2 + \tau\,\|\eta'\|_2 ,$$

where

$$\rho = \left(\frac{\sqrt{5}+1}{2}\right)^2 \left(\delta_{2k} + \delta_{4k} + \frac{\delta_{2k}\sqrt{1+\delta_{2k}} + 3\delta_{3k}\sqrt{1+\delta_k}}{\sqrt{1-\delta_{2k}}}\right) \tag{6.34}$$

and

$$\tau = \left(\frac{\sqrt{5}+1}{2}\right)^2 \left(\frac{\delta_{2k} + 3\delta_k + 6}{\sqrt{1-\delta_{2k}}} + 2\sqrt{1+\delta_{2k}}\right).$$

To guarantee the performance of $\{x^p\}$, it is sufficient to require $\rho < 1$. Due to the fact that $\delta_{2k} \leq \delta_{3k} \leq \delta_{4k}$, the condition $\rho < 1$ is guaranteed by

$$2\left(\frac{\sqrt{5}+1}{2}\right)^2 \delta_{4k} + 4\delta_{4k}\sqrt{\frac{1+\delta_{4k}}{1-\delta_{4k}}} < 1,$$

which is equivalent to

$$(16 + 4\beta^2)\delta_{4k}^3 + (16 - 4\beta^2 - 4\beta)\delta_{4k}^2 + (4\beta + 1) - 1 < 0,$$

where $\beta = \left(\frac{\sqrt{5}+1}{2}\right)^2$. A sufficient condition to guarantee the inequality above is $\delta_{4k} < 0.1033$.

In particular, if x is k-sparse and $\eta = 0$, the sequence $\{x^p\}$ generated by OPGHT satisfies

$$\|x^{p+1} - x\|_2 \leq \rho\|x^p - x\|_2 \leq \rho^{p+1}\|x^0 - x\|_2,$$

where ρ is given as (6.34), which implies that $\{x^p\}$ converges to x when $p \to \infty$. ■

6.3.3 Analysis of ROPGHTP in noisy scenarios

Theorem 6.3.4. *Let $y = Ax + \eta$ be the noisy measurements of the signal x, and $q = 2k$. Suppose that the restricted isometry constant of the measurement matrix A satisfies that*

$$\delta_{3k} < 0.0609.$$

The sequence $\{x^p\}$ generated by the ROPGHTP satisfies

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho^* \left\|x^p - x_S\right\|_2 + \tau^* \left\|Ax_{\overline{S}} + \eta\right\|_2,$$

where

$$\rho^* = \frac{1}{\sqrt{1-\delta_{2k}^2}}\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\delta_{2k} + \delta_{3k} + \frac{\delta_{2k}\sqrt{1+\delta_{2k}} + 3\delta_{3k}\sqrt{1+\delta_k}}{\sqrt{1-\delta_{2k}}}\right) < 1 \tag{6.35}$$

and

$$\tau^* = \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}} + \frac{1}{\sqrt{1-\delta_{2k}^2}}\left(\frac{\sqrt{5}+1}{2}\right)^2\left(\sqrt{1+\delta_k} + \sqrt{1+\delta_{2k}} + \frac{6+3\delta_k+\delta_{2k}}{\sqrt{1-\delta_{2k}}}\right). \tag{6.36}$$

Proof. The ROPGHTP is regarded as a combination of ROPGHT with a pursuit step. Denote by $\overline{x^p}$ the intermediate vector in ROPGHTP was generated by ROPGHT algorithm. Applying Lemma 3.3.8, we have

$$\|x^{p+1} - x_S\|_2 \leq \frac{1}{\sqrt{1-\delta_{2k}^2}}\|x_S - \overline{x^p}\|_2 + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}\|\eta'\|_2. \tag{6.37}$$

We can directly recall the results for ROPGHT, i.e.,

$$\|x_S - \overline{x^p}\|_2 \leq \rho \|x_S - x^p\|_2 + \tau \|\eta'\|_2 , \tag{6.38}$$

where ρ and τ are given as (6.13) and (6.14), respectively. Gathering (6.37) and (6.38) leads to

$$\left\|x^{p+1} - x_S\right\|_2 \leq \rho^* \|x_S - x^p\|_2 + \tau^* \|\eta'\|_2, \tag{6.39}$$

where ρ^* and τ^* are given by (6.35) and (6.36) in the Theorem 6.3.4, respectively. For the guaranteed performance of x_S by RSDHTP, it is sufficient to ensure $\rho^* < 1$, which implies

$$\left(\frac{\sqrt{5}+1}{2}\right)^2 \frac{1}{\sqrt{1-\delta_{4k}^2}} \left(2\delta_{4k} + 4\delta_{4k}\sqrt{\frac{1+\delta_{4k}}{1-\delta_{4k}}}\right) < 1 \tag{6.40}$$

due to the fact that $\delta_{2k} \leq \delta_{3k} \leq \delta_{4k}$. The inequality (6.40) indicates

$$(20\beta^2 + 8\beta + 1)\delta_{4k}^3 + (12\beta^2 - 1)\delta_{4k}^2 - (8\beta + 1)\delta_{4k} + 1 > 0,$$

where $\beta = \left(\frac{\sqrt{5}+1}{2}\right)^2$. Therefore, a sufficient condition to guarantee (6.40) is given by $\delta_{4k} < 0.0609$.

In particular, if the vector x is k-sparse and $\eta = 0$, the relation (6.39) is reduced to

$$\left\|x^{p+1} - x\right\|_2 \leq \rho^* \|x^p - x\|_2 \leq (\rho^*)^{p+1} \left\|x - x^0\right\|_2,$$

where ρ^* is given in (6.35). It implies that the sequence $\{x^p\}$ generated by ROPGHTP converges to x when $p \to \infty$. ■

Prior to performing numerical experiments, we consider the computational complexity of the proposed algorithms.

Remark 6.3.1

Based on Remarks 4.3.3 and 5.3.2, it can be seen that the ROPGHT algorithm requires

about $O\left(mn + q^{3.5}L\right)$ flops, which is comparable to the PGROT algorithm, and the ROPGHTP algorithm requires involves $O\left(mn + q^{3.5}L + k^3\right)$ or $O\left(mn + q^{3.5}L + m^3\right)$ depending on the strategy applied to the pursuit step, where $q > 0$ is the number of elements extracted from the full gradient and L is the size of the problem data encoding in binary.

In Table 6.1, we summarize the computational complexity for each iteration of the main algorithms proposed in the thesis. While an algorithm may have a lower complexity, it may perform worse in practical applications, such as recovering sparse signals and solving sparse optimization problems. The complexity of an algorithm reflects limited knowledge of the algorithm, and the proper choice between different algorithms may depend upon the particular circumstances and specific requirements.

Table 6.1: Complexity each iteration of the main algorithms in the thesis

Algorithm	Complexity
NSIHT	$O\left(mn^2 + n^3\right)$
NSHTP	$O\left(mn^2 + n^3 + k^3\right)$
NTROT	$O\left(mn^2 + n^3 + n^{3.5}L\right)$
NTROTP	$O\left(mn^2 + n^3 + n^{3.5}L + k^3\right)$
PGROT	$O\left(mn + (q+k)^{3.5}L\right)$
PGROTP	$O\left(mn + (q+k)^{3.5}L + k^3\right)$
ROPGHT	$O\left(mn + q^{3.5}L\right)$
ROPGHTP	$O\left(mn + q^{3.5}L + k^3\right)$

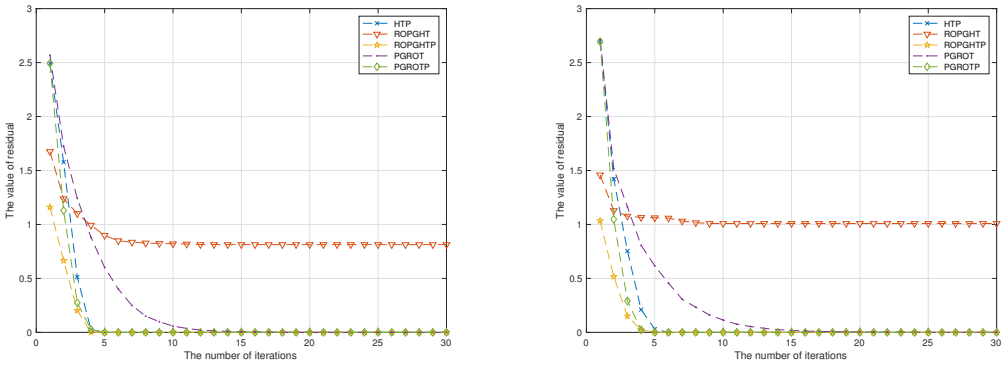
6.4 Numerical Experiments

Experiments were performed to investigate the performance of the proposed algorithms in terms of residual reduction and phase transition for signal reconstruction. The measurement matrices generated in this section are Gaussian random matrices whose entries are i.i.d. and following the normal distribution $\mathcal{N}(0, 1/\sqrt{m})$. Nonzero entries of random sparse

vectors follow the standard normal distribution $\mathcal{N}(0,1)$, and their positions are chosen according to a uniform distribution. Related convex optimization problems are solved by the CVX [66] with solver 'Mosek' [3].

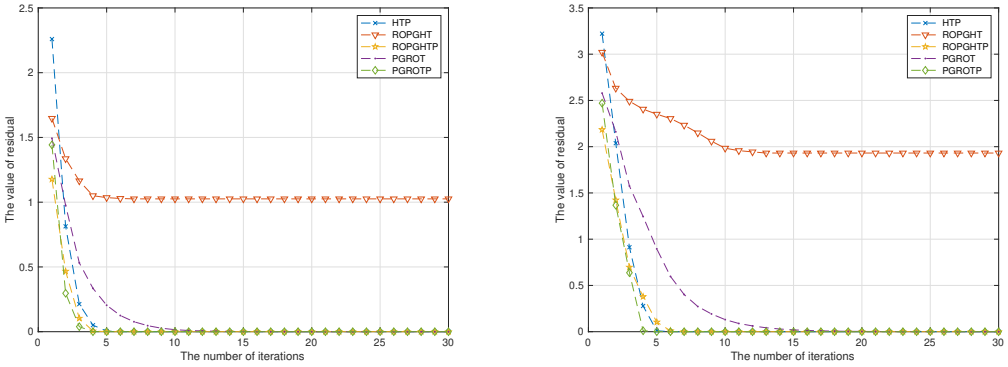
6.4.1 Residual reduction

This experiment compares the residual value of the HTP, ROPGHT, ROPGHTP, PGROT and the PGROTP algorithms in the course of iterations. We set $A \in \mathbb{R}^{256\times 512}$, $y = Ax^*$ and stepsize $\lambda = 1$, and all algorithms start with the initial point $x^0 = 0$. For partial gradient algorithms, we use k-sparse partial gradient directions, i.e., $q = k$ and $q = 2k$, where k is the sparsity level of the generated signal. As illustrated in Fig. 6.1 and 6.2, the ROPGHTP is capable of reducing residual values, and its initial value is smaller than those of the other approaches, whereas the ROPGHT without using pursuit steps is trapped at a local point and thus the residual convergence is not achieved.



(a) $k = 50$ (b) $k = 70$

Figure 6.1: Residual change in the course of iterations between HTP, ROPGHT, ROPGHTP, PGROT and PGROTP algorithms. For partial gradient algorithms, $q = k$.



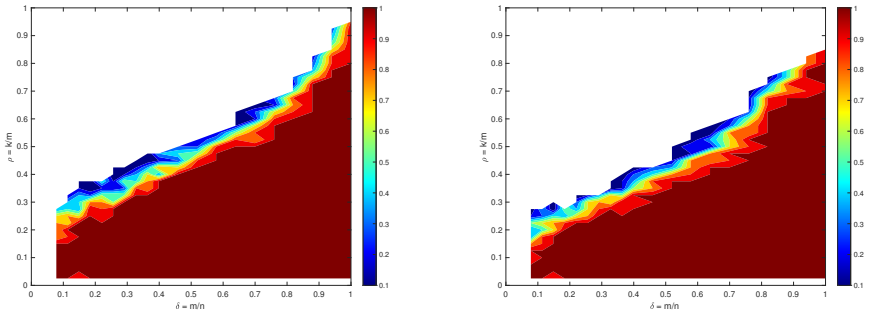
(a) $k = 50$ (b) $k = 70$

Figure 6.2: Residual change in the course of iterations between HTP, ROPGHT, ROPGHTP, PGROT and PGROTP algorithms. For partial gradient algorithms, $q = 2k$.

6.4.2 Phase transition

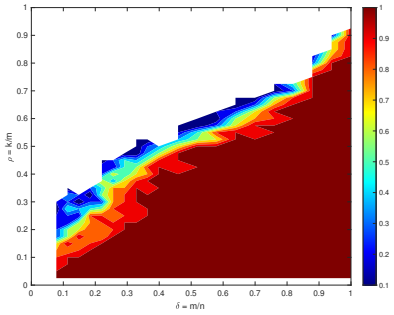
This experiment compares the phase transition phenomenon of HTP and ROPGHTP with different stepsizes in terms of the success rate for sparse signal recovery. Experimental settings are similar to the experiments performed in Section 5.4.4. In this case, we fix the vector length as $n = 512$. The algorithm terminates either when the maximum 30 iterations are reached or the stopping criterion $\|x - x^*\|_2/\|x^*\|_2 \leq 10^{-3}$. At each pair of (m, k), 10 trials are generated to calculate the success rate.

Fig. 6.3 shows the phase transition of HTP and ROPGHTP algorithms, confirming their effectiveness for signal reconstruction, where q is set to be k and $2k$ for ROPGHTP, respectively. It can be seen that HTP is more sensitive to the stepsize λ than ROPGHTP as the shaded area is significantly reduced when λ increases. It appears that the ROPGHTP performs better than HTP when $\lambda = 5$ and 10 under such circumstances.

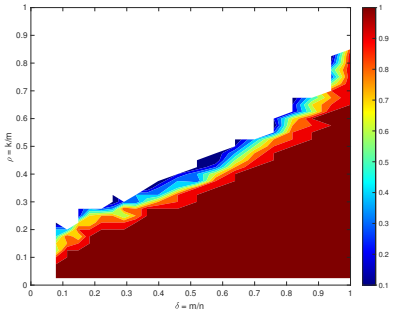


(a) HTP with $\lambda = 1$

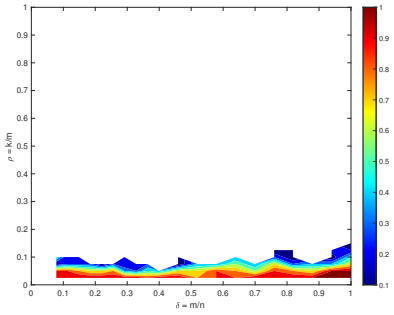
(b) ROPGHTP with $\lambda = 1$



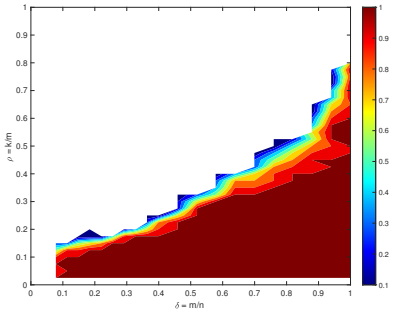
(c) HTP with $\lambda = 2$



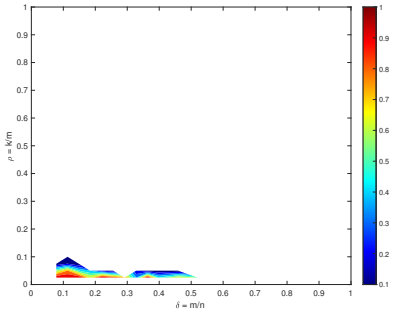
(d) ROPGHTP with $\lambda = 2$



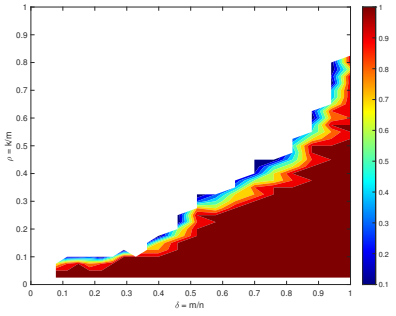
(e) HTP with $\lambda = 5$



(f) ROPGHTP with $\lambda = 5$

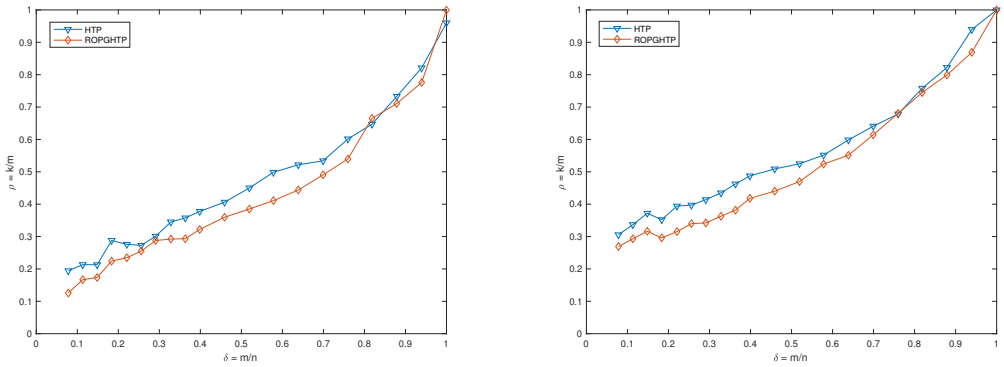


(g) HTP with $\lambda = 10$



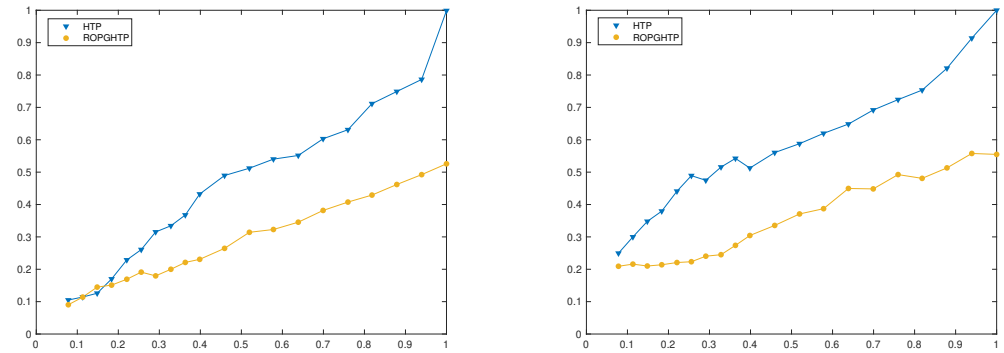
(h) ROPGHTP with $\lambda = 10$

Figure 6.3: Phase transition of HTP and ROPGHTP ($q = k$) with different stepsizes in terms of success rates.



(a) Success rate = 0.9 (b) Success rate = 0.1

Figure 6.4: Phase transition curves of HTP and ROPGHTP ($q = k$) with $\lambda = 1$ in terms of different success rates.



(a) Success rate = 0.9 (b) Success rate = 0.1

Figure 6.5: Phase transition curves of HTP and ROPGHTP ($q = 2k$) with $\lambda = 1$ in terms of different success rates.

Fig. 6.4-6.5 demonstrate the phase transition curves of HTP and ROPGHTP with respect to the success rates 0.9 and 0.1 when $\lambda = 1$. It can be seen from Fig. 6.4 that the curve for ROPGHTP is slightly lower than that of HTP, indicating that using the optimal partial gradient does not contribute to improving the performance of HTP under such circumstances. Moreover, compared with Fig. 6.5, when q is increased to be $2k$, the performance of ROPGHTP becomes even worse. Whether the performance of the proposed algorithms in this chapter can be further improved is an interesting question to address in the near future.

6.5 Conclusion

This chapter investigates the method which combines the so-called optimal partial gradient and hard thresholding. We have proven that the proposed algorithms with $q = 2k$ are guaranteed to recover the sparse signal under the restricted isometry property (RIP). Empirical results show that the ROPGHTP generally works efficient for sparse signal recovery. The algorithms with other choices of q, i.e., $q \neq 2k$, are still worth investigating in the future.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

Four thresholding algorithms are proposed in this thesis by applying novel search directions (Newton-type and partial gradient) to existing hard and optimal thresholding methods. Under the assumption of restricted isometry property, we have shown that the proposed algorithms are guaranteed to reconstruct the sparse signal if the parameters in the algorithms are suitably chosen. Numerical experiments via synthetic data indicate that the proposed algorithms are robust and efficient for signal recovery, and are comparable to several mainstream compressed sensing algorithms.

With respect to the research presented in this thesis, certain potential future research areas merit further examination, excluding those already discussed. Thresholding methods usually adopt a search direction $\mathbf{d}$ from the classic optimization theory, utilize some stepsize λ and specific thresholding operator $\mathcal{T}_k$ to obtain sparse vectors, i.e.,

$$x^{p+1} = \mathcal{T}_k(x^p + \lambda\ \mathbf{d}).$$

Most existing work uses the hard thresholding, optimal k-thresholding or soft thresholding operator and uses gradient, partial gradient, or Newton-type direction as a search direction. The traditional Newton or steepest descent direction is not necessarily the most efficient direction for sparse optimization problems for which many traditional optimization theory and tools are actually lost. Thus an interesting future work might be exploring more efficient thresholding operators and possibly optimal stepsizes or the optimal search

directions in the sparsity-based space or settings.

LIST OF REFERENCES

[1] A. Amini and F. Marvasti. Deterministic construction of binary, bipolar, and ternary compressed sensing matrices. *IEEE Trans. Inf. Theory* 57.4 (2011), pp. 2360–2370.

[2] A. Amini, V. Montazerhodjat, and F. Marvasti. Matrices with small coherence using p-ary block codes. *IEEE Trans. Signal Process.* 60.1 (2011), pp. 172–181.

[3] E. D. Andersen and K. D. Andersen. "The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm". *High performance optimization*. Springer, 2000, pp. 197–232.

[4] L. Applebaum et al. Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery. *Appl. Comput. Harmon. Anal.* 26.2 (2009), pp. 283–290.

[5] A. S. Bandeira et al. Certifying the restricted isometry property is hard. *IEEE Trans. Inf. Theory* 59.6 (2013), pp. 3448–3450.

[6] R. Baraniuk et al. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation* 28.3 (2008), pp. 253–263.

[7] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.* 8.1 (1988), pp. 141–148.

[8] A. Beck. *First-order methods in optimization*. SIAM, 2017.

[9] M. Benzi. Preconditioning techniques for large linear systems: a survey. *Journal of computational Physics* 182.2 (2002), pp. 418–477.

[10] C. R. Berger et al. Application of compressive sensing to sparse channel estimation. *IEEE Communications Magazine* 48.11 (2010), pp. 164–174.

[11] K. Bhatia, P. Jain, and P. Kar. Robust regression via hard thresholding. *arXiv preprint arXiv:1506.02428* (2015).

[12] M. Biswas and H. Om. A new soft-thresholding image denoising method. *Procedia Technol.* 6 (2012), pp. 10–15.

[13] J. D. Blanchard, C. Cartis, and J. Tanner. Compressed sensing: How sharp is the restricted isometry property? *SIAM Rev.* 53.1 (2011), pp. 105–125.

[14] J. D. Blanchard, J. Tanner, and K. Wei. CGIHT: conjugate gradient iterative hard thresholding for compressed sensing and matrix completion. *Information and Inference: A Journal of the IMA* 4.4 (2015), pp. 289–327.

[15] J. D. Blanchard, J. Tanner, and K. Wei. Conjugate gradient iterative hard thresholding: Observed noise stability for compressed sensing. *IEEE Trans. Signal Process.* 63.2 (2014), pp. 528–537.

[16] T. Blumensath. Accelerated iterative hard thresholding. *Signal Process.* 92.3 (2012), pp. 752–756.

[17] T. Blumensath and M. Davies. Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.* 27.3 (2009), pp. 265–274.

[18] T. Blumensath and M. Davies. Iterative thresholding for sparse approximations. *J.Fourier Anal. Appl.* 14.5-6 (2008), pp. 629–654.

[19] T. Blumensath and M. Davies. Normalized iterative hard thresholding: Guaranteed stability and performance. *IEEE J. Sel. Top. Signal Process.* 4.2 (2010), pp. 298–309.

[20] T. Blumensath and M. E. Davies. "A simple, efficient and near optimal algorithm for compressed sensing". *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. 2009, pp. 3357–3360.

[21] H. Boche et al. *Compressed sensing and its applications*. Springer, 2015.

[22] L. Bottou. "Large-scale machine learning with stochastic gradient descent". *Proceedings of COMPSTAT'2010*. Springer, 2010, pp. 177–186.

[23] J.-L. Bouchot, S. Foucart, and P. Hitczenko. Hard thresholding pursuit algorithms: number of iterations. *Appl. Comput. Harmon. Anal.* 41.2 (2016), pp. 412–435.

[24] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[25] J. Cahill, X. Chen, and R. Wang. The gap between the null space property and the restricted isometry property. *Linear Algebra Appl.* 501 (2016), pp. 363–375.

[26] T. T. Cai and L. Wang. Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Trans. Inf. Theory* 57.7 (2011), pp. 4680–4688.

[27] T. Cai and A. Zhang. Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. *IEEE Trans. Inf. Theory* 60.1 (2014), pp. 122–132.

[28] E. Candes and T. Tao. The Dantzig selector: Statistical estimation when p is much larger than n. *The annals of Statistics* 35.6 (2007), pp. 2313–2351.

[29] E. Candès. "Compressive sampling". *Proceedings of the international congress of mathematicians*. Vol. 3. Madrid, Spain. 2006, pp. 1433–1452.

[30] E. Candès. The restricted isometry property and its implications for compressed sensing. *C.R. Math.* 346.9-10 (2008), pp. 589–592.

[31] E. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. Inf. Theory* 51.12 (2005), pp. 4203–4215.

[32] E. Candès, M. Wakin, and S. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *J.Fourier Anal. Appl.* 14.5-6 (2008), pp. 877–905.

[33] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* 52.2 (2006), pp. 489–509.

[34] V. Cevher. "On accelerated hard thresholding methods for sparse approximation". *Wavelets and Sparsity XIV*. Vol. 8138. International Society for Optics and Photonics. 2011, p. 813811.

[35] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision* 40.1 (2011), pp. 120–145.

[36] L.-H. Chang and J.-Y. Wu. An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit. *IEEE Trans. Inf. Theory* 60.9 (2014), pp. 5702–5715.

[37] J. Chen and Q. Gu. "Fast newton hard thresholding pursuit for sparsity constrained nonconvex optimization". *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2017, pp. 757–766.

[38] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM Rev.* 43.1 (2001), pp. 129–159.

[39] A. Cohen, W. Dahmen, and R. DeVore. Compressed sensing and best k-term approximation. *J. Am. Math. Soc.* 22.1 (2009), pp. 211–231.

[40] A. Cohen, W. Dahmen, and R. DeVore. Orthogonal matching pursuit under the restricted isometry property. *Constructive Approximation* 45.1 (2017), pp. 113–127.

[41] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inf. Theory* 55.5 (2009), pp. 2230–2249.

[42] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* 57.11 (2004), pp. 1413–1457.

[43] I. Daubechies et al. Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* 63.1 (2010), pp. 1–38.

[44] M. A. Davenport. “Random observations on random observations: Sparse signal acquisition and processing”. PhD thesis. 2010.

[45] M. A. Davenport and M. B. Wakin. Analysis of orthogonal matching pursuit using the restricted isometry property. *IEEE Trans. Inf. Theory* 56.9 (2010), pp. 4395–4401.

[46] M. E. Davies and R. Gribonval. Restricted Isometry Constants Where ℓ_p Sparse Recovery Can Fail for $0 < p \leq 1$. *IEEE Trans. Inf. Theory* 55.5 (2009), pp. 2203–2214.

[47] G. M. Davis, S. G. Mallat, and Z. Zhang. Adaptive time-frequency decompositions. *Optical engineering* 33.7 (1994), pp. 2183–2191.

[48] R. A. DeVore. Deterministic constructions of compressed sensing matrices. *J. Complexity* 23.4-6 (2007), pp. 918–925.

[49] D. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory* 52.4 (2006), pp. 1289–1306.

[50] D. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* 41.3 (1995), pp. 613–627.

[51] D. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1-minimization. *Proceedings of the National Academy of Sciences* 100.5 (2003), pp. 2197–2202.

[52] D. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inf. Theory* 47.7 (2001), pp. 2845–2862.

[53] D. L. Donoho et al. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inf. Theory* 58.2 (2012), pp. 1094–1121.

[54] M. Elad. *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer Science & Business Media, 2010.

[55] M. Elad and A. M. Bruckstein. A generalized uncertainty principle and sparse representation in pairs of bases. *IEEE Trans. Inf. Theory* 48.9 (2002), pp. 2558–2567.

[56] Y. Eldar and G. Kutyniok. *Compressed sensing: theory and applications*. Cambridge University Press, 2012.

[57] A. Fawzi, M. Davies, and P. Frossard. Dictionary learning for fast classification based on soft-thresholding. *Int. J. Comput. Vision* 114.2 (2015), pp. 306–321.

[58] M. Fornasier and H. Rauhut. Iterative thresholding algorithms. *Appl. Comput. Harmon. Anal.* 25.2 (2008), pp. 187–208.

[59] S. Foucart. A note on guaranteed sparse recovery via ℓ_1-minimization. *Appl. Comput. Harmon. Anal.* 29.1 (2010), pp. 97–103.

[60] S. Foucart. Hard thresholding pursuit: an algorithm for compressive sensing. *SIAM J. Numer. Anal.* 49.6 (2011), pp. 2543–2563.

[61] S. Foucart. "Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants". *Approximation Theory XIII: San Antonio 2010*. Springer, 2010, pp. 65–77.

[62] S. Foucart and M.-J. Lai. Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$. *Appl. Comput. Harmon. Anal.* 26.3 (2009), pp. 395–407.

[63] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Vol. 1. 3. Birkhäuser Basel, 2013.

[64] R. Garg and R. Khandekar. "Gradient descent with sparsification: an iterative algorithm for sparse recovery with restricted isometry property". *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM. 2009, pp. 337–344.

[65] J. H. d. M. Goulart and G. Favier. "An iterative hard thresholding algorithm with improved convergence for low-rank tensor recovery". *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE. 2015, pp. 1701–1705.

[66] M. Grant and S. Boyd. *CVX: Matlab software for disciplined convex programming, version 2.1*. 2014.

[67] L. Grippo, F. Lampariello, and S. Lucidi. A truncated Newton method with nonmonotone line search for unconstrained optimization. *J. Optim. Theory Appl.* 60.3 (1989), pp. 401–419.

[68] T. Hastie, R. Tibshirani, and M. Wainwright. *Statistical learning with sparsity: the lasso and generalizations*. Chapman and Hall/CRC, 2019.

[69] M. R. Hestenes, E. Stiefel, et al. *Methods of conjugate gradients for solving linear systems*. Vol. 49. 1. 1952.

[70] P. W. Holland and R. E. Welsch. Robust regression using iteratively reweighted least-squares. *Communications in Statistics-theory and Methods* 6.9 (1977), pp. 813–827.

[71] R. Horn and C. Johnson. *Matrix analysis*. Cambridge university press, 2012.

[72] S. D. Howard, A. R. Calderbank, and S. J. Searle. "A fast reconstruction algorithm for deterministic compressive sensing using second order Reed-Muller codes". *2008 42nd Annual Conference on Information Sciences and Systems*. IEEE. 2008, pp. 11–15.

[73] M. Jing, X. Zhou, and C. Qi. Quasi-Newton iterative projection algorithm for sparse recovery. *Neurocomputing* 144 (2014), pp. 169–173.

[74] A. Khare and U. S. Tiwary. Soft-thresholding for denoising of medical images—a multiresolution approach. *International Journal of Wavelets, Multiresolution and Information Processing* 3.04 (2005), pp. 477–496.

[75] B. Li et al. Sufficient conditions for generalized orthogonal matching pursuit in noisy case. *Signal Process.* 108 (2015), pp. 111–123.

[76] S. Li and G. Ge. Deterministic sensing matrices arising from near orthogonal systems. *IEEE Trans. Inf. Theory* 60.4 (2014), pp. 2291–2302.

[77] C. Liu, Y. Fang, and J. Liu. Some new results about sufficient conditions for exact support recovery of sparse signals via orthogonal matching pursuit. *IEEE Trans. Signal Process.* 65.17 (2017), pp. 4511–4524.

[78] H. Liu and J. Peng. Sparse signal recovery via alternating projection method. *Signal Process.* 143 (2018), pp. 161–170.

[79] H. Liu, J. Peng, and Z. Lin. A theoretical result of sparse signal recovery via alternating projection method. *Information Sciences* 506 (2020), pp. 51–57.

[80] H. Liu et al. "Image denoising via adaptive soft-thresholding based on non-local samples". *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 484–492.

[81] H. Liu and R. Foygel Barber. Between hard and soft thresholding: optimal iterative thresholding algorithms. *Information and Inference: A Journal of the IMA* 9.4 (2020), pp. 899–933.

[82] Y. Liu et al. Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging. *IEEE Trans. Med. Imaging* 35.9 (2016), pp. 2130–2140.

[83] D. Loghin. A note on constraint preconditioning. *SIAM Journal on Matrix Analysis and Applications* 38.4 (2017), pp. 1486–1495.

[84] N. Meng and Y.-B. Zhao. Newton-step-based hard thresholding algorithms for sparse signal recovery. *IEEE Trans. Signal Process.* 68 (2020), pp. 6594–6606.

[85] N. Meng and Y.-B. Zhao. Newton-Type Optimal Thresholding Algorithms for Sparse Optimization Problems. *Journal of the Operations Research Society of China* (2022), pp. 1–23.

[86] N. Meng et al. Partial gradient optimal thresholding algorithms for a class of sparse optimization problems. *Journal of Global Optimization* (2022), pp. 1–21.

[87] Q. Mo and S. Li. New bounds on the restricted isometry constant $\delta_2 k$. *Appl. Comput. Harmon. Anal.* 31.3 (2011), pp. 460–468.

[88] Q. Mo and Y. Shen. A remark on the restricted isometry property in orthogonal matching pursuit. *IEEE Trans. Inf. Theory* 58.6 (2012), pp. 3654–3656.

[89] D. Needell and J. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* 26.3 (2009), pp. 301–321.

[90] D. Needell and R. Vershynin. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE J. Sel. Top. Signal Process.* 4.2 (2010), pp. 310–316.

[91] D. Needell and R. Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Found. Comput. Math.* 9.3 (2009), pp. 317–334.

[92] T. L. Nguyen and Y. Shin. Deterministic sensing matrices in compressive sensing: a survey. *The Scientific World Journal* 2013 (2013).

[93] J. Nocedal. Theory of algorithms for unconstrained optimization. *Acta Numer.* 1.1 (1992), pp. 199–242.

[94] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[95] V. Patel and R. Chellappa. "Sparse representations, compressive sensing and dictionaries for pattern recognition". *The first Asian conference on pattern recognition*. IEEE. 2011, pp. 325–329.

[96] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition". *Proceedings of 27th Asilomar conference on signals, systems and computers*. IEEE. 1993, pp. 40–44.

[97] M. Powell. Direct search algorithms for optimization calculations. *Acta Numer.* (1998), pp. 287–336.

[98] S. Qaisar et al. Compressive sensing: From theory to applications, a survey. *J. Commun. Networks* 15.5 (2013), pp. 443–456.

[99] K. Qiu and A. Dogandzic. Sparse signal reconstruction via ECME hard thresholding. *IEEE Trans. Signal Process.* 60.9 (2012), pp. 4551–4569.

[100] H. Rauhut, R. Schneider, and Ž. Stojanac. Low rank tensor recovery via iterative hard thresholding. *Linear Algebra Appl.* 523 (2017), pp. 220–262.

[101] R. Rockafellar. *Convex analysis*. Princeton university press, 2015.

[102] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE* 37.1 (1949), pp. 10–21.

[103] J. Shen and P. Li. "On the iteration complexity of support recovery via hard thresholding pursuit". *International Conference on Machine Learning*. PMLR. 2017, pp. 3115–3124.

[104] Z.-J. Shi and J. Shen. New inexact line search method for unconstrained optimization. *J. Optim. Theory Appl.* 127.2 (2005), pp. 425–446.

[105] C.-B. Song, S.-T. Xia, and X.-J. Liu. Improved analysis for subspace pursuit algorithm in terms of restricted isometry constant. *IEEE Signal Process Lett.* 21.11 (2014), pp. 1365–1369.

[106] T. Steihaug. The conjugate gradient method and trust regions in large scale optimization. *SIAM J. Numer. Anal.* 20.3 (1983), pp. 626–637.

[107] W. Sun, J. Han, and J. Sun. Global convergence of nonmonotone descent methods for unconstrained optimization problems. *J. Comput. Appl. Math.* 146.1 (2002), pp. 89–98.

[108] J. Tanner and K. Wei. Normalized iterative hard thresholding for matrix completion. *SIAM J. Sci. Comput.* 35.5 (2013), S104–S125.

[109] R. Tibshirani. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Series B (Methodological)* 58.1 (1996), pp. 267–288.

[110] A. M. Tillmann and M. E. Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Trans. Inf. Theory* 60.2 (2014), pp. 1248–1259.

[111] J. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inf. Theory* 50.10 (2004), pp. 2231–2242.

[112] J. Tropp and A. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inf. Theory* 53.12 (2007), pp. 4655–4666.

[113] P. Tseng et al. A simple polynomial-time algorithm for convex quadratic programming (1988).

[114] J. Wang, S. Kwon, and B. Shim. Generalized orthogonal matching pursuit. *IEEE Trans. Signal Process.* 60.12 (2012), pp. 6202–6216.

[115] J. Wang et al. Recovery of sparse signals via generalized orthogonal matching pursuit: A new analysis. *IEEE Trans. Signal Process.* 64.4 (2015), pp. 1076–1089.

[116] N. Wang and Y. Wang. "An image reconstruction algorithm based on compressed sensing using conjugate gradient". *2010 4th International Universal Communication Symposium*. IEEE. 2010, pp. 374–377.

[117] Q. Wang and G. Qu. A new greedy algorithm for sparse recovery. *Neurocomputing* 275 (2018), pp. 137–143.

[118] J. Weed. Approximately certifying the restricted isometry property is hard. *IEEE Trans. Inf. Theory* 64.8 (2018), pp. 5488–5497.

[119] L. Welch. Lower bounds on the maximum cross correlation of signals (corresp.) *IEEE Trans. Inf. Theory* 20.3 (1974), pp. 397–399.

[120] L. Xiao and T. Zhang. A proximal-gradient homotopy method for the sparse least-squares problem. *SIAM J. Optim.* 23.2 (2013), pp. 1062–1091.

[121] J. Xu and Y.-B. Zhao. Stability analysis of a class of sparse optimization problems. *Optim. Methods Software* 35.4 (2020), pp. 836–854.

[122] X.-S. Yang. *Introduction to algorithms for data mining and machine learning*. Academic press, 2019.

[123] X.-T. Yuan, P. Li, and T. Zhang. "Exact recovery of hard thresholding pursuit". *Advances in Neural Information Processing Systems*. 2016, pp. 3558–3566.

[124] X.-T. Yuan, P. Li, and T. Zhang. Gradient hard thresholding pursuit. *J. Mach. Learn. Res.* 18.1 (2017), pp. 6027–6069.

[125] X.-T. Yuan and Q. Liu. "Newton greedy pursuit: A quadratic approximation method for sparsity-constrained optimization". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 4122–4129.

[126] X.-T. Yuan and Q. Liu. Newton-Type greedy selection methods for ℓ_0-constrained minimization. *IEEE Trans. Pattern Anal. Mach. Intell.* 39.12 (2017), pp. 2437–2450.

[127] Y. Yuan and W. Sun. *Optimization theory and methods*. Science Press, Beijing, 1997.

[128] H. Zhang and W. W. Hager. A nonmonotone line search technique and its application to unconstrained optimization. *SIAM J. Optim.* 14.4 (2004), pp. 1043–1056.

[129] T. Zhang. Sparse recovery with orthogonal matching pursuit under RIP. *IEEE Trans. Inf. Theory* 57.9 (2011), pp. 6215–6221.

[130] X. Zhang et al. "Robust Regression via Heuristic Hard Thresholding." *IJCAI*. 2017, pp. 3434–3440.

[131] Y. Zhang. *A simple proof for recoverability of ℓ_1-minimization (ii): the nonnegativity case*. Tech. rep. 2005.

[132] Y. Zhang. Theory of compressive sensing via ℓ_1-minimization: A non-RIP analysis and extensions. *Journal of the Operations Research Society of China* 1.1 (2013), pp. 79–105.

[133] Y.-B. Zhao. Equivalence and strong equivalence between the sparsest and least ℓ_1-norm nonnegative solutions of linear systems and their applications. *Journal of the Operations Research Society of China* 2.2 (2014), pp. 171–193.

[134] Y.-B. Zhao. Optimal k-thresholding algorithms for sparse optimization problems. *SIAM J. Optim.* 30.1 (2020), pp. 31–55.

[135] Y.-B. Zhao. RSP-based analysis for sparsest and least ℓ_1-norm solutions to underdetermined linear systems. *IEEE Trans. Signal Process.* 61.22 (2013), pp. 5777–5788.

[136] Y.-B. Zhao. *Sparse Optimization Theory and Methods*. CRC Press, 2018.

[137] Y.-B. Zhao and M. Kočvara. A new computational method for the sparsest solutions to systems of linear equations. *SIAM J. Optim.* 25.2 (2015), pp. 1110–1134.

[138] Y.-B. Zhao and D. Li. Reweighted ℓ_1-minimization for sparse solutions to underdetermined linear systems. *SIAM J. Optim.* 22.3 (2012), pp. 1065–1088.

[139] Y.-B. Zhao and Z.-Q. Luo. Analysis of optimal thresholding algorithms for compressed sensing. *Signal Process.* 187 (2021), p. 108148.

[140] Y.-B. Zhao and Z.-Q. Luo. Constructing new weighted ℓ_1-algorithms for the sparsest points of polyhedral sets. *Math. Oper. Res.* 42.1 (2017), pp. 57–76.

[141] Y.-B. Zhao and Z.-Q. Luo. Dynamic Orthogonal Matching Pursuit for Signal Reconstruction. *arXiv preprint arXiv:2108.09892* (2021).

[142] Y.-B. Zhao and Z.-Q. Luo. Improved RIP-based bounds for guaranteed performance of several compressed sensing algorithms. *arXiv preprint arXiv:2007.01451* (2020).

[143] P. Zhou, X.-T. Yuan, and J. Feng. "Efficient stochastic gradient hard thresholding". *Advances in Neural Information Processing Systems*. 2018, pp. 1984–1993.

[144] S. Zhou, L. Pan, and N. Xiu. Subspace Newton method for the ℓ_0-regularized optimization. *arXiv preprint arXiv:2004.05132* (2020).

[145] S. Zhou, N. Xiu, and H.-D. Qi. Global and Quadratic Convergence of Newton Hard-Thresholding Pursuit. *J. Mach. Learn. Res.* 22.12 (2021), pp. 1–45.