Quantitative analysis of algorithms for compressed signal recovery

Andrew J Thompson

Doctor of Philosophy
University of Edinburgh
2012

Abstract

Compressed Sensing (CS) is an emerging paradigm in which signals are recovered from undersampled nonadaptive linear measurements taken at a rate proportional to the signal's true information content as opposed to its ambient dimension. The resulting problem consists in finding a sparse solution to an underdetermined system of linear equations. It has now been established, both theoretically and empirically, that certain optimization algorithms are able to solve such problems. Iterative Hard Thresholding (IHT) (Blumensath and Davies, 2007), which is the focus of this thesis, is an established CS recovery algorithm which is known to be effective in practice, both in terms of recovery performance and computational efficiency. However, theoretical analysis of IHT to date suffers from two drawbacks: state-of-the-art worst-case recovery conditions have not yet been quantified in terms of the sparsity/undersampling trade-off, and also there is a need for average-case analysis in order to understand the behaviour of the algorithm in practice.

In this thesis, we present a new recovery analysis of IHT, which considers the fixed points of the algorithm. In the context of arbitrary matrices, we derive a condition guaranteeing convergence of IHT to a fixed point, and a condition guaranteeing that all fixed points are 'close' to the underlying signal. If both conditions are satisfied, signal recovery is therefore guaranteed. Next, we analyse these conditions in the case of Gaussian measurement matrices, exploiting the realistic average-case assumption that the underlying signal and measurement matrix are independent. We obtain asymptotic phase transitions in a proportional-dimensional framework, quantifying the sparsity/undersampling trade-off for which recovery is guaranteed. By generalizing the notion of fixed points, we extend our analysis to the variable stepsize Normalised IHT (NIHT) (Blumensath and Davies, 2010). For both stepsize schemes, comparison with previous results within this framework shows a substantial quantitative improvement.

We also extend our analysis to a related algorithm which exploits the assumption that the underlying signal exhibits tree-structured sparsity in a wavelet basis (Baraniuk et al., 2010). We obtain recovery conditions for Gaussian matrices in a simplified proportional-dimensional asymptotic, deriving bounds on the oversampling rate relative to the sparsity for which recovery is guaranteed. Our results, which are the first in the phase transition framework for tree-based CS, show a further significant improvement over results for the standard sparsity model. We also propose a dynamic programming algorithm which is guaranteed to compute an exact tree projection in low-order polynomial time.

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Andrew J Thompson*)

Acknowledgments

I would like to express my gratitude for the help and support I have received during my PhD and the writing of this thesis:

First and foremost, to my supervisor Coralia Cartis, for giving so much of her time and expertise in order to guide my research, and also for her patience in teaching me how to write academically. Thanks also go to my second supervisor Jared Tanner for the interest he took in my professional development, and for many illuminating discussions. I would like to thank both my supervisors, and also Jeff Blanchard, for giving me the valuable opportunity of contributing to research projects in the early part of my PhD.

I benefited greatly from involvement in the Edinburgh Compressed Sensing (E-CoS) group, whose regular reading groups helped in broadening my knowledge of the literature. The group was also the occasion for many helpful discussions, most notably with Mike Davies concerning the IHT algorithm, and with Gabriel Rilling concerning the exact tree projection problem.

My gratitude also goes to several of my PhD colleagues for their friendship and encouragement. I shared an office throughout the entirety of my PhD with Stuart Murray, and the humourous wordplays we shared were a source of much entertainment, and perhaps the reason that the turnover of other users of our office was so mysteriously high! I also appreciated many good times spent with Bubacarr Bah and Ke Wei, who shared with me the ups and down of the strange but unforgettable journey that is a mathematics PhD.

I would also like to thank my parents for their personal support during my PhD, my flatmates and friends who enriched my life in Edinburgh in so many ways, and the Morningside Baptist Church community for their great encouragement. Thank you to everyone!

Contents

Chapter 1

Introduction

1.1 Compressed Sensing: an overview

The last fifty years have witnessed an information revolution, namely an explosion in mankind's ability to sample, store, process and transmit information in ever-increasing volumes. The catalyst for this transformation was the pioneering work of Nyquist and Shannon [131, 139] on the digitization of analog signals, which showed that any band-limited signal can be exactly recovered from a set of finitely spaced samples, and much signal processing has now switched to the digital domain. However, in many current signal processing applications, the traditional approach involves the acquisition of extremely high-dimensional data which can be prohibitively expensive, either in terms of time or money. Furthermore, the volume of data acquired may be so large that it exceeds storage or transmission requirements [6], and subsequently needs to be compressed. Compression techniques, such as transform coding, are motivated by the observation that many signals have inherent simplicity structure, so that their true information content is much lower than their dimension might suggest. However, there seems to be something intuitively wasteful about the traditional approach of maximal sampling followed by compression, in which a full set of measurements are obtained, only for the majority to be subsequently discarded. If the underlying signal really is compressible, is it not possible to take less measurements in the first place and build the compression into the measurement process?

The theory of Compressed Sensing (CS) directly addresses this issue, proposing new sensing protocols in which it is possible to recover compressible signals from a reduced number of appropriately chosen nonadaptive measurements. In fact, central to CS is the claim that the number of measurements need only be proportional to the information content of the signal, rather than its dimension. Having obtained the undersampled measurements, the CS approach to recovering the original signal is to use optimization formulations and techniques to find a sparse, approximate solution to the underdetermined system of equations arising from the measurements. At the heart of compressed sensing is the discovery that, for appropriately

chosen measurements, a large number of algorithms succeed in solving this seemingly intractable problem to high accuracy, hence recovering a good approximation to the original signal.

There already exist a wide range of applications of CS theory, spread over areas as diverse as astronomy, medicine, defence technology and communications. CS has opened up new approaches to digitizing analog signals, allowing signals of high bandwidth to be sampled more efficiently [126]. In medical imaging, CS techniques have been used in magnetic resonance imaging (MRI) where it has been shown possible to reduce the amount of time a patient needs to remain in the MRI scanner [120]. Defence technology is another promising area, where storage and transmission capacities on board aircraft are often limited, and the use of CS in Synthetic Aperture Radar (SAR) imaging is already being explored [134]. Seismic imaging, where significant undersampling of signals is unavoidable, is also benefiting greatly from the recent insights of CS [119]. The Herschel telescope launched in 2009 with a camera that implements CS encoding, with the aim of provided more accurate compressed images that can be transmitted back to earth [28]. In communications theory, a wide range of applications of CS have been proposed, and in particular new approaches to sparse channel estimation have been developed [14]. Meanwhile, CS theory is even motivating entirely new technology, such as the single-pixel camera [82], and this particular application is the focus of Section 1.11.

Outline of the chapter. The aim of the present chapter is to set the original work presented in later chapters of the thesis within the context of preceding and related research. While CS itself is less than a decade old, its roots go back much further, and we begin in Section 1.2 by tracing this history. After establishing a notational framework that will be used throughout the thesis in Section 1.3, we next describe some of the main results in CS in Section 1.4. Section 1.5 introduces the phase transition framework as a means of addressing the vital question of how to quantify recovery results in CS. We then present a survey of the various recovery algorithms which have been proposed for CS in Section 1.6, before introducing the algorithm which will largely be the subject of analysis in this thesis, Iterative Hard Thresholding (IHT) [26], in Section 1.7. This thesis also considers an extension of IHT to a *tree-based* signal model: we review the current literature on refined sparsity models for CS in Section 1.8, before introducing the tree-based signal model and the Iterative Tree Projection (ITP) algorithm [24, 7] in Section 1.9. We present the outline and contributions of the thesis in Section 1.10, before illustrating some of the key points of the chapter by means of a case study on the single pixel camera [82] in Section 1.11.

1.2 Before CS: a brief survey

At its heart, compressed sensing involves finding sparse solutions to underdetermined systems of linear equations, and its roots therefore lie in underdetermined inverse problems and sparse

modelling. Central to CS are deep connections between these two areas.

Inverse problems. An *inverse problem* refers to any problem in which it is required to convert observed measurements of a system into information about the system itself. Inverse problems proliferate in many scientific applications, including geophysics, medical imaging, tomography, remote sensing and astronomy. They have been extensively studied ever since a seminal paper by the physicist Ambarzumyan in 1929 [3]. Compressed sensing, and therefore this thesis, is especially concerned with discrete, linear inverse problems, where the task is to determine an unknown vector $x^* \in \mathbb{R}^N$ from the measurements

$$b = Ax^* + e, \tag{1.1}$$

where $b \in \mathbb{R}^n$ is a vector of measurements, $A \in \mathbb{R}^{n \times N}$ is the measurement matrix, and where e is a noise vector.

A *well-posed* problem was defined by Hadamard [105] as one for which there exists a unique solution which is stable to small perturbations in the data. However, many inverse problems are *ill-posed* as they are underdetermined with $n < N$, due to limitations on the number of samples that can be collected, meaning that inferences are required from a relatively small data sample.

Ill-posed inverse problems can often be transformed into well-posed problems by seeking a solution that is consistent not only with the data but also with prior notions of the likely behaviour of the system under exploration. Different strategies for implementing this principle have been developed, with perhaps the most well-known being the method of regularization, first proposed by Tikhonov and Arsenin in 1977 [152], in which the approach is to solve a reformulated optimization problem with an added penalty term.

Sparse inverse problems. Interest in sparse modelling has been gradually developing over many years, especially in the context of sparse regression in statistics. The task in sparse regression is to identify the most influential variables of a model, which generally leads to a well-posed inverse problem, and an early example is found in the work of Garside [97] in 1965 on subset selection. Sparse solutions to inverse problems have also been historically sought in certain scientific disciplines such as medical ultrasound [133], but most notably by geophysicists working in reflection seismology, in which properties of the Earth's subsurface are estimated from reflected seismic waves. An early example is a paper by Claerbout and Muir in 1973 on the recovery of wide-band signals from narrow-band data [53] which, once discretized, amounts to finding a sparse solution to an underdetermined discrete linear inverse problem such as (1.1).

The interest in sparsity naturally fuelled the development of algorithms for sparse solutions to linear systems. In particular, the sparsity-enforcing properties of the l_1-norm, defined as

the sum of the absolute values of the components of a vector, have long been recognized. For example, sparse solutions are obtained in [53] by solving an optimization problem involving the minimization of the l_1-norm. The authors recast the problem as a linear program, and methods for solving linear programs such as the simplex method have existed since the 1940s [57].

Early connections between sparsity, undersampling and uncertainty principles. Levy and Fullagar in 1981 [117] performed extensive testing which demonstrated that accurate recovery of seismic signals is possible from significantly undersampled measurements using an l_1 approach. As a clear phenomenon began to be observed numerically, it was not long before the first theoretical results appeared. The earliest recovery guarantee for l_1-minimization was given by Santosa and Symes in 1983 [138], where it was shown that a signal consisting of a small number of spikes can be recovered from undersampled Fourier measurements, provided the number of spikes multiplied by the number of missing measurements does not exceed half the dimension of the signal, which is equivalent to requiring that the number of measurements n satisfies

$$n \geq N\left(1 - \frac{1}{2k}\right),$$

where k is the number of spikes (sparsity) of the underlying signal.

In 1989, Donoho and Stark [74] first pointed to the essential question of the properties of the measurement matrix, showing that the result in [138] relies upon an *uncertainty principle* satisfied by the Fourier transform. Namely, a signal and its Fourier transform cannot both be arbitrarily sparse. The need for an uncertainty principle is intuitive: we would like the undersampled measurements to capture sufficient information to recover a sparse signal, and so it seems natural to require that its measurements can be distinguished from those of any other sparse signal. But if the Fourier transforms of two sparse signals agree, then the Fourier transform of the difference between the two signals would vanish and hence also be sparse, thus violating the uncertainty principle. The authors also conjectured that much stronger uncertainty principles might hold if the spike positions were scattered at random, a key notion that would later become an integral part of CS theory.

Sparse approximation over redundant bases. In the 1990s, there was a growing interest in the signal processing community around finding sparse representations over redundant bases (in which the number of basis elements is greater than the signal dimension). This could either arise due to the use of overcomplete representations such as Gabor frames or wavelets, or due to several orthogonal bases being concatenated together. Being yet another example of an underdetermined system of linear equations, some strategy is required to select a signal representation from a potentially infinite number of possibilities. Following several proposed solutions which do not promote sparsity, Mallat and Zhang [124] proposed the Matching Pursuit strategy in 1993 which builds up a sequence of sparse approximations in a stepwise, 'greedy'

fashion. The authors also pointed to a particular uncertainty principle, namely the *incoherence* of basis vectors, as playing a crucial role in the success or otherwise of the sparse approximation. This was soon followed by the better-performing but more computationally expensive greedy algorithm, Orthogonal Matching Pursuit (OMP), proposed in [62].

In 1999, Chen, Donoho and Saunders proposed minimizing the l_1-norm to obtain sparse approximations from redundant bases, an approach which they termed Basis Pursuit (BP) [52]. More precisely, the BP criterion is to solve the optimization problem

$$\min_{x \in \mathbb{R}^N} \|x\|_1 \quad \text{subject to} \quad Ax = b. \tag{1.2}$$

The authors also proposed a variant called Basis Pursuit De-Noising, which is designed to carry out sparse approximation in a noisy context, by solving a least-squares problem with an l_1 penalty.

Coherence-based recovery guarantees. Theoretical results for Basis Pursuit then began to follow: first Donoho and Huo [71] showed in 2001 that sparse representations from a concatenation of two orthonormal bases are exactly recovered by Basis Pursuit provided the *mutual coherence* between the bases is sufficiently small. This was soon followed in 2003 by results for more general overcomplete representations based on the related concept of *coherence*, by Donoho and Elad [70], and by Gribonval and Nielson [102]. Given a measurement matrix A, its coherence $\mu(A)$ is defined to be the maximum absolute Euclidean inner product between any two distinct columns of A. Writing A_i for the ith column of A, we therefore define

$$\mu(A) := \max_{i \neq j} \left| A_i^T A_j \right|. \tag{1.3}$$

The authors of [70, 102] give conditions which ensure that the l_1-problem (1.2) shares exactly the same solution as the problem

$$\min_{x \in \mathbb{R}^N} \|x\|_0 \quad \text{subject to} \quad Ax = b, \tag{1.4}$$

where $\|x\|_0$ denotes the l_0 pseudo-norm which simply counts the nonzero coefficients of x. Such a phenomenon is often referred to as l_1-l_0 *equivalence*, which we next formalize.

Definition 1.1 (l_1-l_0 **equivalence**). *Suppose $b = Ax^*$. We say l_1-l_0 equivalence holds if (1.4) and (1.2) share the same solution.*

It is shown in both [70] and [102] that l_1-l_0 equivalence holds provided

$$\mu(A) < \frac{1}{2k-1}. \tag{1.5}$$

At around the same time, Gilbert et al. [98] proved a coherence-based result for OMP, which was

subsequently significantly improved by Tropp [153]. Since many representations yield $\mu(A) = \mathcal{O}(1/\sqrt{n})$, such coherence results imply that $n = \mathcal{O}(k^2)$ measurements are sufficient to recover a k-sparse signal. However, numerical tests by Donoho and Huo led them to conjecture that their coherence-related results "point the way to a phenomenon valid under far less restrictive sparsity assumptions" [71]. In particular, the breakthrough of compressed sensing was to establish that measurement schemes exist for which the number of measurements need only be proportional to the sparsity of the underlying signal.

1.3 Linear measurements of compressible signals

Before describing some of the key results in compressed sensing, we establish some notation and present in more detail a well-known framework of linear measurements of compressible signals.

Linear measurements. We consider the problem of recovering a discrete or digitized signal $z^* \in \mathbb{R}^N$ (at a given time if it is time-varying) from a set of linear measurements. Suppose we obtain $n < N$ measurements as noisy Euclidean inner products of z^* with a series of fixed vectors $\{\phi_1, \phi_2, \ldots, \phi_n\}$, so that

$$b_i = \phi_i^T z^* + e_i \quad \text{for} \quad i = 1, 2, \ldots, n,$$

where e_i models additive noise. Then the measurement process may be represented by the matrix equation

$$b = \Phi z^* + e, \tag{1.6}$$

where Φ is the matrix with $\phi_1^T, \phi_2^T, \ldots, \phi_n^T$ as its rows.

Compressible signals. We assume that the signal z^* has low information content relative to its dimension, in the sense that there exists some basis in which the signal's information is largely captured by relatively few of its coefficients. This means that z^* is well-approximated by a sparse vector in some basis. Suppose $\{\psi_1, \psi_2, \ldots, \psi_N\}$ is such a basis, and let x^* be the coefficients with respect to this basis, so that

$$z^* = \sum_{i=1}^{N} x_i^* \psi_i,$$

which we can equally write in matrix notation as

$$z^* = \Psi x^*, \tag{1.7}$$

where Ψ is the $N \times N$ matrix with $\psi_1, \psi_2, \ldots, \psi_N$ as its columns. We consider both exact and approximate sparse models: either x^* is *k-sparse* for some $k < N$, meaning that it has exactly k nonzeros; or x^* is *k-compressible*, meaning that it is well-approximated by a k-sparse vector. The former is often a helpful simplification in the context of theoretical analysis, while the latter is the more likely in practical applications.

The model. Defining $A := \Phi\Psi$, we may combine (1.6) and (1.7) and write instead

$$b = Ax^* + e, \tag{1.8}$$

where we refer to the $n \times N$ matrix A from now on as the measurement matrix. Recalling (1.1), we clearly have an ill-posed discrete linear inverse problem, or equivalently a system of underdetermined linear equations. But in addition, we are assuming that the linear system has an underlying (exact or approximate) sparse solution. We formalize three possible variants of this problem below.

Problem 1.1 (**Sparse recovery from exact measurements**). *Recover exactly a k-sparse $x^* \in \mathbb{R}^N$ from the noiseless measurements $b = Ax^* \in \mathbb{R}^n$, where $2k \leq n \leq N$.*

Problem 1.2 (**Sparse recovery from noisy measurements**). *Recover an approximation to a k-sparse $x^* \in \mathbb{R}^N$ from the noisy measurements $b = Ax^* + e \in \mathbb{R}^n$, where $2k \leq n \leq N$.*

Problem 1.3 (**Compressible recovery from noisy measurements**). *Recover an approximation to a k-compressible $x^* \in \mathbb{R}^N$ from the noisy measurements $b = Ax^* + e \in \mathbb{R}^n$, where $2k \leq n \leq N$.*

1.4 CS: measurements proportional to sparsity

1.4.1 Results for l_1-l_0 equivalence

The 'birth' of compressed sensing. An important breakthrough came in 2004, when both Candès, Romberg and Tao [36] and Donoho [67] showed that it is possible to design measurement matrices for which l_1-l_0 equivalence as defined in Definition 1.1 holds, provided the number of measurements satisfies $n = \mathcal{O}(k \log N)$. This implies that k-sparse signals can be exactly recovered using l_1-minimization from far fewer measurements than previously thought possible.

In [36], working in the complex domain $x^* \in \mathbb{C}^N$, l_1-l_0 equivalence is shown to hold with overwhelming probability for the case where the matrix A consists of n randomly chosen rows from a Fourier transform matrix. In [67], conditions to be satisfied by the measurement matrix are given, and it is shown that these conditions are satisfied with overwhelming probability for $A = \Phi\Psi$, where the columns of Φ are drawn at random from the unit sphere on $\mathbb{R}^n$, and where Ψ is an orthogonal transform matrix.

What is striking is that in both cases the proposed measurement scheme is either random or randomized, and for this reason probabilistic results are obtained which state that recovery occurs with 'overwhelming probability'. By overwhelming probability, it is meant that, if one considers a sequence of problem sizes $(k, n, N) \to \infty$, then the probability that recovery is unsuccessful decays exponentially with respect to these problem dimensions. The use of random or randomized matrices, which may appear at first to be somewhat counter-intuitive, is in fact crucial for obtaining better recovery guarantees, allowing stronger uncertainty principles to be derived. On the contrary, random matrices can be intutitively seen to be an attractive option: if one applies a random matrix to a sparse vector, one would expect the resulting measurements to be in some sense 'averaged out' across all measurement coefficients, leading to particularly strong uncertainty principles.

Proportional-dimensional results. While [36] and [67] are rightly seen as the foundational papers, improved results followed for random matrices. It was shown by both Candès and Tao [39] and Donoho [64] that recovery of any k-sparse signal is guaranteed provided the number of measurements satisfies

$$n = \mathcal{O}\left[k \log\left(\frac{N}{k}\right)\right]. \tag{1.9}$$

While it might appear that this is a minor adjustment to a log-factor, in fact the difference is significant. Condition (1.9) implies that recovery conditions may be characterized by their asymptotic behaviour as one lets the problem dimensions (k, n, N) grow in proportion to each other.

In [64], Donoho obtains the improved result for matrices satisfying the same conditions as in [67], and in particular for the uniform spherical ensemble. In [39], results were obtained for Gaussian random matrices, in which each entry independently and identically follows a Gaussian distribution. To obtain their results, the authors introduced a concept that would eventually become ubiquitous in CS: the Restricted Isometry Property (RIP), which we here define. Throughout this thesis, we will use $\|\cdot\|$ with no subscript to denote the Euclidean norm.

Definition 1.2 (**Restricted Isometry Property [39]**). *For a given matrix A, define the (symmetric) RIP constant R_s of order s to be*

$$R_s := \min_{r \geq 0} r \quad \textit{subject to} \quad (1-r)\|y\|^2 \leq \|Ay\|^2 \leq (1+r)\|y\|^2 \ \textit{ for all } s\textit{-sparse } y. \tag{1.10}$$

Recovery is guaranteed in [39] provided A satisfies

$$R_k + R_{2k} + R_{3k} < 1. \tag{1.11}$$

The authors then go on to show that, provided A is drawn from the Gaussian ensemble such

that (1.9) holds, the condition (1.11) holds with probability approaching 1 exponentially in n as one grows (k, n, N) proportionally.

Robust recovery guarantees. These early results concerned the simplified case in which noiseless measurements of an exactly sparse signal are taken. In reality, this is unlikely: signals are likely to be only approximately sparse, and measurements cannot be obtained to arbitrary accuracy. It is important, therefore, that recovery guarantees can be extended in both of these two directions. This was first done for the case of noisy measurements by Donoho [68] in a companion paper to [64], in which it was shown that an accurate approximation is guaranteed by solving the problem

$$\min_{x \in \mathbb{R}^N} \|x\|_1 \quad \text{subject to} \quad \|Ax - b\| \leq \sigma, \tag{1.12}$$

which is related to the Basis Pursuit De-Noising formulation previously used in [52], see Section 1.6 for further elaboration. In 2006, Candès, Romberg and Tao [37] extended their RIP-based results from [39] to the case of both noisy measurements and inexactly sparse signals. Since then, the RIP condition guaranteeing recovery in this sense has been gradually weakened by a number of authors, including a now well-known result by Candès [32] which we state next. Small improvements have since been made to this result by Chartrand [46], Foucart and Lai [94] and Cai, Wang and Xu [31], but we choose to state the result in [32] due to its particular simplicity. First, let us define x_k^* to be the best k-sparse approximation to x^*, as follows.

Definition 1.3 (**Best k-sparse approximation**). *Define $x_k^* \in \mathbb{R}^N$ to be*

$$x_k^* = \underset{\|z\|_0 \leq k}{\arg\min} \|z - x^*\|. \tag{1.13}$$

Note that x_k^* is in fact the vector x^* with all but its k largest in magnitude entries set to zero. The result in [32] follows.

Theorem 1.4 (**Robust RIP-based recovery [32]**). *Suppose that $\|e\| \leq \sigma$ and $R_{2k} < \sqrt{2}-1$. Then, given any $x^* \in \mathbb{R}^N$, the solution $\hat{x}$ to (1.12) satisfies*

$$\|\hat{x} - x^*\| \leq \left(\frac{2\sqrt{1 + R_{2k}}}{1 - R_{2k}} \right) \cdot \sigma + \left(\frac{\sqrt{2} R_{2k}}{1 - R_{2k}} \right) \cdot \frac{\|x^* - x_k^*\|_1}{\sqrt{k}}. \tag{1.14}$$

In words, the result guarantees that, provided a sufficiently strong RIP condition is satisfied by the matrix A, the approximation error of the solution $\hat{x}$ to (1.12) does not exceed a fixed multiple of the magnitude of the error due to noise in the measurements and inaccuracy of the sparse model.

1.4.2 Beyond l_1-l_0 equivalence

So far in this review, we have maintained a focus on one particular approach to CS recovery: l_1-minimization. Indeed, the theory of CS was first developed solely within the framework of l_1-l_0 equivalence. However, we have already seen that, even prior to the first CS results, algorithms which do not solve an l_1-minimization problem had already been proposed for sparse signal recovery. In particular, greedy matching pursuit algorithms had been developed, and coherence-based recovery guarantees had also been proved for Orthogonal Matching Pursuit (OMP) [153].

OMP with Gaussian measurements. The first theoretical result for CS with a greedy algorithm was proved by Tropp and Gilbert [154], who considered OMP with Gaussian measurements. They showed that, provided the number of measurements satisfies $n \geq 20k \log(N/a)$ for some $0 < a < 0.36$, then OMP with Gaussian measurements exactly recovers a k-sparse signal with probability exceeding $1 - 2a$. However, like the early results for l_1-l_0 equivalence [36, 67], this result does not permit a proportional-dimensional interpretation, and furthermore the probability of failure is fixed and does not become exponentially small.

RIP analysis beyond l_1-minimization. During 2008 and 2009, three more algorithms were shown to exactly recover k-sparse signals from noiseless measurements, or equivalently to recover compressible signals from noisy measurements to within a given accuracy, provided the measurement matrix satisfies an RIP condition. In [127], Needell and Tropp proposed Compressive Sampling Matching Pursuit (CoSaMP), obtaining the recovery condition $R_{4k} < 0.1$, while Dai and Milenkovic proposed Subspace Pursuit (SP) [56] and derived the condition $R_{3k} < 0.165$. In fact, CoSaMP and SP are very similar to each other, and are both closely related to OMP and StOMP. Around the same time, Blumensath and Davies [23] proposed Iterative Hard Thresholding (IHT), the algorithm which will largely become the focus of this thesis, and proved recovery under the condition $R_{3k} < 1/\sqrt{8}$. These original recovery guarantees have all subsequently been tightened: $R_{4k} < 0.384$ for CoSaMP by Foucart [92], $R_{3k} < 0.35$ for SP by Jain et al. [109], and $R_{3k} < 1/\sqrt{3}$ for IHT by Foucart [93]. More recently, an algorithm called Orthogonal Matching Pursuit with Replacement (OMPR) has been shown by Jain, Tewari and Dhillon to provide similar recovery guarantees under the condition $R_{2k} < 0.499$ [109]. It should be noted that each of these results parallels Theorem 1.4 for recovery using l_1-minimization, and implies that recovery is successful with Gaussian measurement matrices provided the number of measurements satisfies (1.9). For a brief description of OMP and related methods, we refer the reader to the overview of CS algorithms in Section 1.6. We give a thorough introduction to IHT and related variants in Section 1.7.

But what is the constant of proportionality? The results presented in this and the previous section establish that it is possible to recover compressible signals by taking a number

of measurements proportional to the sparsity of the signal's approximation. However, the crucial question for a practitioner is a quantitative question: how many measurements should I take to recover a given signal? Recovery conditions such as (1.9) are proportionality statements which leave this question unanswered unless the constant of proportionality is itself determined. A natural way to explore this important quantitative question is by using the *phase transition framework*, to which we now turn our focus in the next section.

1.5 Quantitative results via phase transitions

The results in [39] and [64] mentioned in Section 1.4 are asymptotic results for random matrices in which the dimensions (k, n, N) are allowed to grow to infinity while remaining proportional to each other. We may represent the latter by defining two new variables as the limiting values of the ratios n/N and k/n. We will refer to such a framework as the *proportional-growth asymptotic*, which we now formally define.

Definition 1.5 (**Proportional-growth asymptotic [21]**). *We say that a sequence of problem sizes (k, n, N), where $0 < k \leq n \leq N$, grows proportionally if, for some $\delta \in (0, 1]$ and $\rho \in (0, 1]$,*

$$\frac{n}{N} \to \delta \quad \textit{and} \quad \frac{k}{n} \to \rho \quad \textit{as } (k, n, N) \to \infty.$$

This framework defines a two-dimensional phase space for asymptotic analysis in which the variables δ and ρ have a simple practical interpretation. δ is the ratio by which the signal is undersampled (an *undersampling ratio*), while the ratio ρ effectively determines how many measurements need to be taken as a multiple of the sparsity (an *oversampling ratio*). A matter of vital practical importance is understanding the relationship between these variables. In CS, we want to recover signals with sparsity as high as possible, while taking as few measurements as possible: in other words, we want δ to be small and ρ to be large.

In particular, results such as those in [39] and [64] imply that recovery is asymptotically guaranteed for the random matrix ensemble in question, provided that $\rho < \hat{\rho}(\delta)$ for some recovery threshold $\hat{\rho}(\delta)$ which the authors do not determine. However, both authors make clear that the recovery thresholds obtainable from their analysis are rather pessimistic, and suggest that improved quantitative results might be possible.

Precise phase transitions for Gaussian matrices. In 2005, Donoho [66] showed that l_1-l_0 equivalence holds for all sparse vectors provided the measurement matrix satisfies a certain geometrical property. Namely, one requires the quotient polytope obtained by applying the matrix A to the l_1-ball or *cross-polytope* in $\mathbb{R}^N$,

$$\{x \in \mathbb{R}^N \ : \ \|x\|_1 \leq \tau\},$$

to be *k-neighbourly*. A k-neighbourly polytope is one where any $k+1$ vertices, not including an antipodal pair, span a face of the polytope. The condition means that, even though the cross-polytope is being projected into a lower dimension, all of its k-faces survive the projection, and none are swallowed up into the interior of the polytope.

The concept of k-neighbourliness had previously been a much-studied topic in combinatorial geometry. Affentranger and Schneider [2] derived and analyzed a formula for the expected number of k-faces surviving a random orthogonal projection, and further analysis was carried out by Böröczky and Henk [29], and also by Vershik and Sporyshev [159] who pioneered the study of the asymptotic case in which the dimensions grow proportionally. Drawing upon this rich body of work, Donoho [65] derived a threshold $\rho < \rho_S(\delta)$ for random orthogonal projection matrices which ensures that, in the proportional-growth asymptotic, with probability tending to 1 exponentially fast in n, all k-faces survive the projection, thus implying l_1-l_0 equivalence. The equivalence threshold $\rho_S(\delta)$ represented a radical improvement on those obtainable from the analysis in [39] and [64].

The subscript S in $\rho_S(\delta)$ stands for *strong*, since weaker forms of k-neighbourliness can also be considered [159]. By *weak* neighbourliness we mean that the overwhelming majority of k-faces survive projection, or alternatively that the probability that a single independently-chosen k-face survives projection is overwhelmingly high. Donoho [65] also derived a weak equivalence threshold $\rho < \rho_W(\delta)$ for random orthogonal matrices which guarantees that l_1-l_0 equivalence holds for the overwhelming majority of k-sparse vectors, or alternatively that l_1-l_0 equivalence holds with overwhelming probability for any k-sparse vector whose k non-zero coefficients are located on a single independently-chosen support set and have a particular sign pattern. Weak equivalence may therefore be viewed as a particular kind of average-case analysis, as opposed to strong equivalence which is worst-case analysis. The weak threshold represents a substantial improvement on the strong threshold, which is not surprising since it requires the satisfaction of a less stringent condition.

A more comprehensive treatment of the above results is given by Donoho and Tanner in [77]. In particular, it was proved that both the strong and weak equivalence thresholds are precise, with equivalence also failing to hold with overwhelming probability above the respective curves, thus giving *phase transitions*. These results are significant since they were the first to describe quantitatively and precisely the allowable rates of undersampling in compressed sensing. Furthermore, it follows from the work of Baryshnikov and Vitale [9] that their phase transitions, as well as applying to random orthogonal matrices, in fact apply equally well to Gaussian random matrices. Plots of the strong and weak phase transitions for Gaussian measurement matrices are shown in Figure 1.1. Recovery is asymptotically certain, in the strong and weak senses respectively, for a (δ, ρ) pair falling below the curve, and asymptotically certain to fail for a (δ, ρ) pair falling above the curve.

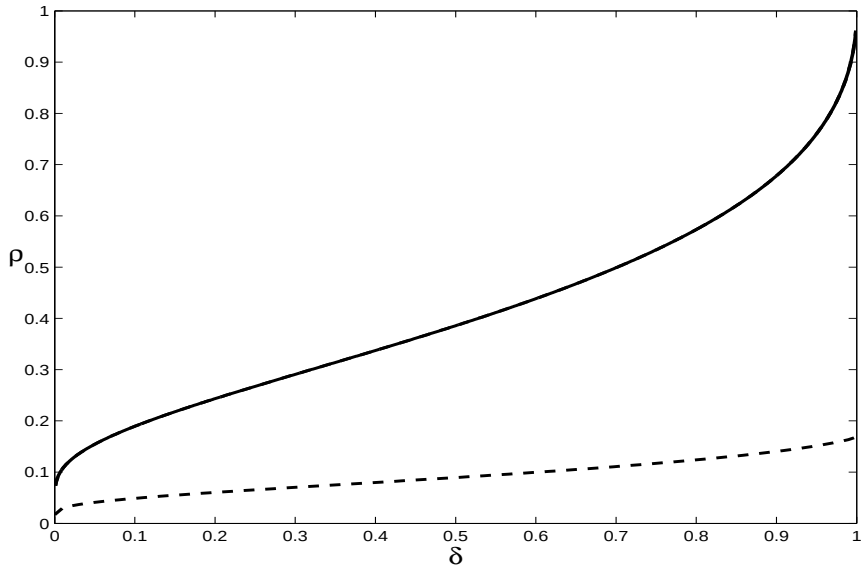


Figure 1.1: Phase transitions for Gaussian matrices and l_1-minimization: weak (unbroken); strong (dashed) [77].

Empirical phase transitions. Not long after the deep connections between compressed sensing and combinatorial geometry were made, it was shown empirically that a phase transition is also observed in practice for l_1-l_0 equivalence. One of the earliest examples to appear in print was in 2006 [80], where Donoho et al. presented the results of numerical testing in which high-dimensional problems were tested with the aim of well approximating the asymptotic limit. By considering a mesh of equispaced points in the (δ, ρ) plane, multiple trials were performed where in each case n measurements were taken of a k-sparse vector x^* of dimension $N = 800$, such that $n = \delta N$ and $k = \rho n$. In each trial, the nonzero coefficients of x^* were drawn i.i.d. Gaussian, and the measurement matrix was drawn from the uniform spherical ensemble (for which theoretical results were obtained in [67] and [64]). In each trial, the solution $\hat{x}$ to (1.2) was obtained, and recovery was deemed successful if the error between $\hat{x}$ and x^* was below some small tolerance level. Plots were obtained of the proportion of successful recoveries throughout the (δ, ρ) plane. The plots showed that a phase transition phenomenon occurs in practice, though not entirely sharp. Instead, one may observe a narrow band within which the probability of success transitions from zero to one. But the most striking finding was that, if one overlays the theoretical phase transition derived in [65], there is close agreement with the level curve for the empirical proportion of successful recoveries equal to 1/2. In other words, for high-dimensional problems, one observes empirical behaviour that is essentially described by the weak phase transition. This is in fact to be expected since, in the experiment just described, the signal (and therefore its support) is drawn independently of the measurement matrix. This shows that the strong phase transition, while being a theoretically interesting concept, is pessimistic concerning how many measurements should be taken in practice.

Universality of empirical phase transitions. The other intriguing aspect of the numerical results just described is that one observes close agreement with the theoretical phase transition for Gaussian matrices, even though the matrix used in the tests is not Gaussian, but drawn from the uniform spherical ensemble. On the other hand, these two distributions are intimately connected, since a uniform spherical matrix may be obtained by normalizing the columns of a Gaussian matrix. But it nevertheless raises the question: might similar empirical results be observed for other families of matrices? Indeed, some of the early theoretical results pointed in this direction: we have seen, for instance, that the foundational paper of Candès, Romberg and Tao [36] obtained results for random partial Fourier matrices in the complex domain.

In [78], Donoho and Tanner performed extensive numerical testing of l_1-l_0 equivalence similar to those in [80] for measurement matrices drawn from a number of random or randomized matrix ensembles. Empirical phase transitions were obtained for various random matrix ensembles, including Gaussian matrices, Bernoulli matrices where all entries are independent and equally likely to be zero or one, and Rademacher matrices where all entries are independent and equally likely to be plus or minus one. Phase transitions were also obtained for random partial deterministic transform matrices, including the Discrete Cosine Transform (real counterpart of the Fourier transform) and the Hadamard transform. A plot of these empirical phase transitions is given in Figure 1.2 [78]. The lower curve is the one of immediate interest: the upper curve is the for the case in which all coefficients of the sparse signal are assumed to be nonnegative. A close agreement can be observed between each respective empirical phase transition and the theoretical weak phase transition for Gaussian matrices. The result for Gaussian matrices indicates that a close approximation to the asymptotic limit is reached for finite dimensions of reasonable size (the authors used $N = 1600$). That the other non-Gaussian matrix ensembles should exhibit the same phase transition is remarkable, and points to the universality of the weak phase transition of [77] over a wide class of different random or randomized matrix ensembles.

Phase transitions beyond l_1-l_0 equivalence. The concept of a phase transition can also be extended to describe the region of phase space in which any CS algorithm exactly recovers a sparse signal from noiseless measurements (Problem 1.1). Quantitative results in the phase transition framework have also been obtained for three algorithms which do not rely upon l_1-l_0 equivalence, CoSaMP, SP and IHT (see also Section 1.4.2). Building upon the existing RIP [39] analysis for each algorithm, lower bounds on the theoretical strong phase transition for each of the three algorithms were derived by Blanchard et al. in 2011 [17]. One of the contributions of the work was the weakening of the RIP conditions from [127, 56, 23] by considering asymmetric RIP constants, which we next define.

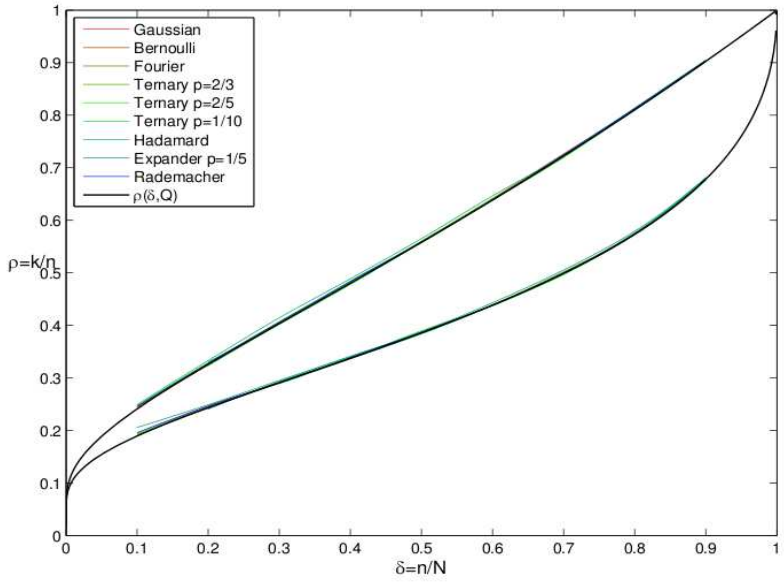


Figure 1.2: Empirical phase transitions for l_1-minimization used in conjunction with various random and randomized matrix ensembles [78].

Definition 1.6 (**Asymmetric RIP [21]**). *For a given matrix A, define L_s and U_s, the lower and upper RIP constants of order s, to be, respectively,*

$$L_s := 1 - \min_{1 \leq \|y\|_0 \leq s} \frac{\|Ay\|^2}{\|y\|^2} \quad \text{and} \quad U_s := \max_{1 \leq \|y\|_0 \leq s} \frac{\|Ay\|^2}{\|y\|^2} - 1. \tag{1.15}$$

We note that symmetric and asymmetric RIP constants satisfy the relation $R_s = \max(L_s, U_s)$. It follows that any asymmetric RIP condition is made stricter in general by the imposition of symmetry upon the RIP constants. Indeed, this imposition is in many cases artificial, since RIP constants often do not exhibit symmetry. For example, it was observed in [21] that U_s is often significantly larger than L_s for Gaussian matrices. From now on in this thesis, when we refer to the RIP, we will assume it to be the asymmetric variant.

By making use of upper bounds on asymmetric RIP constants for Gaussian matrices derived by Blanchard et al. in [21], lower bounds on the strong phase transition for exact recovery were obtained in [17] in the proportional-growth asymptotic for each of CoSaMP, SP and IHT. The phase transition derived in [17] are displayed in Figure 1.3. The reader's attention should be drawn in particular to the scale on the ρ-axis, showing that, for each algorithm, critical values of the ratio ρ are everywhere of the order 10^{-3} or worse. Since $\rho = k/n$, these phase transitions imply that a number of measurements of the order at least 1000 times the sparsity is needed to guarantee recovery using any of the three algorithms. Empirical studies reported in [21], in which lower bounds on RIP constants were calculated numerically, show that the bounds are quite sharp, being always within a factor of 1.83 of empirically observed lower bounds

for $n = 400$, with the factor being even lower for small ρ. The conclusion is therefore that, at least for Gaussian measurement matrices, the RIP gives phase transitions for these three algorithms which are much lower than those obtained by means of combinatorial geometry for l_1-l_0 equivalence. Since the work in [17] was carried out, Bah and Tanner [4] derived improved upper bounds on RIP constants for Gaussian matrices, which are always within a factor of 1.57 of empirically observed lower bounds for $n = 400$. We will have cause to make use of Gaussian RIP bounds later in this thesis (Chapters 2 and 5), and we will use the bounds from [4] in order to obtain the best possible quantitative results.

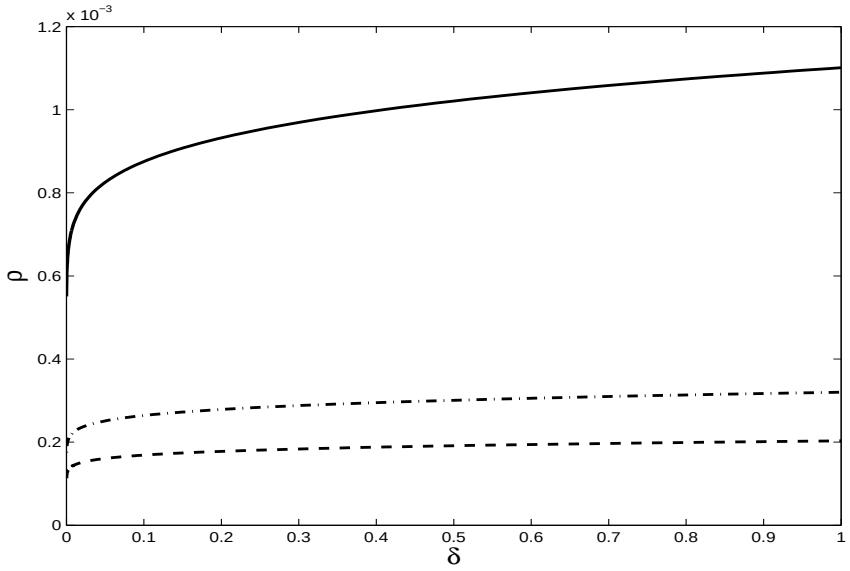


Figure 1.3: Lower bounds on the strong phase transition based upon the RIP for recovery using Gaussian matrices: IHT (unbroken); CoSaMP (dashed); SP (dash-dot) [17].

These theoretical results are in marked contrast to the observed behaviour of CoSaMP, SP and IHT. Numerical testing of recovery properties similar to that described earlier in this section was carried out for each of the three algorithms in question in [72], revealing phase transitions reasonably comparable to the weak phase transition for l_1-l_0 equivalence, though slightly lower. More recently, GPUs have been used to empirically test IHT algorithms upon problems of high dimension [18]. The authors tested a variable stepsize variant of IHT known as Normalised IHT (NIHT) [27], which is the subject of theoretical analysis in this thesis. Figure 1.4 displays empirical phase transitions for different problem sizes obtained for NIHT used in conjunction with partial DCT measurement matrices [18]. Superimposed in blue is the theoretical weak phase transition for l_1-minimization. The plot demonstrates something remarkable: NIHT exhibits an empirical phase transition which is at least as good as that for l_1-minimization, making it competitive with l_1-minimization from the point of view of recovery performance. See Sections 1.7 and 2.1 for more details on NIHT.

From Figures 1.1 and 1.3, we conclude that there exists a considerable gap between theo-

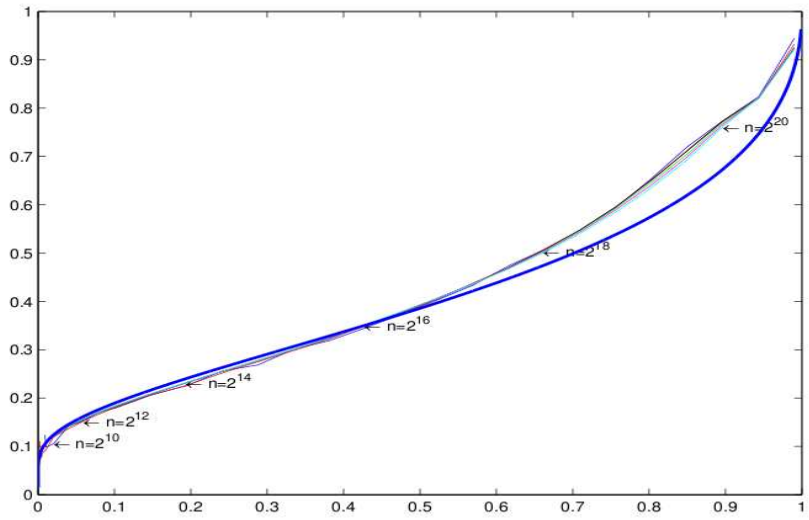


Figure 1.4: Empirical phase transitions for NIHT for different problem sizes, alongside the weak phase transition for l_1-minimization (blue).

retical recovery guarantees and empirically observed performance for IHT and other algorithms which do not rely upon l_1-minimization. We can see, therefore, that the RIP is something of a 'blunt instrument' when it comes to proving quantitative results in compressed sensing. It is versatile and easy to use, but in some sense fails to capture the essence of what makes a CS algorithm successful or otherwise. Part of the problem is that it is a worst-case technique and is therefore limited by the (at present unknown) strong phase transition for each algorithm, whereas we have seen that it is the average-case weak phase transition which describes the practical behaviour of a given algorithm. In this thesis, we will obtain the first average-case recovery guarantees for IHT and a related variant, in the form of lower bounds on a particular kind of weak phase transition for Gaussian measurement matrices (see Chapter 5).

1.6 Algorithms for CS: a survey

In this section, we present a brief survey of the various algorithms that have been proposed for compressed sensing. We have established that the task in CS is to recover an underlying sparse solution from an underdetermined system of linear equations. Considering first the noiseless case, the problem of finding the *sparsest* solution to the linear system is found by solving (1.4). In this simplified case, any k-sparse solution with $2k \leq n$ is also the sparsest solution, provided the measurement matrix satisfies a somewhat mild RIP condition, which we next state.

Lemma 1.7 (**Sparsest solution condition [70]**). *Consider Problem 1.1 and suppose that $L_{2k} < 1$ holds for the matrix A. Then x^* is the solution to (1.4).*

Proof: Let y denote the sparsest solution to $b = Ax$, and suppose $y \neq x^*$. Then $\|y\|_0 \leq k$

which implies that $\|y - x^*\|_0 \leq 2k$. Applying (1.15), $\|A(y - x^*)\| \geq (1 - L_{2k})\|y - x^*\|$. Since $y \neq x^*$, and supposing $L_{2k} < 1$, we have $\|A(y - x^*)\| > 0$. One the other hand, $A(y - x^*) = Ay - Ax^* = b - b = 0$, which is a contradiction. □

The condition $L_{2k} < 1$ is equivalent to requiring that any $2k$ columns of A are linearly independent. Equivalently, Donoho and Elad [70] defined the *spark* of a matrix as the smallest s such that there exists a linearly dependent collection of s columns, and therefore the condition may also be expressed as $\text{spark}(A) > 2k$. Provided $2k \leq n$, this condition is satisfied by a Gaussian matrix with probability 1 [42], and also by any matrix which is in *general position* (or equivalently $L_n < 1$). The conclusion is that it is often possible to assume, at least in the noiseless setting, that a sparse solution is in fact the most sparse solution, thus reducing the task of recovery to solving (1.4). For this reason, we will often assume in this thesis that $L_{2k} < 1$ holds, which we formalize as the following assumption.

Assumption 1 ($2k$**-column linear independence**)**.** *Assume that the matrix A satisfies $L_{2k} < 1$, that is, any $2k$ of its columns are linearly independent.*

If the measurements are noisy, or if the signal is inexactly sparse, (1.4) may be infeasible, but we can instead consider solving the problem

$$\min_{x \in \mathbb{R}^N} \|x\|_0 \quad \text{subject to} \quad \|Ax - b\| \leq \sigma, \tag{1.16}$$

which finds the sparsest solution which satisfies the measurements to within some noise tolerance parameter $\sigma > 0$. CS algorithms can be naturally divided into two categories: those which solve convex relaxations in which the l_0-norm in (1.16) is replaced with the l_1-norm, and those which solve the original problem (1.16) more directly. We will present a survey of both categories, turning our focus first to algorithms for convex relaxation.

1.6.1 Algorithms for solving the l_1 convex relaxation

First, let us make precise what we mean by the convex relaxation of (1.16). Given a function $f(x) : X \to \mathbb{R}$ defined on some subset $X \subseteq \mathbb{R}^N$, a *convex underestimator* of f is a convex function c such that $c(x) \leq f(x)$ for all $x \in X$. The *convex envelope* of f is then defined as the convex function $\hat{c}$ which is the supremum over all convex underestimators, that is

$$\hat{c}(x) := \sup\{c(x) \;\; : \;\; c(x) \leq f(x) \;\; \forall\, x \in X, \;\; c \text{ convex}\}.$$

The l_1-norm can be interpreted as the convex envelope of the l_0 pseudo-norm in the following sense, where we define the infinity norm of $x \in \mathbb{R}^N$ in the usual way as $\|x\|_\infty := \max_{1 \leq i \leq N} |x_i|$.

Lemma 1.8 (**Convex envelope of the l_0-norm**). *Let $x \in \mathbb{R}^N$ and let $R > 0$. Then $(1/R)\|x\|_1$ is the convex envelope of $\|x\|_0$ on the set $\{x : \|x\|_\infty \leq R\}$.*

It follows that the closest convex relaxation to (1.16) is obtained by replacing the l_0-norm with the l_1-norm, thus obtaining (1.12).

Next we introduce two further convex formulations, which can be shown to be essentially equivalent to (1.12). The first is obtained by interchanging the objective and constraint functions in (1.12), to give

$$\min_{x\in\mathbb{R}^N} \frac{1}{2}\|Ax-b\|^2 \quad \text{subject to} \quad \|x\|_1 \leq \tau, \tag{1.17}$$

where $\tau \geq 0$. One may view (1.17) as a linear least-squares problem with an l_1-ball constraint. The second formulation is the unconstrained problem

$$\min_{x\in\mathbb{R}^N} \frac{1}{2}\|Ax-b\|^2 + \lambda\|x\|_1, \tag{1.18}$$

where $\lambda \geq 0$, which follows an often-used technique in optimization of removing the l_1 constraint in (1.17) and instead penalizing it in the objective. Each of the formulations (1.12), (1.17) and (1.18) is convex but non-smooth, due to the presence of the $\|x\|_1$ term. However, should it be required, it is straightforward to convert each formulation into a smooth problem by splitting x into its positive and negative parts and introducing further non-negativity constraints, which transforms (1.12) into a second-order cone problem (SOCP), and (1.17) and (1.18) into bound-constrained quadratic programs (BCQPs). The following theorem establishes that the three formulations are equivalent under an appropriate choice of parameters.

Theorem 1.9 (**Equivalence of l_1 formulations [147, Theorem 7]**). *Consider problems (1.12), (1.17) and (1.18).*

1. *Given $\lambda \geq 0$, the global solution $\hat{x}(\lambda)$ of (1.18) is also the global solution of (1.17) for some $\tau \geq 0$. Conversely, given $\tau \geq 0$, the global solution $\hat{x}(\tau)$ of (1.17) is also the global solution of (1.18) for some $\lambda \geq 0$.*

2. *Given $\lambda \geq 0$, the global solution $\hat{x}(\lambda)$ of (1.18) is also the global solution of (1.12) for some $\sigma \geq 0$. Conversely, given $\sigma \geq 0$, the global solution $\hat{x}(\sigma)$ of (1.12) is either $\hat{x}(\sigma) = 0$ or else is the global solution of (1.18) for some $\lambda \geq 0$.*

Interior point methods. The advent of the interior point method (IPM) has revolutionized large-scale convex optimization, and IPMs are now considered to be the most computationally efficient option in many applications, especially when dimensions are large. In contrast to many algorithms for constrained optimization which move around the boundary of the feasible constraint set, IPMs take a path through the interior of the feasible set by solving a sequence of

perturbed problems which converge to the original problem. As was noted in Section 1.2, the problem (1.12) was first proposed in the context of sparse approximation by Donoho et al. in 1999 [52] under the name of Basis Pursuit De-Noising (BPDN). The authors proposed recasting (1.12) as a perturbed linear program and using an IPM originally designed for regularized linear programs [99]. Later, as the first results for robust CS began to appear [68, 37], Candès and Romberg proposed l_1-MAGIC [35], an IPM which solves the second-order cone problem (SOCP) into which (1.12) can easily be recast. The authors make clear that their algorithm is not designed to be cutting-edge, but more a proof of concept: a more computationally efficient IPM algorithm was proposed in 2007 by Kim et al. [112], which they called `l1_ls`, which stands for 'l_1-regularized least squares'. This time, (1.18) is recast as a bound-constrained quadratic problem (BCQP), and greater computational efficiency is achieved through preconditioning. A truncated conjugate-gradient method is used to solve the inner iterations which, like the algorithm used in [52], allows the measurement matrix to be defined implicitly as a matrix-vector multiplication. Many of the transforms frequently used in CS, such as the DCT or wavelets, have fast implicit transforms, making this an important issue for any CS algorithm. Interior point methods are likely to offer a considerable advantage over gradient-based methods if the matrices defining the problem are ill-conditioned, and they are especially effective at exploiting sparsity in the matrices involved. However, neither of these two features characterize the CS problem: good CS measurement matrices are likely to be very well-conditioned and highly dense, meaning that IPMs do not offer the same advantages for CS problems as they often do for other problems.

Iteratively reweighted least squares (IRLS). In IRLS, a sequence of l_2-regularized least-squares problems are solved in which the l_1-term in (1.18) is replaced by a weighted l_2-norm term which is designed to converge to the l_1-norm as the algorithm progresses. In each iteration, a least-squares fit must be solved which involves the inversion of the full measurement matrix A, which unfortunately makes the computational cost somewhat high. IRLS for l_1-minimization is one of a more general family of methods, which date back to the work of Lawson in 1961 [116]. The first IRLS methods for l_1-minimization were developed in the 1970s, and it was first proposed as an algorithm for sparse approximation by Gorodnitsky and Rao in 1997 [101], under the name FOCUSS (Focal Underdetermined System Solver). Interest in IRLS was renewed with the advent of CS, and a new weighting scheme was proposed in 2008 [49] by Chartrand and Yin, which led to an improvement in performance, and for which Daubechies et al. [61] obtained theoretical recovery guarantees using the RIP. Wipf and Nagarajan [160] proposed a new non-separable weighting scheme in 2010.

Stagewise pivoting. In statistical regression, (1.18) was proposed for the overdetermined case in 1996 by Tibshirani [151] as a technique for sparse linear regression, and was referred to

as Least Absolute Shrinkage and Selection (LASSO). In this context, Osborne et al. proposed an algorithm called Homotopy in 2000 [132] which exploits the fact that the solution to (1.18) is piecewise-linear as a function of λ. Starting with the solution at $\lambda = 0$ (which is just the zero vector), the next value of λ at which the support set of the solution changes is located at each stage, and the least-squares solution on the current support is then implicitly calculated by updating a QR factorization. Later, a version of Homotopy adapted to the underdetermined setting of CS was proposed in [121]. In 2004, Efron et al. proposed a similar algorithm, called Least Angle Regression (LARS) [86], which approximates the solution path of (1.18) by a series of greedy updates. In fact, LARS may be viewed as an l_1 counterpart of the greedy OMP algorithm which was encountered in Sections 1.2 and 1.4.2.

Gradient methods. By far the most popular approach to l_1-minimization problems in CS is to use a gradient method, a wide selection of which are now available. In a gradient method, the competing goals of satisfying the measurements and maintaining sparsity are separated from each other and performed sequentially, with the algorithm alternating between gradient steps to fit the linear system, and projection or shrinkage steps to drive down the l_1-norm. This is an attractive approach since projection onto the l_1-ball can be accomplished by a straightforward component-wise shrinkage operation often referred to as the *soft threshold*. More precisely, writing $\mathcal{S}_\tau(x)$ for the projection of $x \in \mathbb{R}^N$ onto the l_1-ball of radius τ for some $\tau \geq 0$, namely

$$\mathcal{S}_\tau(x) := \min_{y \in \mathbb{R}^N} \|y - x\| \quad \text{subject to} \quad \|y\|_1 \leq \tau,$$

then $\mathcal{S}_\tau(x)$ is also the solution to the unconstrained problem

$$\min_{y \in \mathbb{R}^N} \frac{1}{2}\|y - x\|^2 + \lambda \|y\|_1,$$

for some $\lambda \geq 0$, which has the closed-form solution

$$\{\mathcal{S}_\tau(x)\}_i = \operatorname{sgn}(x_i) \cdot \min(|x_i| - \lambda, 0) \quad \text{for all } i = 1, 2, \ldots, N.$$

The designation 'soft' distinguishes the operation from the nonconvex *hard threshold* which projects onto the l_0-ball, see Section 1.7.

The most computationally expensive steps in a typical gradient method are the matrix-vector products required to perform the gradient step, and the soft threshold projections. By contrast, other algorithms for l_1-minimization require the repeated solution of large system of equations and the equivalent inversion of a matrix (for example, IPMs and IRLS), or the repeated updating of matrix factorizations (for example, stagewise pivoting methods). For this reason, gradient methods are often viewed as the most computationally efficient algorithms for l_1-minimization. They are, however, especially sensitive to the conditioning of the measurement

matrix; although we have seen that CS measurement matrices are often very well conditioned. Also applicable is the more general point that faster implementations are possible by making use of transform matrices for which implicit matrix-vector products are available.

A number of iterative soft thresholding (IST) methods have been proposed for the unconstrained problem (1.18) in recent years, masquerading under the various names of iterative thresholding [60], forward-backward splitting [54], fixed-point iteration [106] and sparse reconstruction by separable approximation (SpaRSA) [161], with the algorithm being first developed in 2003 as a specific case of an expectation-maximization (EM) algorithm for Bayesian image reconstruction [90]. Standard convergence results for these methods require small gradient steps to be taken, whereas optimal performance is often attained by taking much larger steps, such as the Barzilai-Borwein criterion, for example in [161]. Two-stage enhancements to these standard IST methods have also been proposed, in which use is made of information from the previous two iterations. Faster empirical performance is reported for TwIST (two-step IST) [15], and a faster rate of global convergence is both proved and observed for FISTA (fast IST algorithm) [11], which relies upon recent work on optimal gradient methods in [129]. Most recently, a two-stage soft thresholding algorithm called Approximate Message Passing (AMP), inspired by ideas from belief propagation in graphical models, was proposed in [73]. Empirical phase transitions were obtained for AMP, which were in close agreement with those for l_1-l_0 equivalence, suggesting that the algorithm has optimal recovery performance for an l_1 algorithm. These empirical findings were followed by a proof in [10] that, in the proportional-dimensional framework for Gaussian matrices, AMP converges to the solution of (1.18) in the high-dimensional limit.

Not all gradient methods for CS solve the formulation (1.18) however: SPGL1 [157], which is also based on soft thresholding, instead solves (1.17) by gradient projection, and connections between the solutions of (1.17) and (1.12) are also exploited to embed the algorithm within a scalar equation solver which enables the solution of (1.12). There also exist gradient methods which take an entirely different approach to soft thresholding: NESTA [12] (named after Nesterov) makes use of optimal gradient ideas from [129], but solves instead the formulation (1.12). Meanwhile, GPSR (gradient projection for sparse reconstruction) [91] is a gradient projection algorithm which solves the smooth reformulation of (1.18) as a bound-constrained quadratic program (BCQP), by splitting the variables into their positive and negative parts.

1.6.2 Alternatives to l_1-minimization

l_p-relaxation. The l_p-norm for $p \geq 1$ of a vector $x \in \mathbb{R}^N$ is defined as

$$\|x\|_p := \left(\sum_{i=1}^{N} |x_i|^p \right)^{\frac{1}{p}}, \tag{1.19}$$

where the l_1-norm is recovered by setting $p = 1$ in (1.19). One may also extend the definition to any $0 < p < 1$, although the resulting function is no longer a norm. However, it is well-known that minimizing the l_p-'norm' also promotes sparsity, and another possible approach to CS recovery is to solve the nonconvex problem

$$\min_{x \in \mathbb{R}^N} \frac{1}{2}\|Ax - b\|^2 + \lambda \|x\|_p, \tag{1.20}$$

or a related variant. Furthermore, certain recovery results for arbitrary measurement matrices have also been improved by switching from l_1 to l_p: In 2007, Chartrand [45] generalized an RIP condition for l_1-l_0 equivalence of Candès et al. [38] to the general case of l_p-l_0 equivalence, obtaining a condition that weakens as p decreases. Shortly afterwards, Chartrand and Staneva [48] proved that Gaussian matrices satisfy the RIP condition in [45] with overwhelming probability provided the number of measurements satisfies

$$n \geq C_1(p) \cdot k + p \cdot C_2(p) \cdot k \log\left(\frac{N}{k}\right),$$

where $C_1(p)$ and $C_2(p)$ are bounded functions of p in $0 < p \leq 1$.

Quasi-Newton methods [45] have been proposed for (1.20), while in addition the IRLS [49] and iterative shrinkage [47] techniques for l_1-minimization have been extended to the l_p setting. However, due to (1.20) being nonconvex, these algorithms are not guaranteed to find its global minimizer, but only a local minimizer. On the other hand, numerical experiments in [45, 49, 47] suggest that the global solution is often obtained, and that an improvement in empirical recovery performance is observed for the algorithms in question as one decreases p towards zero. However, there is no evidence that there exists an algorithm for l_p-minimization with superior recovery performance to optimal algorithms for l_1-minimization.

Greedy algorithms. We have already met the greedy LARS [86] algorithm in the l_1 setting. However, most greedy algorithms for CS are designed to directly solve the l_0-constrained problem

$$\min_{x \in \mathbb{R}^N} \frac{1}{2}\|Ax - b\|^2 \quad \text{subject to} \quad \|x\|_0 \leq k, \tag{1.21}$$

where k is a positive integer sparsity parameter. The problem (1.21) will be the one of most interest in this thesis, and for this reason we will simply refer to it from now onwards as the *l_0-problem*. Greedy algorithms for l_0, often referred to as *matching pursuits*, approximate the solution path $\hat{x}(k)$ of (1.21) by means of a sequence of local updates in which the sparsity of the solution is gradually increased. Or viewed another way, matching pursuits seek to gradually identify the support set of the signal. In a typical greedy algorithm, new support set elements are usually selected according to the gradient information of the current iterate, and the least-squares solution on the new support is either calculated explicity, or implicitly by updating a

QR factorization.

Matching pursuit algorithms were first used in statistical regression, and can be traced back as far as 1960 [87]. Mallat and Zhang [124] proposed the first such algorithm in a signal processing context in 1993, which they simply referred to as Matching Pursuit (MP), though it had been known to the statistics community for some time previously as Projection Pursuit Regression [95]. The better-performing but more computationally expensive Orthogonal Matching Pursuit (OMP) soon followed, originally proposed in [51] and introduced to the signal processing community by [62], see also theoretical results for OMP in Section 1.4.2. Blumensath and Davies later proposed a truncated conjugate gradient algorithm in [22] for approximating the least-squares projections, thus reducing computational expense. Other algorithms extended OMP by allowing several elements to be selected in each iteration, most notably Stagewise Orthogonal Matching Pursuit (StOMP) [80], Regularized Orthogonal Matching Pursuit (ROMP) [128] and Stagewise Weak Orthogonal Matching Pursuit (SWOMP) [25].

Later support identification algorithms were proposed which do not build up the solution in a stagewise manner, but instead maintain a support set of size k which is repeatedly both added to and pruned back. Algorithms in this category are Compressive Sampling Matching Pursuit (CoSaMP) [127] and Subspace Pursuit (SP) [56], see Section 1.5 for recovery guarantees and phase transitions for Gaussian matrices. Since least-squares projections are performed both after adding to and pruning back the support set, these algorithms have been referred to as two-stage pursuit algorithms [72]. Various alterations to OMP and two-stage pursuits continue appear, for example Cycling Matching Pursuit [145], Stepwise Optimal Subspace Pursuit [158], OMP with Replacement (OMPR) [109] and A* OMP [110].

1.7 Iterative Hard Thresholding

Iterative Hard Thresholding (IHT) is a *gradient projection* method for (1.21). A gradient projection method can be applied to any problem of the form

$$\min_{x\in\mathbb{R}^N} \; f(x) \quad \text{subject to} \quad x \in F,$$

where $f(x)$ is differentiable, and where the Euclidean projection onto F exists, is well-defined and can be computed. In each iteration of a gradient projection algorithm, one takes a step in the direction of instantaneous steepest descent, namely the negative gradient of $f(x)$ at the current iterate, before projecting back onto the feasible set F, namely

$$x^{m+1} := \mathcal{P}_F\{x^m - \alpha^m \nabla f(x^m)\}, \tag{1.22}$$

where we define the Euclidean projection onto the set F as

$$\mathcal{P}_F(x) = \underset{z\in\mathbb{R}^N}{\arg\min} \|z - x\| \quad \text{subject to} \quad z \in F, \tag{1.23}$$

and where α^m is a (possibly variable) stepsize.

In the context of the l_0-problem, writing

$$\Psi(x) := \frac{1}{2}\|b - Ax\|^2 \tag{1.24}$$

for the objective function of (1.21), we have $\nabla\Psi(x) = -A^T(b - Ax)$. Meanwhile, if the feasible set is the nonconvex set of k-sparse vectors in $\mathbb{R}^N$, then the projection in (1.23), which we denote by $\mathcal{H}_k(\cdot)$, is often referred to in the signal processing community as the *hard threshold* operator, defined to be

$$\mathcal{H}_k(x) = \underset{\|z\|_0\leq k}{\arg\min} \|z - x\|. \tag{1.25}$$

As the name suggests, $\mathcal{H}_k(\cdot)$ may also be viewed as a thresholding operator which keeps the k largest in magnitude coefficients of a vector while setting the rest to zero, as the following lemma establishes. We will assume that, if the kth and $(k+1)$th entries are of equal magnitude, some predefined ordering is used to determine the kth largest magnitude.

Lemma 1.10 (**Equivalence of the hard threshold and l_0 projection**). *Let $\mathcal{H}_k(\cdot)$ be defined as in (1.25). Then*

$$\{\mathcal{H}_k(x)\}_i := \begin{cases} x_i & i \in \Gamma \\ 0 & i \notin \Gamma \end{cases} \quad \textit{where } \Gamma := \{\textit{indices of the } k \textit{ largest in magnitude entries of } x\}. \tag{1.26}$$

Proof: Let $\|z\|_0 \leq k$ and let $\text{supp}(z) = \Gamma$, so that $|\Gamma| \leq k$. Then

$$\|z - x\|^2 = \|(z - x)_\Gamma\|^2 + \|x_{\Gamma^C}\|^2,$$

and $\|(z - x)_\Gamma\|^2$ is minimized by setting $z_i = x_i$ for all $i \in \Gamma$. Meanwhile, $\|x_{\Gamma^C}\|$ is minimized by choosing Γ to be the set corresponding to the k largest in magnitude coefficients of x. □

Applying (1.22) to (1.21), using (1.25), gives the generic IHT iteration

$$x^{m+1} := \mathcal{H}_k\{x^m + \alpha^m A^T(b - Ax^m)\}. \tag{1.27}$$

Tracing the algorithm's history, gradient projection algorithms were first proposed by Goldstein in 1964 [100], thus extending the steepest descent method for unconstrained optimization, first used as far back as 1847 by Cauchy [44], to the setting of constrained optimization. Due

to a seminal paper by Landweber in 1951 [114], gradient projection methods became known as *projected Landweber* algorithms in the signal processing community in the 1990s, although the constraint set in question was usually convex, see for example [135].

A method very similar to IHT was proposed by Herrity et al. in 2006 [107], where a hard threshold in which all coefficients with absolute value above some fixed value are kept in each iteration. Interestingly, this method is in fact a counterpart to the IST method for (1.18) in the l_1 setting, which uses the soft threshold, see section 1.6.1. In 2008, Blumensath and Davies [26] proposed IHT with unit stepsize as a variant of the algorithm in [107] in which exactly k nonzero coefficient are retained by the hard threshold in each iteration, for some fixed sparsity parameter k. The authors also proved convergence of the algorithm under the somewhat restrictive condition that the spectral norm of the matrix is less than 1. In 2009, Blumensath and Davies proposed IHT in the context of CS [23], proving recovery of sparse signals, and stable recovery of compressible signals from noisy measurements, provided A satisfies the RIP condition $R_{2k} < 1/\sqrt{8}$. By analysing this condition in the proportional-dimensional asymptotic for Gaussian matrices, Blanchard et al. [17] derived a lower bound on the strong phase transition for IHT (see Section 1.5 for further elaboration).

In [27], Blumensath and Davies proposed a variant with variable stepsize, which they termed Normalised Iterative Hard Thresholding (NIHT). In NIHT, α^m is chosen to give the maximum possible decrease in the objective function of (1.21) in the case that the support set remains unchanged, with a backtracking strategy used to ensure sufficient decrease in the case of a change in support. The authors prove that NIHT is guaranteed to always converge, and they also give an RIP condition guaranteeing the convergence of NIHT to the underlying signal x. These results highlight the important issue that a recovery algorithm may converge, but to an undesired solution: in particular, algorithms for nonconvex optimization often converge to local minima, but not necessarily to the global minimum.

Several improvements to the first RIP result for constant stepsize IHT in [23] have since appeared: Garg and Khandekar [96] obtained $R_{2k} < 1/3$ for a constant stepsize of $\alpha := 1/(1 + U_{2k})$. More recently, Foucart obtained $R_{3k} < 1/2$ [92], and subsequently $R_{3k} < 1/\sqrt{3}$ [93], both for unit stepsize. See Chapter 2 for an analysis of the current best RIP conditions for IHT and NIHT within the same phase transition framework used in [17]. The point was made in Section 1.5 that the earliest RIP-based recovery guarantee for IHT leads to pessimistic quantitative predictions concerning the degree of undersampling permitted by the algorithm in the case of Gaussian matrices, and we will show that the same is true for these more recent RIP conditions in Chapter 2.

We conclude this section by making the important observation that, since IHT algorithms are gradient projection algorithms, they share many similarities with the gradient-based algorithms for l_1-minimization which were surveyed in Section 1.6.1 and which are frequently used by practitioners. In particular, each iteration of an IHT algorithm requires only a handful of

matrix-vector products and a hard threshold projection, and so the computational cost for IHT is similar to that for l_1-based gradient methods.

1.8 Beyond sparsity: background

We have seen that what enables the recovery of signals from undersampled linear measurements is the use of the prior assumption that the signal is sparse in some basis. Sometimes, however, a more refined signal model is more appropriate. For example, in image compression, the JPEG2000 standard exploits the fact that the values and locations of wavelet coefficients of natural images tend to have a particular structure. From the point of view of signal recovery, it may be vital to consider a more refined model in order to obtain a sensible solution. Furthermore, because a structured sparsity model represents more prior information about the signal, recovery is often possible from a reduced number of measurements.

Nonnegativity and other additional constraints. The use of additional assumptions in sparse inverse problems certainly did not originate with CS. Even as far back as 1907, a result of Carathéodory [40, 41] implies that the assumption of nonnegativity can dramatically reduce the number of measurements needed to recover a signal which has a sparse Fourier transform: provided the nonzero Fourier coefficients are positive, $n = 2k + 1$ equally-spaced measurements uniquely determines the signal.

The first theoretical CS results for the case of nonnegative coefficients were obtained in 2005 by Donoho and Tanner [75, 76]. The authors showed in [75] that l_1-l_0 equivalence for nonnegative signals is equivalent to k-neighbourliness of the projected nonnegatively-constrained l_1-ball, which is in fact another regular polytope, the *simplex*. We recall from Section 1.5 that both a strong and weak notion of neighbourliness can be considered, where strong neighbourliness means that all k-faces survive projection and weak neighbourliness means that the majority of k-faces survive projection. In the same proportional-dimensional asymptotic with which we are now familiar, Vershik and Sporyshev derived the weak phase transition $\hat{\rho}_W^+(\delta)$ for the simplex way back in 1992 [159], without appreciating its implication for l_1-l_0 equivalence at the time. The authors of [76] also derived a strong neighbourliness threshold $\hat{\rho}_S^+(\delta)$, which the authors went on to show to be a sharp phase transition in [77]. These results, valid for random orthogonal or Gaussian measurement matrices, establish precise quantitative limits on undersampling for nonnegative signals. Both phase transitions are higher than the respective phase transitions for the unconstrained case, which implies that fewer measurements need to be taken, which is entirely as expected given that the nonnegativity constraint represents additional prior information about the signal.

In [79], the same authors analyzed the neighbourliness of the positive orthant $\{x \in \mathbb{R}^N : x \geq 0\}$ when projected by means of a matrix which is in general position and has centrosymmetric ex-

changeable columns almost surely, and derived phase transitions in the (δ, ρ)-plane below which a system of linear equations has a unique solution under nonnegativity constraints – namely the original sparse nonnegative solution x^* from which the measurements were generated. Notably, these results apply not just to Gaussian matrices, but to a wide range of matrices satisfying the above assumptions. Results were also obtained for the hypercube $\{x \in \mathbb{R}^N : 0 \leq x \leq 1\}$, under the assumption that the projection matrix is in general position. Within these regions of phase space, therefore, signals with nonnegativity or box constraints can be recovered using any algorithm capable of generating a feasible point of the problem.

Block sparsity. In the decade prior to the emergence of CS, techniques were already being developed for recovering signals in which the locations of the nonzero coefficients have a particular structure. For example, motivated by an application in brain imaging, Rao [137] proposed a variant of the IRLS algorithm FOCUSS [101] in 1996 which was adapted to recover *block-sparse* signals, in which the nonzero coefficients cluster together in blocks. Around the time that CS was emerging, various extensions of existing algorithms for CS with simple sparsity were proposed for recovering block-sparse signals, including a variant of basis pursuit in which the l_1 vector norm is replaced with the mixed l_p/l_1 norm for some $p > 0$ [146, 55], and the simultaneous OMP (S-OMP) algorithm [155]. Tropp [155, 156], Chen and Huo [50] and Eldar et al. [88] extended existing coherence-based recovery conditions for basis pursuit and OMP, obtaining improved recovery guarantees for their respective block-sparse variants.

In 2008, Eldar and Mishali introduced the concept of *block-RIP* [89], in which RIP bounds are only required to hold for vectors with valid supports under the block-sparse model. The authors showed that robust recovery by l_2/l_1-minimization is guaranteed under the same RIP condition as in [32], where standard RIP constants are replaced by their block-based equivalents. Furthermore, it is shown that the number of measurements needed to guarantee recovery from random measurements (i.i.d. subgaussian entries) satisfies

$$n = \mathcal{O}\left[k + \frac{k}{J}\log\left(\frac{N}{k}\right)\right], \tag{1.28}$$

where J is the number of blocks in the signal model. The reason that it is possible to obtain improved results is that RIP bounds are now only required to hold for vectors with certain supports, namely those which satisfy the block-sparsity model. This result suggests that l_2/l_1-minimization might boast an improved recovery phase transition over l_1-minimization in the (δ, ρ) proportional-growth asymptotic. In this respect, Stojnic et al. [143] showed that, given $\tau > 0$, for fixed J satisfying $J > 1/(1-\delta)$ and $J > \log(1/\tau)/\tau$, a lower bound on the strong phase transition is given by $\hat{\rho}_S(\delta) = 1/2 - \tau$. Since $\rho < 1/2$ is required to ensure a unique solution to the system of equations (see Lemma 1.7), this means that the highest possible strong phase transition is achieved provided the size of the blocks is suitably large.

In 2010, Baraniuk et al. proposed and analysed algorithms within a general *model-based* framework for structured sparsity [7]. It is observed that any CS algorithm which uses hard threshold operations to perform sparse projections may be mutated into an algorithm for structured sparse recovery by replacing the hard threshold with the projection onto the feasible set for the model. Based upon this observation, the authors proposed block-based versions of CoSaMP [127] and IHT [26]. The authors obtain recovery guarantees for these algorithms by substituting block-based RIP constants into the same RIP conditions previously given for the case of simple sparsity [92, 93]. The authors also extend their results to signals which only approximately fit the structured sparsity model. For random matrices with subgaussian entries, it follows that recovery is guaranteed for both algorithms provided the number of measurements satisfies (1.28).

Other refined sparsity models. This thesis will address another form of structured sparsity: *tree-based models*, to which we devote Section 1.9. We have sought to describe the refinements to simple sparsity which have been the subject of most research attention, and for which improved theoretical results have been obtained. Mention should also be given to the recent work by Duarte et al. on *spectral compressive sensing* [81], where signals are modelled as sparse in the Fourier domain, with the additional restriction that selected basis vectors are sufficiently uncorrelated. In a different direction, a weighted l_1-minimization method has been proposed by Khajehnejad et al. [111] for recovering signals with nonuniform sparsity, where coefficients are assigned different prior probabilities of being nonzero. Research into more refined sparsity models is likely to continue to be driven by the widely varying needs of practitioners. We have concentrated on refinements to sparsity models, but it should also be added that other models of low signal information content exist. In particular, there has been an explosion of interest in modelling signals as low-rank matrices [34] in the last five years, an emerging area of research with many connections to CS theory.

1.9 Tree-based models for CS

The advent of the wavelet transform in the 1980s has undoubtedly revolutionized signal processing. In particular, it is now well-known that wavelets provide sparse representations for piecewise-smooth signals [123]. They are widely used in many applications, especially in image processing, and they provide the basis for the JPEG-2000 compression standard [141]. Wavelet coefficients have a *multiscale* tree structure, in which each coefficient has a single 'parent' coefficient and a small number of 'children' coefficients. Furthermore, wavelets coefficient tend to exhibit the property of *persistence across scales*, meaning that large coefficients tend to be propagated down branches of the tree. It follows that the large wavelet coefficients of piecewise-smooth signals can often be effectively modelled as forming a connected subtree.

Tree-based sparsity models have been extensively used in the signal processing community ever since they were first proposed in the context of wavelet compression [140]. CS algorithms adapted to tree-based models began to appear in 2005, with Duarte et al. proposing variants of MP and OMP [83], and La and Do proposing a variant of OMP [113]. These algorithms build up the support set of the signal in a greedy manner, using gradient information to select new coefficients, while ensuring that the support set remains a connected subtree at each iteration. In [84], a weighted l_1-minimization strategy was proposed, in which tree-structure dependencies are enforced by means of a weighting scheme based on a Hidden Markov Tree (HMT) model.

In [24], Blumensath and Davies considered a more general model of finite-dimensional unions of subspaces which includes a variety of structured sparsity models, including the tree-based model. A variant of IHT [26] was proposed in which the hard threshold projection $\mathcal{H}_k$ is replaced by the projection onto the appropriate union of subspaces. The authors also introduced the *model-based RIP*, showing that any existing RIP condition guaranteeing recovery using IHT [26] also holds for the model-based version of IHT, upon the substitution of model-based RIP constants into the condition. Moreover, they argued that such results represent a qualitative improvement since the model imposes a restriction on the number of permissible support sets.

Tree-based models were also considered within a general model-based framework by Baraniuk et al. in [7] (see Section 1.8). Variants of IHT and CoSaMP [127] were considered in which $\mathcal{H}_k$ is replaced by the projection onto the set of vectors supported on connected subtrees of cardinality k. Under the assumption that the projection is computed exactly, explicit robust recovery conditions for both algorithms were obtained in terms of the *tree-based RIP*. In fact, it follows from their analysis that any RIP conditions for recovery using these algorithms can be extended in this way, including the state-of-the-art results at the time of writing in [93] and [92]. It is observed that tree-based RIP is not sufficient to ensure that the results extend to signals which are only approximately tree-sparse, but it is shown that the results can be extended in this way provided the measurement matrix satisfies a certain *Restricted Amplification Property* (RAmP). For random matrices with subgaussian entries, it is proved that both tree-based RIP and the RAmP hold, and therefore recovery is guaranteed for both algorithms, with overwhelming probability provided the number of measurements satisfies $n = \mathcal{O}(k)$. Comparing this result with (1.9), we observe an improvement in the form of the removal of the log-factor. In the phase transition framework, the result also suggests that recovery is asymptotically determined by the oversampling factor $\rho = k/n$ alone, and not by the undersampling factor $\delta = n/N$. However, the result exhibits the same ambiguity that was observed in Section 1.4: it is equally computationally intractable to calculate tree-based RIP constants, leaving unanswered the quantitative question as to the constant of proportionality involved.

In addition to IHT for standard CS, this thesis focusses upon its tree-based variant just described, which we will refer to as Iterative Tree Projection (ITP). The algorithm differs from

IHT only in the projection, so that one proceeds by means of the iteration

$$x^{m+1} := \mathcal{P}_k\{x^m + \alpha^m A^T(b - Ax^m)\}, \tag{1.29}$$

where $\mathcal{P}_k(\cdot)$ is the Euclidean projection onto the set of vectors supported on connected trees of cardinality k. See Section 6.2 for a more precise definition of $\mathcal{P}_k(\cdot)$ in the context of dyadic wavelet trees. ITP may be viewed as a gradient projection algorithm for the tree-based l_0-problem

$$\min_{x\in\mathbb{R}^N} \Psi(x) \quad \text{subject to} \quad x \in \mathcal{T}_k, \tag{1.30}$$

where $\Psi(\cdot)$ is defined in (1.24).

The results in [7] make the assumption that it is possible to perform the exact tree projection $\mathcal{P}_k(\cdot)$. Various strategies exist for approximately computing this projection, including the *Condensing Sort and Select Algorithm* (CSSA) [8] and an approach based upon Lagrangian relaxation [63, 69]. However, to the best of our knowledge, no algorithm with sub-exponential complexity has been shown to exactly perform tree projection for a given sparsity k.

1.10 Outline and contribution of the thesis

In Chapter 2, we first bring the quantitative analysis of RIP conditions for IHT algorithms up to date, by quantifying the current state-of-the-art RIP conditions for IHT [93] in the phase transition framework for Gaussian measurement matrices, and by obtaining the first such phase transition for the variable stepsize variant NIHT. We show that, while there is an improvement in the phase transition for IHT, the results are still pessimistic compared to average-case behaviour, and especially so for NIHT. The chapter concludes with some considerations on the choice of support sizes in RIP conditions.

RIP analysis is by nature worst-case analysis and results are therefore limited by the strong phase transition. In contrast, we have seen in Section 1.5 that it is the weak phase transition which captures the average-case, practical case of interest. The main contribution of this thesis is to pioneer of a new method of recovery analysis which leads to average-case recovery guarantees and lower bounds on a particular kind of weak phase transition for the case of Gaussian matrices. By making the realistic assumption that the original signal and measurement matrix are independent, we derive phase transitions for both IHT and NIHT which more closely describe the true performance of the algorithm.

In Chapter 3, we first present our new analysis in the more general context of arbitrary measurement matrices. Our approach, in contrast to previous RIP-based analysis [23, 96, 92, 93, 27], considers the fixed points of IHT. We derive necessary conditions for IHT to have a fixed point on a given support set, and we prove convergence of IHT to some fixed point under an

RIP condition significantly weaker than those that have appeared previously. We also extend our analysis to the variable step-size NIHT by generalizing the notion of a fixed point to a *stable point*. While our main aim is to obtain average-case recovery guarantees, for the sake of comparison we conclude the chapter by translating our stable point analysis into worst-case RIP-based recovery conditions for both IHT and NIHT.

Under the assumption of independence between the original signal and measurement matrix, the necessary conditions for a stable point derived in Chapter 3 turn out to be especially amenable to probabilistic analysis in the case of Gaussian measurement matrices. To perform this analysis, we require some results concerning distributions of certain Rayleigh quotients, and large deviations results for certain Gaussian-related random variables in the proportional-dimensional asymptotic introduced in Section 1.5, and we develop these tools in Chapter 4.

We begin Chapter 5 by quantifying in the phase transition framework the worst-case RIP-based recovery conditions derived from our stable point analysis in Chapter 3, and by motivating the need for average-case analysis. We then apply the probabilistic tools developed in Chapter 4 to the conditions derived in Chapter 3, obtaining lower bounds on a weak phase transition guaranteeing exact recovery of the original signal in the case of noiseless measurements and exactly sparse signals. We also extend our results to the more realistic model of noisy measurements and compressible signals, showing that the same phase transition guarantees convergence to a stable point which approximates the original signal within some error tolerance, which we quantify as a multiple of the unrecoverable energy due to compression error and measurement noise.

Chapters 6 and 7 extend our recovery analysis of Chapters 3, 4 and 5 to the tree-based setting and ITP algorithms. Our recovery results, in common with previous RIP-based results [24, 7], require that the tree projection in ITP be performed exactly. However, we noted in Section 1.9 that currently available algorithms are only guaranteed to approximately calculate a tree projection for a given sparsity [8, 63, 69]. In Chapter 6, after first giving a more technical introduction to the tree-based signal model and the ITP algorithm, we propose a dynamic programming (DP) algorithm which is guaranteed to perform an exact tree projection, and we also prove that it has low-order polynomial complexity. We then adapt the stable point analysis of Chapter 3 for arbitrary matrices to ITP algorithms.

In Chapter 7, we build upon the stable point analysis of Chapter 6 to obtain quantitative recovery results for ITP algorithms in the context of Gaussian measurement matrices. Since the number of permissible support sets is much reduced in the tree-based model, tighter tail bound results are derived in which there is no dependence upon the δ variable. This allows us to obtain results in a simplified proportional-dimensional asymptotic in which we entirely dispense with the δ variable. We derive critical oversampling thresholds $\hat{\rho}$, such that recovery is guaranteed provided $\rho < \hat{\rho}$. These results give a significant improvement upon those for simple sparsity and IHT algorithms.

Chapter 8 gives final conclusions and potential related research directions for the future. To summarize, the contributions of this thesis are as follows.

- We quantify existing state-of-the-art RIP-based recovery analysis of IHT algorithms in the context of Gaussian measurement matrices by means of the phase transition framework. Our analysis leads to the highest known lower bound on the strong phase transition for recovery using IHT (Chapter 2).
- We introduce an entirely new recovery analysis of IHT algorithms which considers the fixed or *stable* points of the algorithms. We derive from our analysis new RIP conditions guaranteeing robust recovery. We quantify these recovery conditions by means of the phase transition framework for Gaussian matrices, and in doing so obtain the highest known lower bound on the strong phase transition for recovery using NIHT (Chapter 3).
- We obtain the first quantitative average-case recovery guarantees for IHT algorithms in the phase transition framework for Gaussian matrices. We give lower bounds on a particular kind of weak phase transition for recovery using IHT and NIHT, which assumes independence between the signal and measurement matrix. We obtain significant improvements upon the best RIP-based phase transitions for both algorithm variants. Our results consider the realistic model of noisy measurements and inexactly sparse signals, guaranteeing an improved robustness to these inaccuracies. Our results narrow the gap between worst-case recovery guarantees and average-case performance of IHT algorithms (Chapters 4 and 5).
- We extend our recovery analysis to ITP algorithms, the tree-based mutations of IHT algorithms, with the same stepsize schemes (ITP and NITP). We obtain quantitative average-case recovery guarantees for Gaussian matrices by introducing a simplified asymptotic framework. This is the first time that recovery guarantees for a tree-based algorithm have been quantified in this way (Chapters 6 and 7).
- We propose a dynamic programming (DP) algorithm for exact tree projection. We also prove that our algorithm has low-order polynomial complexity (in fact $\mathcal{O}(Nk)$ for binary tree structures), making it the first algorithm with sub-exponential complexity guaranteed to perform an exact tree projection (Chapter 6).

1.11 An illustration: the single-pixel camera

As an illustration, we conclude this first chapter by describing one particular example of an application of Compressed Sensing in the context of imaging. In 2008, Duarte et al. [82] proposed a new design for a camera which exploits CS theory. In contrast to a conventional camera which would have a vast array of photon detectors – one for each pixel – the single-pixel

camera is so-called because it uses only a single photon detector or 'pixel'. The incident light field is directed onto a Digital Micromirror Device (DMD) which consists of an array of tiny mirrors, one corresponding to each pixel. Each mirror can be oriented in one of two directions: the 'on' position directs the light for that pixel towards the single detector, while the 'off' position directs the light away from the detector. The light from all the pixels set to the 'on' position is then summed and a measurement is recorded as an output voltage on the photon detector. The measurement is then subsequently digitized. A series of measurements can be obtained by flipping the mirrors and repeating the process a number of times. See Figure 1.5 for a diagram of the camera design.

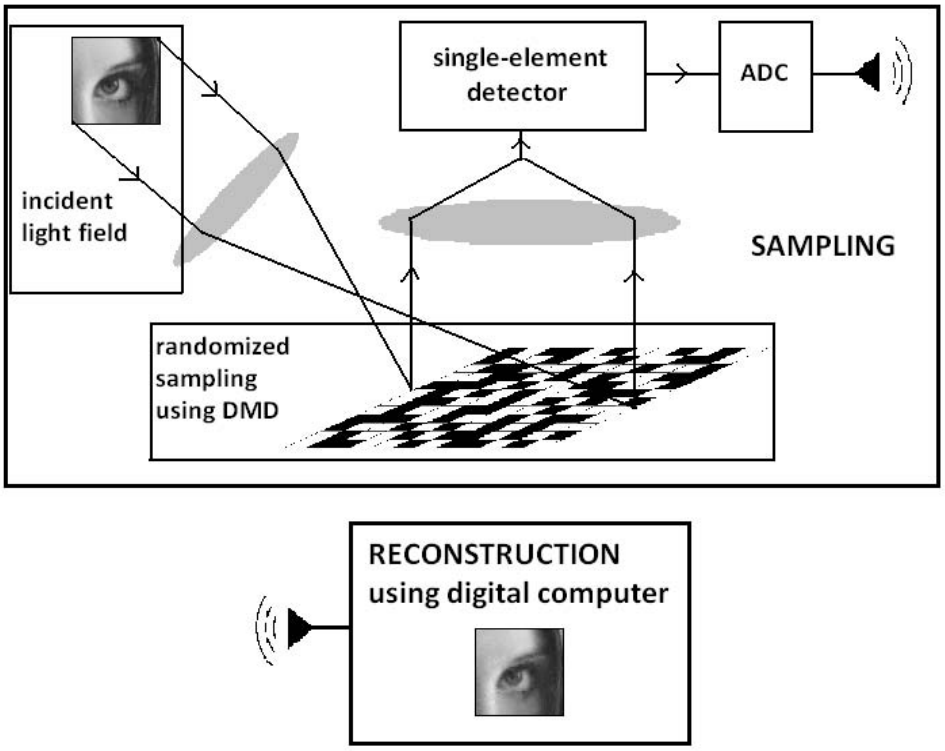


Figure 1.5: A schematic diagram of the Rice CS single-pixel camera.

The concept of a single-pixel camera is not a new one, and it fits into the broad category of multiplexing imaging methods in which a series of consecutive measurements are made by a single detector. What sets the CS single-pixel camera model apart is a novel sampling approach which means that it is possible to take fewer measurements. It therefore offers an alternative approach to obtaining compressed images: rather than taking a full set of samples and subsequently compressing, compressed samples are acquired in the first place. To achieve this, CS theory motivates the use of a random sampling procedure in which each mirror is set randomly to either the 'on' or 'off' position with equal probability.

Modelling the incident light field as a discrete pixelated array consisting of N pixels, we may also represent this 'original image' as a vector $z^* \in \mathbb{R}^N$. By means of the DMD, the light

from all the pixels set to the 'on' position is summed to give a single measurement. We may model this summation as an inner product of the original image z with some vector $\phi_i \in \mathbb{R}^N$ consisting of random ones and zeros. Sampling error is likely to occur, particularly as a result of photon counting noise, and subsequently due to quantization error in the digitization, which may be modelled as additive noise. We may therefore write

$$b_i = \phi_i^T z^* + e_i.$$

In keeping with the CS ethos, we choose to undersample the image and take n such measurements, where $n \ll N$. Writing Φ for the $n \times N$ matrix whose rows are the vectors $\{\phi_i^T\}$, we may represent the entire measurement process by the matrix equation (1.6). In this case, the matrix Φ is a random Bernoulli matrix consisting of equiprobable ones and zeros.

We note that the sampling model here described fits precisely into the general framework of undersampled linear measurements introduced in Section 1.3. Furthermore, many images are compressible in some appropriate transform basis, such as a wavelet basis, and so it is quite realistic to assume that the image in question may be compressible. The single-pixel camera therefore represents a very natural application of CS theory. Furthermore, if an image is compressible using a wavelet transform, it is also likely that its wavelet coefficients have a rooted tree structure, which suggests that improved recovery may be possible by using a tree-based recovery algorithm.

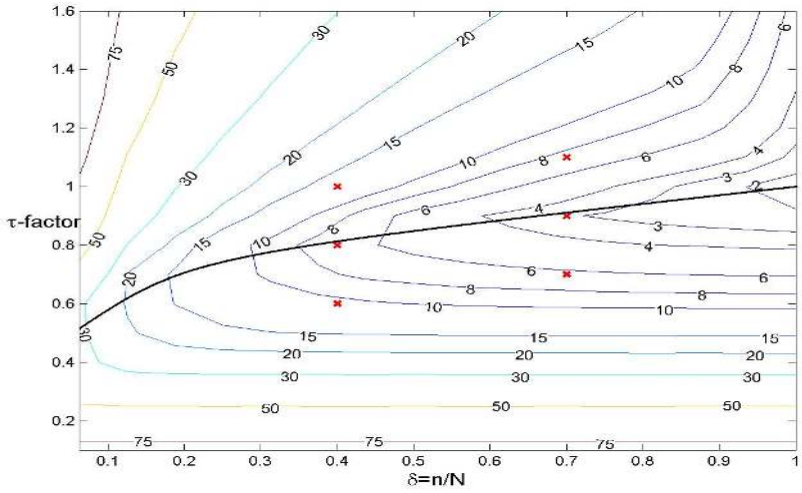
During the 2nd year of my PhD studies, I undertook an internship in the defence technology company *SELEX Galileo Ltd.* The aim of the project [149, 150] was to explore the potential of the single-pixel camera by building a computer model of the camera design, incorporating several CS recovery algorithms, and subjecting the model to extensive numerical testing. The final model offered a variety of choices for the measurement scheme (including Bernoulli matrices and Gaussian matrices), the sparsifying transform (including several commonly-used wavelet transforms) and also the recovery algorithm. Three CS recovery algorithms were tested, two of which are the subject of recovery analysis in this thesis: NIHT [27] (see Sections 1.7 and 2.1) and its tree-based variant normalised ITP [7] (see Sections 1.9 and 6.1). Tests were also carried out upon a gradient projection method for l_1-minimization, based upon the SPGL1 method [157] (see Section 1.6), which we refer to as l_1-*projection*.

Reflecting the phase transition framework with which we are now familiar, there are two parameters which the user is free to vary: the undersampling ratio $\delta = n/N$ and the algorithm's tuning parameter. For NIHT and NITP, the tuning parameter is simply the assumed sparsity of the image, whereas for l_1-projection the tuning parameter τ similarly constrains the sparsity of the algorithm's output. We report here the results of systematic testing throughout this two-dimensional parameter space, which may be instructively compared with the phase transitions from Section 1.5. Recovery of a test image by each of the three algorithms was tested by

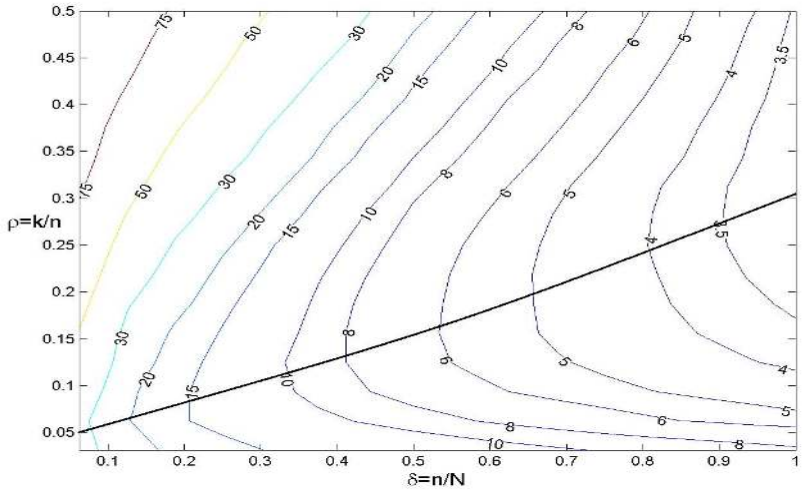
means of multiple trials on a mesh of equally-spaced grid points in the parameter space. In the experiments which we report here, the measurement scheme was a close variant of the Bernoulli sampling scheme in which each entry of the measurement matrix is ± 1 with equal probability, and Daubechies 9-7 wavelets [59] – on which the JPEG2000 compression standard is partly based [141] – were used as the sparsifying transform.

Figure 1.6 gives plots of recovery accuracy in the form of the *root mean-squared error* (RMSE) of the algorithm output in relation to the original image data used to generate the measurements. We observe an *optimal tuning* behaviour, namely that for a given undersampling ratio δ, there exists some value of the respective tuning parameters which minimize the RMSE. Underlying phase transitions provide the explanation for the optimal tuning behaviour: below the optimal tuning curve, the algorithm has good recovery performance, meaning that results improve as the sparsity is increased; however, as the sparsity is increased such that the phase transition is breached, recovery performance begins to deteriorate. Tracing out the optimal tuning curve (the superimposed black curve) from right to left, we see that recovery accuracy degrades gracefully as the image is more aggressively undersampled, and that similar recovery performance may be observed for each of the algorithms. We wish to emphasize in particular that NIHT and NITP, which are not based upon l_1-minimization, have competitive recovery performance with l_1-projection. We also observe that the tree-based NITP gives a modest improvement for small values of δ, with the optimal tuning curve being higher and exhibiting less dependence upon δ, which is in keeping with the results alluded to in Section 1.9.

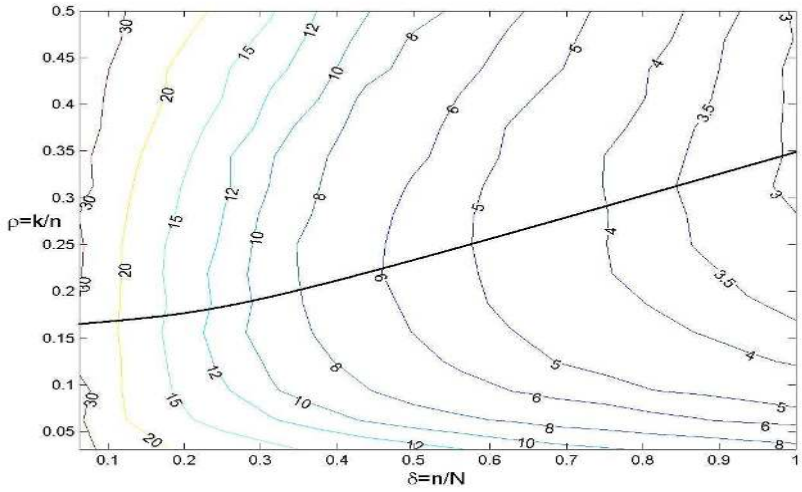
In conclusion, based upon my experience fully documented in [149, 150], the CS single-pixel camera is a promising new technology, offering the potential for image recovery from significantly fewer measurements. Furthermore, the effectiveness of CS recovery algorithms, including NIHT and NITP, was empirically demonstrated.



(a)



(b)



(c)

Figure 1.6: Phase plots of average RMSE for a 64x64 test image (a) l_1-projection (b) NIHT (c) NITP.

Chapter 2

RIP-based phase transitions for IHT/NIHT

In subsequent chapters of this thesis, by introducing a new method of analysis, we will derive lower bounds on the weak recovery phase transition for IHT [26] and its variant NIHT [27] in the case of Gaussian matrices. In order to show that our approach leads to improved recovery guarantees, it is necessary to compare our phase transitions with those presently obtainable from RIP analysis. The RIP condition for IHT obtained by Blumensath and Davies in [23] was analysed in the phase transition framework in [17] by Blanchard et al., and the resulting phase transition was presented in Section 1.4 (see Figure 1.3). However, the RIP condition for NIHT derived in [27] has not been analysed in the phase transition framework for Gaussian matrices. Furthermore, since the work in [17], other RIP-based recovery conditions have been derived for constant stepsize IHT [96, 92, 93], and these conditions must also be examined in the same framework in order to determine the current best recovery results based upon RIP analysis. We address both of these issues in the present chapter.[1] We begin in Section 2.1 by formally introducing the generic IHT algorithm and the IHT and NIHT stepsize schemes. Then, in Section 2.2, we extend a recent analysis of IHT with unit stepsize and noiseless measurements by Foucart [93] in several directions: we consider an arbitrary constant stepsize and noisy measurements, we weaken the condition by asymmetrizing the RIP constants, and we also extend the result to NIHT. In Section 2.3, we translate the conditions derived in Section 2.2 into the phase transition framework for Gaussian matrices, showing that the resulting phase transitions are higher than those arising from previous analysis, including the analysis of NIHT in [27]. We conclude the chapter with a discussion of the results in Section 2.4, and some observations concerning support sizes of RIP constants in Section 2.5.

[1] Material in this chapter applies similar proof techniques to those published in [17], which was a joint authorship with J. Blanchard, C. Cartis and J. Tanner, in order to quantify more recent RIP analysis.

2.1 Stepsize schemes for IHT

We recall from Section 1.3 that the problem we are interested in is the recovery of a compressible signal $x^* \in \mathbb{R}^N$ from the noisy linear measurements $b = Ax + e \in \mathbb{R}^n$, where $A \in \mathbb{R}^{n \times N}$ is the measurement matrix, and where $e \in \mathbb{R}^n$ is a noise vector. Since the signal is compressible, it can be well-approximated by a k-sparse vector for some sparsity k, which motivates seeking to recover the signal by solving the *l_0-problem* (1.21). The generic family of IHT algorithms, which solves (1.21) by gradient projection, has already been described in Section 1.7, but we now summarize it more formally in Algorithm 2.1.

Algorithm 2.1 Generic IHT [26]

Inputs: A, b, k.
Initialize $x^0 = 0,\ m = 0$.
While some termination criterion is not satisfied, do:

1. Compute a stepsize α^m.

2. Compute $x^{m+1} := \mathcal{H}_k\left\{x^m + \alpha^m A^T(b - Ax^m)\right\}$,
where $\mathcal{H}_k(\cdot)$ is defined in (1.25).

3. Set $m := m + 1$.

End; output $\hat{x} = x^m$.

Often in this thesis, we will consider the infinite sequence of iterates generated by IHT, but in practice a termination criterion would need to be employed. There are various possibilities for a practical termination criterion, such as requiring that the iterates approach each other, *i.e.* $\|x^m - x^{m-1}\| \leq \eta$ for some $\eta > 0$, or requiring that a good approximation to the linear system be obtained, *i.e.* $\|b - Ax^m\| \leq \eta$ for some $\eta > 0$.

For Algorithm 2.1 to be well-defined, it remains to define a stepsize scheme $\{\alpha^m\}$, and we will consider two options, which give rise to the IHT (*constant* stepsize) and NIHT (*normalised* stepsize) variants respectively.

Algorithm 2.2 IHT [23]

Given some $\alpha > 0$, on **Step 1** of each iteration $m \geq 0$ of generic IHT, set

$$\alpha^m := \alpha. \tag{2.1}$$

The NIHT stepsize choice comprises two steps. Firstly, one initializes with the *exact line-search* criterion [130], which ensures maximum decrease in the objective of (1.21) provided the support set remains unchanged. However, this choice may not even ensure a decrease in the objective of (1.21) in the case when there is a change in the support set. For this reason, a second *backtracking* step is used in which the stepsize is repeatedly shrunk by some shrinkage parameter $\kappa > 1$ until a condition is satisfied which guarantees a decrease in the objective of (1.21). In practice, the choice of κ constitutes a trade-off between recovery performance and computa-

Algorithm 2.3 NIHT [27]

Given some $c \in (0,1)$ and $\kappa > 1/(1-c)$, on **Step 1** of each iteration $m \geq 0$ of generic IHT, do:

1.1. Exact linesearch.

(a) Set $\Gamma^m := \text{supp}(x^m)$.

(b) Compute

$$\alpha^m := \frac{\|A_{\Gamma^m}^T(b - Ax^m)\|^2}{\|A_{\Gamma^m}A_{\Gamma^m}^T(b - Ax^m)\|^2}. \tag{2.2}$$

(c) Let $\tilde{x}^{m+1} := \mathcal{H}_k\left\{x^m + \alpha^m A^T(b - Ax^m)\right\}$.

1.2. Backtracking. If $\text{supp}(\tilde{x}^{m+1}) = \text{supp}(x^m)$, end; output α^m.
Else, while $\alpha^m \geq (1-c)\frac{\|\tilde{x}^{m+1} - x^m\|^2}{\|A(\tilde{x}^{m+1} - x^m)\|^2}$, do:

(a) $\alpha^m := \alpha^m/(\kappa(1-c))$.

(b) $\tilde{x}^{m+1} := \mathcal{H}_k\left\{x^m + \alpha^m A^T(b - Ax^m)\right\}$.

End; output α^m.

tional efficiency: for optimal performance, κ close to 1 should be chosen, while increasing κ will lead to fewer backtracking steps, making the algorithm more computationally efficient. The backtracking step ensures a potentially desirable property of the NIHT algorithm, namely that, provided the measurement matrix satisfies mild linear independence assumptions, it is guaranteed to converge [27]. However, it is also possible to conceive of a variant of NIHT in which the backtracking step is entirely omitted and in which the exact linesearch stepsize is always selected. Furthermore, RIP-based recovery guarantees have also been proven for this simplified variant. We will make reference to these results later in the chapter, while largely restricting our focus to the algorithm originally proposed in [27] in which backtracking is included.

Bounds on the NIHT stepsize are given in the following lemma.

Lemma 2.1 (**NIHT stepsize bounds**). *Suppose Assumption 1 holds and let α^m be chosen according to Algorithm 2.3. Then*

$$\frac{1}{\kappa(1 + U_{2k})} \leq \alpha^m \leq \frac{1}{1 - L_k}. \tag{2.3}$$

Proof: If (2.2) is accepted, then $\alpha^m \leq 1/(1 - L_k)$ by (1.15), which is well-defined by Assumption 1. On the other hand, if (2.2) is rejected, the backtracking phase can only reduce the stepsize further, which proves the upper bound in (2.3). To prove the lower bound, we also distinguish two cases. If (2.2) is accepted, then $\alpha^m \geq 1/(1 + U_k)$ by (1.15). Since $\kappa > 1$, and since $U_{2k} \geq U_k$ by the nonincreasing property of RIP constants, the lower bound in (2.3) holds in this case. On the other hand, if (2.2) is rejected, the penultimate stepsize calculated in the backtracking phase must also have been rejected. Writing $\tilde{\alpha}^m$ for the penultimate stepsize,

since $\tilde{\alpha}^m$ was rejected, we have

$$\tilde{\alpha}^m \geq (1-c)\frac{\|\tilde{x}^{m+1} - x^m\|^2}{\|A(\tilde{x}^{m+1} - x^m)\|^2} \geq \frac{1-c}{1+U_{2k}}, \tag{2.4}$$

where the last step follows from (1.15). But $\alpha^m = \tilde{\alpha}^m/[\kappa(1-c)]$, which combines with (2.4) to give the lower bound in (2.3) in this case also. □

It follows that, for both IHT and NIHT, there exist lower and upper bounds on the stepsize, $\underline{\alpha} > 0$ and $\overline{\alpha} > 0$ respectively, such that $\underline{\alpha} \leq \alpha^m \leq \overline{\alpha}$ for all $m \geq 0$, by (2.3) in the case of NIHT, and trivially by (2.1) in the case of IHT.

2.2 An RIP analysis of IHT for arbitrary matrices

In this section, we obtain robust recovery guarantees in terms of asymmetric RIP constants for both IHT and NIHT, thereby extending the noiseless, symmetric RIP analysis of IHT by Foucart in [93].

Before beginning our analysis, we first present a lemma which allows us to restrict our attention to the case of exactly sparse signals. The result, which bounds the factor by which the measurement matrix amplifies nonsparse vectors, is proved in [127].

Lemma 2.2 (**Amplification bound [127, Proposition 3.5]**). *Given some positive integer s, suppose that $A \in \mathbb{R}^{n\times N}$ has upper RIP constant U_s, as defined in (1.6). Then, for any $x \in \mathbb{R}^N$,*

$$\|Ax\| \leq \sqrt{1+U_s}\left[\|x\| + \sqrt{\frac{1}{s}}\|x\|_1\right]. \tag{2.5}$$

It was shown in [127] that Lemma 2.2 allows the noisy measurements of a compressible signal to be viewed as measurements of a sparse signal with a different noise vector which incorporates the compression error. To see this, let x_k^* be the best k-sparse approximation to the original signal x^* as defined in (1.13), and let us define

$$E := A(x^* - x_k^*) + e. \tag{2.6}$$

Then we have

$$b = Ax^* + e = Ax_k^* + A(x^* - x_k^*) + e = Ax_k^* + E,$$

and Lemma 2.2 implies that

$$\begin{aligned} \|E\| &= \|A(x^* - x_k^*) + e\| \leq \|e\| + \|A(x^* - x_k^*)\| \\ &\leq \|e\| + \sqrt{1+U_k}\left[\|x^* - x_k^*\| + \sqrt{\tfrac{1}{k}}\|x^* - x_k^*\|_1\right], \end{aligned}$$

which may be viewed as a bound for $\|E\|$ in terms of a measure of the unrecoverable energy of the signal, due both to measurement noise and the signal being only approximately sparse. It follows that any result for exactly sparse signals which bounds the approximation error in terms of $\|e\|$ may be extended to compressible signals. Therefore, in this chapter, we restrict our attention to the case in which noisy measurements are taken of an exactly k-sparse signal.

Notation. Before proceeding, we introduce some notation that will be used throughout the rest of the thesis. Given some index set $\Gamma \subseteq \{1, 2, \dots N\}$, we define the complement of Γ to be $\Gamma^C = \{1, 2, \dots N\} \setminus \Gamma$. We write x_Γ for the restriction of the vector x to the coefficients indexed by the elements of Γ, and we write A_Γ for the restriction of the matrix A to those columns indexed by the elements of Γ. When it exists, we denote by $A_\Gamma^\dagger$ the Moore-Penrose pseudoinverse, namely

$$A_\Gamma^\dagger := (A_\Gamma^T A_\Gamma)^{-1} A_\Gamma^T. \tag{2.7}$$

The following lemma gives some further consequences of the RIP and is proved in [16].

Lemma 2.3 (**Consequences of the RIP [16, Lemma 15]**). *Given some positive integer s, suppose that $A \in \mathbb{R}^{n \times N}$ has lower and upper RIP constants L_s and U_s respectively, as defined in (1.6). Let Ω be a set of cardinality s, and let $\Omega = \Omega_1 \cup \Omega_2$ where $|\Omega_1| = s_1$, $|\Omega_2| = s_2$ and $s = s_1 + s_2$. Then*

$$\|A_\Omega^T y\| \leq \sqrt{1 + U_s}\|y\| \quad \textit{for all } y \in \mathbb{R}^n; \tag{2.8}$$

$$(1 - L_s)\|x\| \leq \|A_\Omega^T A_\Omega x\| \leq (1 + U_s)\|x\| \quad \textit{for all } x \in \mathbb{R}^s; \tag{2.9}$$

$$\frac{1}{1 + U_s}\|x\| \leq \|(A_\Omega^T A_\Omega)^{-1} x\| \leq \frac{1}{1 - L_s}\|x\| \quad \textit{for all } x \in \mathbb{R}^s; \tag{2.10}$$

$$\|A_\Omega^\dagger y\| \leq \frac{1}{\sqrt{1 - L_s}}\|y\| \quad \textit{for all } y \in \mathbb{R}^n, \ \textit{provided } A_\Omega^\dagger \textit{ is well-defined}; \tag{2.11}$$

$$\|A_{\Omega_1}^T A_{\Omega_2} z\| \leq \frac{1}{2}(L_s + U_s)\|z\| \quad \textit{for all } z \in \mathbb{R}^{s_2}; \tag{2.12}$$

$$\|(I - \omega A_\Omega^T A_\Omega)x\| \leq \max\{\omega(1 + U_s) - 1, 1 - \omega(1 - L_s)\}\|x\| \quad \textit{for all } x \in \mathbb{R}^s \textit{ and all } \omega > 0. \tag{2.13}$$

Next, by largely following the analysis in [93], we use the RIP to obtain a result for generic IHT with bounded stepsize.

Lemma 2.4 (**Iteration invariant for bounded stepsize**). *Consider Problem 1.2. Let the stepsizes of generic IHT satisfy*

$$\underline{\alpha} \leq \alpha^m \leq \overline{\alpha} \tag{2.14}$$

for all $m \geq 0$. Then

$$\|x^{m+1} - x^*\| \leq \sqrt{3}\max\{\overline{\alpha}(1 + U_{3k}) - 1, 1 - \underline{\alpha}(1 - L_{3k})\}\|x^m - x^*\| + \overline{\alpha}\sqrt{3(1 + U_{2k})}\|e\|. \tag{2.15}$$

Proof: Let us write $a^m := x^m + \alpha^m A^T(b - Ax^m)$, which can be rearranged to give

$$a^m = x^m + \alpha^m A^T(Ax^* + e - Ax^m) = x^* + (I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e. \tag{2.16}$$

Let us write $\Lambda = \mathrm{supp}(x^*)$, $\Gamma^m = \mathrm{supp}(x^m)$, $\Gamma^{m+1} = \mathrm{supp}(x^{m+1})$ and let us further define

$$\Omega = \Lambda \cup \Gamma^m \cup \Gamma^{m+1}. \tag{2.17}$$

By (1.26), we have

$$\|a^m_\Lambda\|^2 \leq \|a^m_{\Gamma^{m+1}}\|^2,$$

which cancels to give

$$\|a^m_{\Lambda\setminus\Gamma^{m+1}}\|^2 \leq \|a^m_{\Gamma^{m+1}\setminus\Lambda}\|^2. \tag{2.18}$$

Substituting (2.16) into (2.18) gives

$$\begin{aligned}&\left\|\left\{x^* + (I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e\right\}_{\Lambda\setminus\Gamma^{m+1}}\right\| \\ &\quad\leq \left\|\left\{x^* + (I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e\right\}_{\Gamma^{m+1}\setminus\Lambda}\right\|,\end{aligned}$$

and the triangle inequality, along with $x^*_{\Gamma^{m+1}\setminus\Lambda} = 0$, implies

$$\begin{aligned}&\left\|x^*_{\Lambda\setminus\Gamma^{m+1}}\right\| - \left\|\left\{(I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e\right\}_{\Lambda\setminus\Gamma^{m+1}}\right\| \\ &\quad\leq \left\|\left\{(I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e\right\}_{\Gamma^{m+1}\setminus\Lambda}\right\|.\end{aligned} \tag{2.19}$$

The sets $\Lambda \setminus \Gamma^{m+1}$ and $\Gamma^{m+1} \setminus \Lambda$ are disjoint, and we may therefore apply the Cauchy-Schwarz inequality, namely $(a+b)^2 \leq \sqrt{2}(a^2 + b^2)$, to (2.19), yielding

$$\left\|x^*_{\Lambda\setminus\Gamma^{m+1}}\right\| \leq \sqrt{2}\left\|\left\{(I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e\right\}_{\Lambda\cup\Gamma^{m+1}}\right\|,$$

from which a further application of the triangle inequality and (2.17) leads us to deduce

$$\left\|x^*_{\Lambda\setminus\Gamma^{m+1}}\right\| \leq \sqrt{2}\left\{\left\|(I - \alpha^m A^T_\Omega A_\Omega)(x^m - x^*)_\Omega\right\| + \alpha^m \left\|A^T_{\Lambda\cup\Gamma^{m+1}} e\right\|\right\}. \tag{2.20}$$

Meanwhile, splitting on Γ^{m+1} and $\Lambda \setminus \Gamma^{m+1}$, and using the definition of $\Gamma^{m+1} = \mathrm{supp}(x^{m+1})$,

$$\begin{aligned}\|x^{m+1} - x^*\|^2 &= \left\|(x^{m+1} - x^*)_{\Gamma^{m+1}}\right\|^2 + \left\|(x^{m+1} - x^*)_{\Lambda\setminus\Gamma^{m+1}}\right\|^2 \\ &= \left\|\left\{(I - \alpha^m A^T A)(x^m - x^*) + \alpha^m A^T e\right\}_{\Gamma^{m+1}}\right\|^2 + \left\|x^*_{\Lambda\setminus\Gamma^{m+1}}\right\|^2,\end{aligned}$$

where the second inequality follows from (2.16). We then apply the triangle inequality and

(2.17) to deduce

$$\begin{aligned}\|x^{m+1}-x^*\|^2 &\leq \left\{\left\|\left\{(I-\alpha^m A^TA)(x^m-x^*)\right\}_{\Gamma^{m+1}}\right\| + \left\|\left\{\alpha^m A^Te\right\}_{\Gamma^{m+1}}\right\|\right\}^2 + \left\|x^*_{\Lambda\setminus\Gamma^{m+1}}\right\|^2 \\ &\leq \left\{\left\|(I-\alpha^m A_\Omega^TA_\Omega)(x^m-x^*)_\Omega\right\| + \alpha^m\left\|A^T_{\Lambda\cup\Gamma^{m+1}}e\right\|\right\}^2 + \left\|x^*_{\Lambda\setminus\Gamma^{m+1}}\right\|^2. \quad (2.21)\end{aligned}$$

Substituting (2.20) into (2.21) then gives

$$\|x^{m+1}-x^*\|^2 \leq 3\left\{\left\|(I-\alpha^m A_\Omega^TA_\Omega)(x^m-x^*)_\Omega\right\| + \alpha^m\left\|A^T_{\Lambda\cup\Gamma^{m+1}}e\right\|\right\}^2. \quad (2.22)$$

Since $|\Omega| \leq 3k$ and $|\Lambda\cup\Gamma^{m+1}| \leq 2k$, the result now follows by applying (2.8), (2.13) and (2.14) to (2.22), and taking square roots. □

Both the IHT and NIHT stepsize schemes have bounded stepsizes, and we may therefore deduce the following results.

Theorem 2.5 (**Iteration invariant for IHT**). *Consider Problem 1.2. Then the iterates of IHT with stepsize α satisfy*

$$\|x^{m+1}-x^*\| \leq \mu_k^{IHT_\alpha}\|x^m-x^*\| + \xi_k^{IHT_\alpha}\|e\|, \quad (2.23)$$

where

$$\mu_k^{IHT_\alpha} := \sqrt{3}\max\{\alpha(1+U_{3k})-1, 1-\alpha(1-L_{3k})\} \quad (2.24)$$

and

$$\xi_k^{IHT_\alpha} := \alpha\sqrt{3(1+U_{2k})}. \quad (2.25)$$

Proof: For IHT with stepsize α, we have $\underline{\alpha} = \alpha$ and $\overline{\alpha} = \alpha$, and the result follows by applying Lemma 2.4. □

Theorem 2.6 (**Iteration invariant for NIHT**). *Consider Problem 1.2 and suppose Assumption 1 holds. Then the iterates of NIHT with shrinkage parameter κ satisfy*

$$\|x^{m+1}-x^*\| \leq \mu_k^{NIHT_\kappa}\|x^m-x^*\| + \xi_k^{NIHT_\kappa}\|e\|, \quad (2.26)$$

where

$$\mu_k^{NIHT_\kappa} := \sqrt{3}\max\left\{\frac{1+U_{3k}}{1-L_k}-1, 1-\frac{1-L_{3k}}{\kappa(1+U_{2k})}\right\} \quad (2.27)$$

and

$$\xi_k^{NIHT_\kappa} := \frac{\sqrt{3(1+U_{2k})}}{1-L_k}. \quad (2.28)$$

Proof: For a given $\kappa > 1$, the stepsize bounds (2.3) apply to NIHT, and the result follows by applying Lemma 2.4 with $\underline{\alpha} := 1/(1-L_k)$ and $\overline{\alpha} := 1/[\kappa(1+U_{2k})]$. □

In order to prove recovery results, we will need the following lemma.

Lemma 2.7. *Suppose there exist $\mu \in [0,1)$ and $\xi > 0$ such that the sequence of iterates $\{x^m\}$ satisfies, for each $m \geq 0$,*

$$\|x^{m+1} - x^*\| \leq \mu\|x^m - x^*\| + \xi\|e\|. \tag{2.29}$$

Then, for all $m \geq 0$,

$$\|\bar{x} - x^*\| \leq \mu^m\|x^*\| + \frac{\xi}{1-\mu}\|e\|. \tag{2.30}$$

Proof: We first prove by induction that, for all $m \geq 0$,

$$\|x^m - x^*\| \leq \mu^m\|x^*\| + \xi\left(\frac{1-\mu^m}{1-\mu}\right)\|e\|. \tag{2.31}$$

Supposing (2.31) holds for some $m \geq 0$, then we may apply (2.29) to (2.31) to deduce

$$\begin{aligned}\|x^{m+1} - x^*\| &\leq \mu\left[\mu^m\|x^*\| + \xi\left(\tfrac{1-\mu^m}{1-\mu}\right)\|e\|\right] + \xi\|e\| \\ &= \mu^{m+1}\|x^*\| + \left[\xi\left(1 + \tfrac{\mu-\mu^{m+1}}{1-\mu}\right)\right]\|e\| \\ &= \mu^{m+1}\|x^*\| + \xi\left(\tfrac{1-\mu^{m+1}}{1-\mu}\right)\|e\|,\end{aligned}$$

and so (2.31) also holds for $m+1$. Since $x^0 = 0$, the result holds trivially for $m = 0$, and therefore for all $m \geq 0$ by induction. Since $\mu^m \in (0,1)$ for all $m \geq 0$, (2.30) now follows. □

Provided $\mu < 1$, the $\mu^m\|x^*\|$ term in (2.30) tends to zero, and the expression $\xi/(1-\mu)$ may be viewed as a *stability factor*, giving a limiting bound on the approximation error as a multiple of the noise level $\|e\|$. We now proceed to the recovery results for both IHT and NIHT for arbitrary matrices.

Corollary 2.8 (**IHT**). *Consider Problem 1.2. Let $\mu_k^{IHT_\alpha}$ and $\xi_k^{IHT_\alpha}$ be defined as in Theorem 2.5. Then, provided $\mu_k^{IHT_\alpha} < 1$, the output, $\hat{x}$, of IHT with stepsize α at iteration m, satisfies*

$$\|\hat{x} - x^*\| \leq \left(\mu_k^{IHT_\alpha}\right)^m\|x^*\| + \frac{\xi_k^{IHT_\alpha}}{1-\mu_k^{IHT_\alpha}}\|e\|. \tag{2.32}$$

Proof: The result follows by combining Theorem 2.5 and Lemma 2.7. □

Corollary 2.9 (**NIHT**). *Consider Problem 1.2. Let $\mu_k^{NIHT_\kappa}$ and $\xi_k^{NIHT_\kappa}$ be defined as in Theorem 2.6. Then, provided $\mu_k^{NIHT_\kappa} < 1$, the output, $\hat{x}$, of NIHT with shrinkage parameter κ at iteration m, satisfies*

$$\|\hat{x} - x^*\| \leq \left(\mu_k^{NIHT_\kappa}\right)^m\|x^*\| + \frac{\xi_k^{NIHT_\kappa}}{1-\mu_k^{NIHT_\kappa}}\|e\|. \tag{2.33}$$

Proof: The result follows by combining Theorem 2.6 and Lemma 2.7. □

Though both Corollaries 2.8 and 2.9 give a limiting bound on the approximation error, they do not necessarily imply convergence of the algorithm. In the simplified noiseless case however, both results can be used to deduce convergence to x^* at a linear rate. In the next section, we give explicit results concerning convergence in the noiseless case, in the phase transition framework for Gaussian matrices. See also Section 3.2 in which we prove robust convergence results for both stepsize schemes as part of our new analysis.

A similar recovery result may also be obtained for the simplified version of NIHT in which the backtracking step is omitted, since in this case we may take $\underline{\alpha} := 1/(1+U_k)$ and $\overline{\alpha} := 1/(1-L_k)$. Indeed, since we could also sub-optimally take $\underline{\alpha} := 1/(1+U_{2k})$, it follows that Corollary 2.9 holds for NIHT without backtracking by setting $\kappa := 1$ in $\mu_k^{NIHT_\kappa}$ and $\xi_k^{NIHT_\kappa}$. Furthermore, since decreasing κ weakens the recovery condition, we see that removing the backtracking step actually leads to an improved recovery condition.

2.3 Improved phase transitions for IHT and NIHT

In [17], RIP recovery conditions were translated into phase transitions for Gaussian matrices in the proportional-dimensional asymptotic of Definition 1.5, using upper bounds on RIP constants derived in [21]. In this section, we follow essentially the same approach as in [17], applying Gaussian RIP bounds to the conditions derived in Section 2.2 for IHT and NIHT. Since the work in [17], tighter bounds on RIP constants for Gaussian matrices have been derived in [4], and we shall switch to using these improved bounds in order to optimize the resulting phase transitions.

We now consider the RIP for an $n \times N$ Gaussian matrix, whose RIP constants therefore depend not just upon the sparsity k, but also its dimensions (n, N). While it would be more precise to follow [17] and [19] in denoting the lower and upper RIP constants by $L(k, n, N)$ and $U(k, n, N)$ respectively, for the sake of uniformity we continue to use the notation introduced in Definition 1.6. Note that the functions $\mu_k^{IHT_\alpha}$, $\xi_k^{IHT_\alpha}$, $\mu_k^{NIHT_\kappa}$ and $\xi_k^{NIHT_\kappa}$ defined in (2.24), (2.25), (2.24) and (2.25) respectively also now depend upon (n, N).

Many of the asymptotic bounds in this thesis will involve the Shannon entropy, which we next define.

Definition 2.10 (**Shannon entropy [21]**). *Given $p \in (0, 1)$, define the Shannon entropy with base e logarithms as*

$$H(p) := -p \ln(p) - (1-p) \ln(1-p). \tag{2.34}$$

The following RIP bounds $\mathcal{L}(\delta, \rho)$ and $\mathcal{U}(\delta, \rho)$ were defined in [4].

Definition 2.11 (**RIP bounds [4, Definition 2.2]**). *Let $(\delta, \rho) \in (0, 1)^2$, $\gamma \in [\rho, \delta^{-1}]$. Let*

$$\psi_{min}(\lambda, \gamma) := H(\gamma) + \frac{1}{2} \left[(1-\gamma) \ln \lambda + \gamma \ln \gamma + 1 - \gamma - \lambda \right],$$

$$\psi_{max}(\lambda,\gamma) := \frac{1}{2}\left[(1+\gamma)\ln\lambda - \gamma\ln\gamma + 1 + \gamma - \lambda\right],$$

where $H(p)$ is defined in (2.34). Define $\lambda^{min}(\delta,\rho;\gamma)$ and $\lambda^{max}(\delta,\rho;\gamma)$ to be the solutions to

$$\delta\psi_{min}(\lambda^{min}(\delta,\rho;\gamma),\gamma) + H(\delta\rho) - \delta\gamma H(\rho/\gamma) = 0 \quad \textit{for} \quad \lambda^{min}(\delta,\rho;\gamma) \le 1-\gamma,$$

and

$$\delta\psi_{max}(\lambda^{max}(\delta,\rho;\gamma),\gamma) + H(\delta\rho) - \delta\gamma H(\rho/\gamma) = 0 \quad \textit{for} \quad \lambda^{max}(\delta,\rho;\gamma) \ge 1+\gamma.$$

Let $\lambda^{min}(\delta,\rho) := \max_\gamma \lambda^{min}(\delta,\rho;\gamma)$ and $\lambda^{max}(\delta,\rho) := \min_\gamma \lambda^{max}(\delta,\rho;\gamma)$, and define

$$\mathcal{L}(\delta,\rho) := 1 - \lambda^{min}(\delta,\rho) \quad \textit{and} \quad \mathcal{U}(\delta,\rho) := \lambda^{max}(\delta,\rho) - 1.$$

The next result, from [4], proves that the numerically computable functions defined in Definition 2.11 are upper bounds on RIP constants for Gaussian matrices with exponentially high probability.

Theorem 2.12 (**Validity of RIP bounds [4, Theorem 2.3]**). *Let A be a matrix of size $n \times N$ whose entries are drawn i.i.d. from $\mathcal{N}(0,1/n)$. Let $\mathcal{L}(\delta,\rho)$ and $\mathcal{U}(\delta,\rho)$ be defined as in Definition 2.11. For any fixed $\epsilon > 0$, in the proportional-growth asymptotic,*

$$\mathbb{P}[L_k < \mathcal{L}(\delta,\rho) + \epsilon] \to 1 \quad \textit{and} \quad \mathbb{P}[U_k < \mathcal{U}(\delta,\rho) + \epsilon] \to 1, \tag{2.35}$$

exponentially in n.

The next two lemmas are needed to enable a translation of Corollaries 2.8 and 2.9 for arbitrary matrices into the phase transition framework for Gaussian matrices.

Lemma 2.13 (**[17, Lemma 12]**). *For some $\tau < 1$, define the set $\mathcal{Z} := (0,\tau)^p \times (0,\infty)^q$ and let $F : \mathcal{Z} \to \mathbb{R}$ be continuously differentiable on $\mathcal{Z}$. Let $A \in \mathbb{R}^{n\times N}$ be a Gaussian matrix with RIP constants $L_k,\dots,L_{pk}$ and $U_k,\dots,U_{qk}$, and let $\mathcal{L}(\delta,\rho),\dots,\mathcal{L}(\delta,p\rho)$ and $\mathcal{U}(\delta,\rho),\dots,\mathcal{U}(\delta,q\rho)$ be defined as in Definition 2.11. Define $\mathbf{1}$ to be the vector of all ones, and*

$$\begin{aligned} z(k,n,N) &:= [L_k,\dots,L_{pk},U_k,\dots,U_{qk}],\\ z(\delta,\rho) &:= [\mathcal{L}(\delta,\rho),\dots,\mathcal{L}(\delta,p\rho),\mathcal{U}(\delta,\rho),\dots,\mathcal{U}(\delta,q\rho)]. \end{aligned}$$

Suppose, for all $t \in \mathcal{Z}$, $(\nabla F[t])_i \ge 0$ for all $i = 1,\dots,p+q$ and there exists $j \in \{1,\dots,p\}$ such that $(\nabla F[t])_j > 0$. Then, for any $\epsilon \in (0,1)$, in the proportional-growth asymptotic,

$$\mathbb{P}\left(F[z(k,n,N)] < F[z(\delta,(1+\epsilon)\rho)]\right) \to 1 \quad \textit{as } n \to \infty, \tag{2.36}$$

exponentially in n on the draw of A. Also, $F[z(\delta,\rho)]$ is strictly increasing in ρ.

Proof: A proof was given in [17] for the case where $\mathcal{L}(\delta,\cdot)$ and $\mathcal{U}(\delta,\cdot)$ are the earlier RIP bounds of Blanchard et al. [21], in which the only assumption made concerning the RIP bounds, in addition to Theorem 2.12, is that $\mathcal{L}(\delta,\rho)$ is strictly increasing in ρ and $\mathcal{U}(\delta,\rho)$ is nondecreasing in ρ. Both properties also hold for the bounds defined in Definition 2.11 [4], and therefore the same argument in [17, Lemma 12] also holds in our case. □

Lemma 2.14. *For some $\tau < 1$, define the set $\mathcal{Z} := (0,\tau)^p \times (0,\infty)^q$ and let $F, G, H : \mathcal{Z} \to \mathbb{R}$ satisfy the conditions of Lemma 2.13. Suppose that*

$$\mu(k,n,N) = \max\big\{F[z(k,n,N)], G[z(k,n,N)]\big\}, \quad \xi(k,n,N) = H[z(k,n,N)], \tag{2.37}$$

and

$$\mu(\delta,\rho) = \max\big\{F[z(\delta,\rho)], G[z(\delta,\rho)]\big\}, \quad \xi(\delta,\rho) = H[z(\delta,\rho)]. \tag{2.38}$$

Then $\mu(\delta,\rho)$ and $\xi(\delta,\rho)$ are both strictly increasing in ρ and, for any $\epsilon \in (0,1)$, in the proportional-growth asymptotic,

$$\mathbb{P}\big\{\mu(k,n,N) \geq \mu(\delta,(1+\epsilon)\rho)\big\} \to 0, \tag{2.39}$$

and

$$\mathbb{P}\big\{\xi(k,n,N) \geq \xi(\delta,(1+\epsilon)\rho)\big\} \to 0, \tag{2.40}$$

both exponentially in n. Furthermore, define $\hat{\rho}(\delta)$ as the unique solution to $\mu(\delta,\rho) = 1$, and suppose that

$$\rho < (1-\epsilon)\hat{\rho}(\delta). \tag{2.41}$$

Then

$$\mu(\delta,(1+\epsilon)\rho) < 1, \tag{2.42}$$

and, in the proportional-growth asymptotic,

$$\mathbb{P}\big\{\mu(k,n,N) \geq 1\big\} \to 0, \tag{2.43}$$

exponentially in n.

Proof: By assumption, we may apply Lemma 2.13 to each of $F(z)$, $G(z)$ and $H(z)$, deducing from (2.36) that

$$\mathbb{P}\left(F[z(k,n,N)] < F[z(\delta,(1+\epsilon)\rho)]\right) \to 1 \quad \text{as } n \to \infty, \tag{2.44}$$

$$\mathbb{P}\left(G[z(k,n,N)] < G[z(\delta,(1+\epsilon)\rho)]\right) \to 1 \quad \text{as } n \to \infty, \tag{2.45}$$

$$\mathbb{P}\left(H[z(k,n,N)] < H[z(\delta,(1+\epsilon)\rho)]\right) \to 1 \quad \text{as } n \to \infty, \tag{2.46}$$

exponentially in n, and that $F[z(\delta,\rho)]$, $G[z(\delta,\rho)]$ and $H[z(\delta,\rho)]$ are each strictly increasing in

ρ, from which it immediately follows that both $\mu(\delta,\rho)$ and $\xi(\delta,\rho)$ are also strictly increasing in ρ. Combining (2.37), (2.38), (2.45) and (2.46), we have

$$\begin{aligned} & \mathbb{P}\big\{\mu(k,n,N) \geq \mu(\delta,(1+\epsilon)\rho)\big\} \\ = \;& \mathbb{P}\Big\{ \max\big\{F[z(k,n,N)], G[z(k,n,N)]\big\} \geq \max\big\{F[z(\delta,(1+\epsilon)\rho)], G[z(\delta,(1+\epsilon)\rho)]\big\}\Big\} \\ \leq \;& \mathbb{P}\big\{F[z(k,n,N)] \geq F[z(\delta,(1+\epsilon)\rho)]\big\} + \mathbb{P}\big\{G[z(k,n,N)] \geq G[z(\delta,(1+\epsilon)\rho)]\big\} \\ & \to 0 \;\text{ as }\; n\to\infty, \end{aligned} \tag{2.47}$$

and therefore (2.39) holds. Meanwhile, combining (2.37), (2.38) and (2.46) immediately yields (2.40). Now suppose (2.41) holds. Since $1-\epsilon < (1+\epsilon)^{-1}$ for any $\epsilon \in (0,1)$, (2.41) implies that

$$(1+\epsilon)\rho < \hat{\rho}(\delta), \tag{2.48}$$

Since $\mu(\delta,\rho)$ is strictly increasing in ρ, it follows from (2.48) and the definition of $\hat{\rho}(\delta)$ that

$$\mu(\delta,(1+\epsilon)\rho) < \mu\left(\delta,\hat{\rho}(\delta)\right) = 1,$$

which proves (2.42), and from which it also follows that

$$\mathbb{P}\big\{\mu(k,n,N) \geq 1\big\} \leq \mathbb{P}\big\{\mu(k,n,N) \geq \mu(\delta,(1+\epsilon)\rho)\big\},$$

to which we may apply (2.47) to deduce (2.43). □

We now proceed to our phase transition results.

Theorem 2.15 (IHT). *Consider Problem 1.2 and suppose the entries of $A \in \mathbb{R}^{n\times N}$ are drawn i.i.d. from $\mathcal{N}(0,1/n)$. Define*

$$\mu^{IHT_\alpha}(\delta,\rho) := \sqrt{3}\max\{\alpha[1+\mathcal{U}(\delta,3\rho)]-1, 1-\alpha[1-\mathcal{L}(\delta,3\rho)]\} \tag{2.49}$$

and

$$\xi^{IHT_\alpha}(\delta,\rho) := \alpha\sqrt{3[1+\mathcal{U}(\delta,2\rho)]}, \tag{2.50}$$

and define $\hat{\rho}_{RIP}^{IHT_\alpha}(\delta)$ as the unique solution to $\mu^{IHT_\alpha}(\delta,\rho)=1$. Choose $\epsilon\in(0,1)$ and suppose that

$$\rho < (1-\epsilon)\hat{\rho}_{RIP}^{IHT_\alpha}(\delta). \tag{2.51}$$

Suppose $\hat{x}$ is the output of IHT with stepsize α at iteration m. Then

$$\mu^{IHT_\alpha}(\delta,(1+\epsilon)\rho) < 1, \tag{2.52}$$

and, in the proportional-growth asymptotic,

$$\|\hat{x} - x^*\| \leq \left(\mu^{IHT_\alpha}(\delta, (1+\epsilon)\rho)\right)^m \|x^*\| + \frac{\xi^{IHT_\alpha}(\delta, (1+\epsilon)\rho)}{1 - \mu^{IHT_\alpha}(\delta, (1+\epsilon)\rho)}\|e\|, \tag{2.53}$$

with probability tending to 1 *exponentially in* n*.*

Proof: Select $\epsilon \in (0,1)$, fix $\tau < 1$ and let

$$z(k,n,N) := [L_{3k}, U_{2k}, U_{3k}] \quad \text{and} \quad z(\delta,\rho) := [\mathcal{L}(\delta, 3\rho), \mathcal{U}(\delta, 2\rho), \mathcal{U}(\delta, 3\rho)].$$

Define $\mathcal{Z} := (0,\tau) \times (0,\infty)^2$, and define the functions $F_\alpha(z), G_\alpha(z), H_\alpha(z) : \mathcal{Z} \to \mathbb{R}$ as

$$F_\alpha(z) = F_\alpha(z_1, z_2, z_3) := \sqrt{3}[\alpha(1+z_3) - 1], \tag{2.54}$$

$$G_\alpha(z) = G_\alpha(z_1, z_2, z_3) := \sqrt{3}[1 - \alpha(1-z_1)], \tag{2.55}$$

$$H_\alpha(z) = H_\alpha(z_1, z_2, z_3) := \alpha\sqrt{3(1+z_2)}, \tag{2.56}$$

noting that

$$\mu_k^{IHT_\alpha} = \max\left\{F_\alpha[z(k,n,N)], G_\alpha[z(k,n,N)]\right\}, \quad \xi_k^{IHT_\alpha} = H_\alpha[z(k,n,N)],$$

where $\mu_k^{IHT_\alpha}$ and $\xi_k^{IHT_\alpha}$ are defined in (2.24) and (2.25) respectively, and

$$\mu^{IHT_\alpha}(\delta,\rho) = \max\left\{F_\alpha[z(\delta,\rho)], G_\alpha[z(\delta,\rho)]\right\}, \quad \xi^{IHT_\alpha}(\delta,\rho) = H_\alpha[z(\delta,\rho)],$$

where $\mu^{IHT_\alpha}(\delta,\rho)$ and $\xi^{IHT_\alpha}(\delta,\rho)$ are defined in (2.49) and (2.50) respectively. Now $F_\alpha(z)$, $G_\alpha(z)$ and $H_\alpha(z)$ are continuously differentiable and nondecreasing in $(z_1, z_2, z_3) \in \mathcal{Z}$, and strictly increasing in z_3, z_1 and z_2 respectively due to $\alpha > 0$, and therefore each satisfies the conditions of Lemma 2.13. We may therefore apply Lemma 2.14, deducing

$$\mathbb{P}\left\{\mu_k^{IHT_\alpha} \geq \mu^{IHT_\alpha}(\delta, (1+\epsilon)\rho)\right\} \to 0, \tag{2.57}$$

and

$$\mathbb{P}\left\{\xi_k^{IHT_\alpha} \geq \xi^{IHT_\alpha}(\delta, (1+\epsilon)\rho)\right\} \to 0, \tag{2.58}$$

exponentially in n, and furthermore that $\mu^{IHT_\alpha}(\delta,\rho)$ and $\xi^{IHT_\alpha}(\delta,\rho)$ are both strictly increasing in ρ, from which it follows that $\hat{\rho}_{RIP}^{IHT_\alpha}(\delta)$ is unique. Since (2.51) holds, we may also use Lemma 2.14 to deduce (2.52), and furthermore that

$$\mathbb{P}\left\{\mu_k^{IHT_\alpha} \geq 1\right\} \to 0, \tag{2.59}$$

exponentially in n, and we may apply Corollary 2.8 to deduce (2.32) with probability tending

to 1 exponentially in n. Since $\mu^{IHT_\alpha}(\delta,\rho)$ and $\xi^{IHT_\alpha}(\delta,\rho)$ are strictly increasing in ρ, (2.53) now follows from (2.32), (2.57) and (2.58). □

Theorem 2.15 gives a continuous range of phase transitions for any $0 < \alpha < 2$. For $\alpha \geq 2$, the result gives $\hat{\rho}_{RIP}^{IHT_\alpha}(\delta) = 0$ for all $\delta \in (0,1)$. It is clear that $\mu^{IHT_\alpha}(\delta,\rho)$ takes its minimum value where both expressions inside the maximum in (2.49) are equal, which implies that the optimum phase transition is obtained when the stepsize is taken to be

$$\hat{\alpha} := 2/[2+\mathcal{U}(\delta,3\rho)-\mathcal{L}(\delta,3\rho)], \tag{2.60}$$

for which we have

$$\mu^{IHT_{\hat{\alpha}}}(\delta,\rho) := \sqrt{3}\left[\frac{\mathcal{L}(\delta,3\rho)+\mathcal{U}(\delta,3\rho)}{2+\mathcal{U}(\delta,3\rho)-\mathcal{L}(\delta,3\rho)}\right], \quad \xi^{IHT_{\hat{\alpha}}}(\delta,\rho) := 2\sqrt{3}\left[\frac{\sqrt{1+\mathcal{U}(\delta,2\rho)}}{2+\mathcal{U}(\delta,3\rho)-\mathcal{L}(\delta,3\rho)}\right]. \tag{2.61}$$

The stepsize choice (2.60) was first adopted in [17]. It is instructive to point out that $\mu^{IHT_{\hat{\alpha}}}(\delta,\rho)$ differs only from the corresponding function in [17] by a scaling factor of $\sqrt{3}/(2\sqrt{2})$, from which it is immediately clear that the present recovery condition represents an improvement. However, as in [17], the result degrades as α is either increased or decreased away from $\hat{\alpha}$.

In the idealized case of zero measurement noise, we can deduce from Theorem 2.15 guaranteed convergence of IHT at a linear rate.

Corollary 2.16 (**Noiseless case**). *Consider Problem 1.1 and suppose the entries of $A \in \mathbb{R}^{n\times N}$ are drawn i.i.d. from $\mathcal{N}(0,1/n)$. Choose $\epsilon \in (0,1)$ and suppose that (2.51) holds, where $\hat{\rho}_{RIP}^{IHT_\alpha}(\delta)$ and $\mu^{IHT_\alpha}(\delta,\rho)$ are defined as in Theorem 2.15. Then, in the proportional-growth asymptotic, the iterates of IHT with stepsize α converge to x^* at a linear rate, with probability tending to 1 exponentially in n.*

Proof: Since we consider Problem 1.1, we have $e := 0$. Provided (2.51) holds, we can apply Theorem 2.15 with $e := 0$, deducing that, for any $m \geq 0$,

$$\|x^m - x^*\| \leq \left(\mu^{IHT_\alpha}(\delta,(1+\epsilon)\rho)\right)^m \|x^*\|,$$

where

$$\mu^{IHT_\alpha}(\delta,(1+\epsilon)\rho) < 1,$$

and so we have convergence to x^* with convergence rate $\mu^{IHT_\alpha}(\delta,(1+\epsilon)\rho)$. □

We next obtain a recovery result for NIHT.

Theorem 2.17 (**NIHT**). *Consider Problem 1.2 and suppose the entries of* $A \in \mathbb{R}^{n\times N}$ *are drawn i.i.d. from* $\mathcal{N}(0, 1/n)$. *Define*

$$\mu^{NIHT_\kappa}(\delta,\rho) := \sqrt{3}\max\left\{\frac{1+\mathcal{U}(\delta,3\rho)}{1-\mathcal{L}(\delta,\rho)} - 1, 1 - \frac{1-\mathcal{L}(\delta,3\rho)}{\kappa[1+\mathcal{U}(\delta,2\rho)]}\right\} \tag{2.62}$$

and

$$\xi^{NIHT_\kappa}(\delta,\rho) := \frac{\sqrt{3[1+\mathcal{U}(\delta,2\rho)]}}{1-\mathcal{L}(\delta,\rho)}, \tag{2.63}$$

and define $\hat{\rho}^{NIHT_\kappa}_{RIP}(\delta)$ *as the unique solution to* $\mu^{NIHT_\kappa}(\delta,\rho) = 1$. *Choose* $\epsilon \in (0,1)$ *and suppose that*

$$\rho < (1-\epsilon)\hat{\rho}^{NIHT_\kappa}_{RIP}(\delta). \tag{2.64}$$

Suppose $\hat{x}$ *is the output of NIHT with shrinkage parameter* κ *at iteration* m. *Then*

$$\mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho) < 1, \tag{2.65}$$

and, in the proportional-growth asymptotic,

$$\|\hat{x}-x^*\| \le \left(\mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)\right)^m \|x^*\| + \frac{\xi^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)}{1-\mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)}\|e\|, \tag{2.66}$$

with probability tending to 1 *exponentially in* n.

Proof: Select $\epsilon \in (0,1)$, fix $\tau < 1$ and let

$$z(k,n,N) := [L_k, L_{3k}, U_{2k}, U_{3k}] \quad \text{and} \quad z(\delta,\rho) := [\mathcal{L}(\delta,\rho), \mathcal{L}(\delta,3\rho), \mathcal{U}(\delta,2\rho), \mathcal{U}(\delta,3\rho)].$$

Define $\mathcal{Z} := (0,\tau)^2 \times (0,\infty)^2$, and define the functions $F_\alpha(z), G_\alpha(z), H_\alpha(z) : \mathcal{Z} \to \mathbb{R}$ as

$$\begin{aligned}
F_\kappa(z) &= F_\kappa(z_1,z_2,z_3,z_4) := \sqrt{3}\left[\frac{1+z_4}{1-z_1} - 1\right],\\
G_\kappa(z) &= G_\kappa(z_1,z_2,z_3,z_4) := \sqrt{3}\left[1 - \frac{1-z_2}{\kappa(1+z_3)}\right],\\
H_\kappa(z) &= H_\kappa(z_1,z_2,z_3,z_4) := \frac{\sqrt{3(1+z_3)}}{1-z_1},
\end{aligned}$$

noting that

$$\mu^{NIHT_\kappa}_k = \max\left\{F_\kappa[z(k,n,N)], G_\kappa[z(k,n,N)]\right\}, \quad \xi^{NIHT_\kappa}_k = H_\kappa[z(k,n,N)],$$

where $\mu^{NIHT_\kappa}_k$ and $\xi^{NIHT_\kappa}_k$ are defined in (2.27) and (2.28) respectively, and

$$\mu^{NIHT_\kappa}(\delta,\rho) = \max\left\{F_\kappa[z(\delta,\rho)], G_\kappa[z(\delta,\rho)]\right\}, \quad \xi^{NIHT_\kappa}(\delta,\rho) = H_\kappa[z(\delta,\rho)],$$

where $\mu^{NIHT_\kappa}(\delta,\rho)$ and $\xi^{NIHT_\kappa}(\delta,\rho)$ are defined in (2.62) and (2.63) respectively. Now $F_\kappa(z)$, $G_\kappa(z)$ and $H_\kappa(z)$ are continuously differentiable and nondecreasing in (z_1,z_2,z_3,z_4), and strictly increasing componentwise in (z_1,z_4), (z_2,z_3) and (z_1,z_3) respectively, and therefore each satisfies the conditions of Lemma 2.13. We may therefore apply Lemma 2.14, deducing

$$\mathbb{P}\big\{\mu_k^{NIHT_\kappa} \geq \mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)\big\} \to 0 \tag{2.67}$$

and

$$\mathbb{P}\big\{\xi_k^{NIHT_\kappa} \geq \xi^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)\big\} \to 0, \tag{2.68}$$

exponentially in n, and furthermore that $\mu^{NIHT_\kappa}(\delta,\rho)$ and $\xi^{NIHT_\kappa}(\delta,\rho)$ are both strictly increasing in ρ, from which it follows that $\hat{\rho}_{RIP}^{NIHT_\kappa}(\delta)$ is unique. Since (2.64) holds, we may also use Lemma 2.14 to deduce (2.65), and furthermore that

$$\mathbb{P}\big\{\mu_k^{NIHT_\kappa} \geq 1\big\} \to 0, \tag{2.69}$$

exponentially in n, and we may apply Corollary 2.9 to deduce (2.33) with probability tending to 1 exponentially in n. Since $\mu^{NIHT_\kappa}(\delta,\rho)$ and $\xi^{NIHT_\kappa}(\delta,\rho)$ are strictly increasing in ρ, (2.66) now follows from (2.33), (2.67) and (2.68). □

In the idealized case of zero measurement noise, we can deduce from Theorem 2.17 guaranteed convergence of NIHT at a linear rate.

Corollary 2.18 (**Noiseless case**). *Consider Problem 1.1 and suppose the entries of $A \in \mathbb{R}^{n\times N}$ are drawn i.i.d. from $\mathcal{N}(0,1/n)$. Choose $\epsilon \in (0,1)$ and suppose that (2.64) holds, where $\hat{\rho}_{RIP}^{NIHT_\kappa}(\delta)$ and $\mu^{NIHT_\kappa}(\delta,\rho)$ are defined as in Theorem 2.17. Then, in the proportional-growth asymptotic, the iterates of NIHT with shrinkage parameter κ converge to x^* at a linear rate, with probability tending to 1 exponentially in n.*

Proof: Since we consider Problem 1.1, we have $e := 0$. Provided (2.64) holds, we can apply Theorem 2.17 with $e := 0$, deducing that, for any $m \geq 0$,

$$\|x^m - x^*\| \leq \big(\mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)\big)^m \|x^*\|,$$

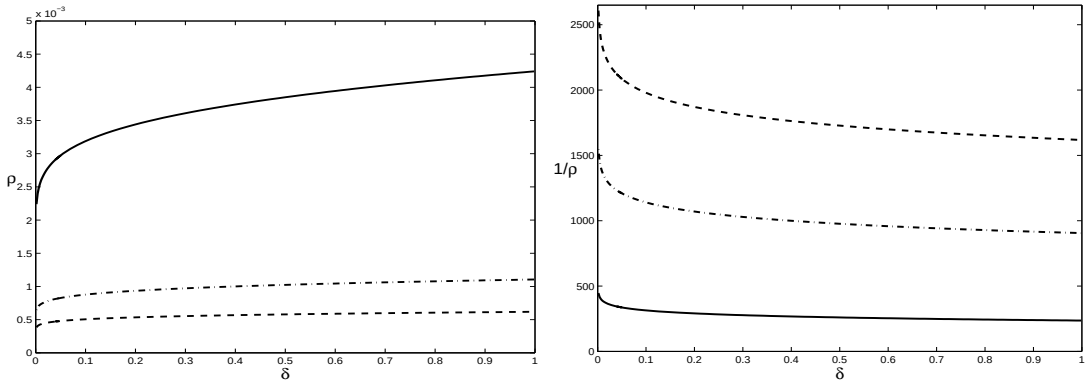
where

$$\mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho) < 1,$$

and so we have convergence to x^* with convergence rate $\mu^{NIHT_\kappa}(\delta,(1+\epsilon)\rho)$. □

2.4 Discussion of results

Current state-of-the-art phase transitions. The functions $\hat{\rho}_{RIP}^{IHT_\alpha}(\delta)$ and $\hat{\rho}_{RIP}^{NIHT_\kappa}(\delta)$ define curves in the (δ, ρ)-plane, which are plotted in Figure 2.1(a). For IHT, we take $\alpha := \hat{\alpha}$ as defined in (2.60), and a shrinkage parameter of $\kappa := 1.1$ is used for NIHT. For a sequence of problem instances with dimensions (k, n, N), provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{RIP}^{IHT_{\hat{\alpha}}}(\delta)$ or $\hat{\rho}_{RIP}^{NIHT_{1.1}}(\delta)$, then with probability approaching 1 exponentially fast in n on the draw of the matrix A, the algorithm in question achieves the bound (2.53) or (2.66) respectively for any k-sparse vector x^*. In the idealized case of zero measurement noise, Corollaries 2.16 and 2.18 imply exact recovery of any k-sparse x^*. The curves therefore define lower bounds on the strong phase transition for the two variants of IHT. Also plotted in Figure 2.1(a) is the corresponding lower bound on the phase transition for IHT resulting from the previous analysis in [17]. We observe an improvement on the phase transition for IHT by around a factor of 4, while the phase transition is in fact lower for NIHT than for IHT.



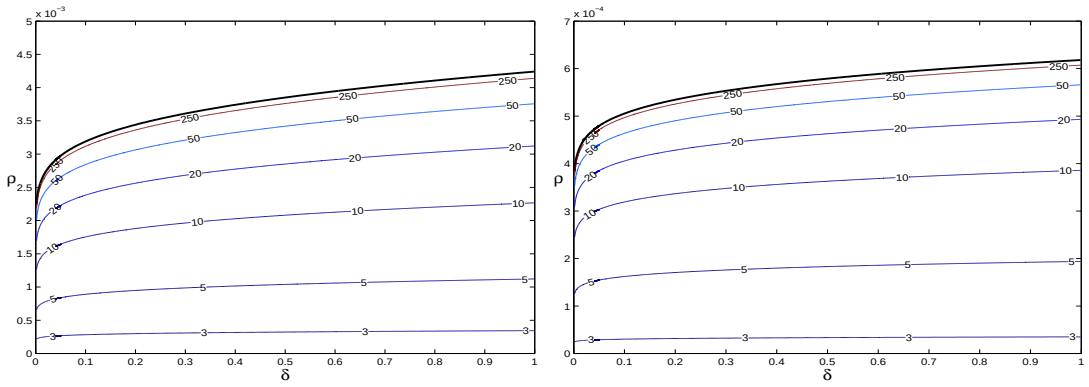
(a) (b)

Figure 2.1: (a) Current best lower bounds on the strong phase transition based upon the RIP for recovery using Gaussian matrices: IHT (unbroken); NIHT (dashed). The corresponding result for IHT from [17] is shown for comparison (dash-dot). (b) The inverse of each respective phase transition.

Figure 2.1(b) plots the reciprocal of each phase transition, which may be interpreted as the number of measurements that must be taken as a multiple of the sparsity in order to guarantee recovery. In particular, a minimum of $n \geq 234k$ and $n \geq 1617k$ measurements must be taken to guarantee recovery using IHT and NIHT respectively. The figure for IHT is, of course, an improvement upon the corresponding figure resulting from the analysis of IHT in [23], which is $n \geq 905k$.

Theorems 2.15 and 2.17 define a *stability factor*, namely a maximum factor by which the l_2-norm of the noise is amplified in the final approximation error. Plots of the stability factors for both IHT and NIHT are displayed in Figure 2.2. We observe, as in [17], that the stability factor becomes unbounded as the phase transition is approached. It follows that phase transitions lower than those defined by $\hat{\rho}_{RIP}^{IHT_\alpha}(\delta)$ and $\hat{\rho}_{RIP}^{NIHT_\kappa}(\delta)$ are actually required to ensure a

satisfactory stability in the presence of noise.



(a) (b)

Figure 2.2: Contour plots of the stability factors for IHT and NIHT: (a) $\xi^{IHT_{\hat{\alpha}}}(\delta,\rho)/[1-\mu^{IHT_{\hat{\alpha}}}(\delta,\rho)]$ (b) $\xi^{NIHT_{1.1}}(\delta,\rho)/[1-\mu^{NIHT_{1.1}}(\delta,\rho)]$.

In the idealized case of noiseless measurements, an inspection of Corollaries 2.16 and 2.18 reveals that, below the phase transition, each algorithm has guaranteed linear convergence, with convergence factor given by $\mu^{IHT_{\hat{\alpha}}}(\delta,\rho)$ and $\mu^{NIHT_{1.1}}(\delta,\rho)$ respectively. Therefore, in order to guarantee convergence at a given rate, even lower phase transitions than those given by $\hat{\rho}^{IHT_{\hat{\alpha}}}_{RIP}(\delta)$ and $\hat{\rho}^{NIHT_{1.1}}_{RIP}(\delta)$ are in fact required.

Comparison with other RIP results for IHT. Let us now justify the claim that the phase transitions derived in Section 2.3 represent the current state-of-the-art. Concerning IHT, we have already argued that the result of Theorem 2.15 is necessarily an improvement upon the corresponding result in [17], based on the analysis of Blumensath and Davies [23], since the only change to the function $\mu^{IHT_{\hat{\alpha}}}(\delta,\rho)$ is a scaling down by a factor of $\sqrt{3}/(2\sqrt{2})$. The same may be observed for the analysis in [92], where Foucart obtains the symmetric condition $R_{3k} < 1/2$ for unit stepsize. Like in the analysis in Section 2.2, R_{3k} is used to bound the expression

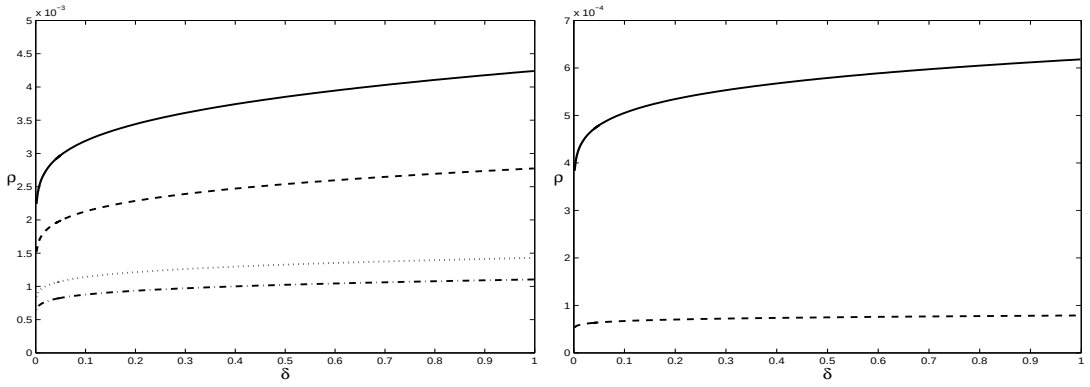
$$\left\|(I-\alpha^m A_\Omega^T A_\Omega)(x^m-x^*)_\Omega\right\|,$$

where $|\Omega| \leq 3k$, and so in the asymmetric context for arbitrary stepsize, the bound (2.13) may be used to finally obtain

$$\mu^{IHT_{\hat{\alpha}}}(\delta,\rho) := 2\left[\frac{\mathcal{L}(\delta,3\rho)+\mathcal{U}(\delta,3\rho)}{2+\mathcal{U}(\delta,3\rho)-\mathcal{L}(\delta,3\rho)}\right],$$

which represents a simple scaling up of $\mu^{IHT_{\hat{\alpha}}}(\delta,\rho)$ by a factor of $2/\sqrt{3}$ compared to the result presented here.

While the methods of analysis in [23, 92, 93] are very similar, one other somewhat different approach was taken in [96], where it was shown that $R_{2k} < 1/3$ guarantees recovery for the



(a) (b)

Figure 2.3: A comparison of RIP-based phase transitions for IHT algorithms; (a) IHT: [93] (unbroken); [92] (dashed); [96] (dotted); [23] (dash-dot); (b) NIHT: our analysis (unbroken); [27] (dashed).

stepsize choice $\tilde{\alpha} := 1/[1 + \mathcal{U}(\delta, 2\rho)]$. The approach in [17] was also applied to this analysis in [148], where the optimal expression for $\mu^{IHT_{\tilde{\alpha}}}$ was found to be

$$\mu^{IHT_{\tilde{\alpha}}}(\delta, \rho) := \sqrt{\frac{\mathcal{L}(\delta, 2\rho) + \mathcal{U}(\delta, 2\rho)}{1 - \mathcal{L}(\delta, 2\rho)}}.$$

Lower bounds on the strong phase transition for Gaussian matrices for each of the analyses of IHT [23, 96, 92, 93] are shown in Figure 2.3(a) to allow a comparison. We observe that the highest phase transition is obtained from the result presented here based upon Foucart's analysis in [93].

Concerning NIHT, we must translate the result of Blumensath and Davies from [27] into the phase transition framework to enable a comparison with the phase transition displayed in this section. This is particularly straightforward, since the authors performed an asymmetric RIP analysis, obtaining

$$\mu_k^{NIHT_\kappa} := 4 \max \left\{ \frac{1 + U_{2k}}{1 - L_{2k}} - 1, 1 - \frac{1 - L_{2k}}{\kappa(1 + U_{2k})} \right\}.$$

The authors upper bound the stepsize by $\alpha^m \leq 1/(1 - L_{2k})$, which may in fact be tightened to $\alpha^m \leq 1/(1 - L_k)$ in accordance with (2.3). This tightening leads to

$$\mu_k^{NIHT_\kappa} := 4 \max \left\{ \frac{1 + U_{2k}}{1 - L_k} - 1, 1 - \frac{1 - L_{2k}}{\kappa(1 + U_{2k})} \right\},$$

which, after the substitution of Gaussian RIP bounds, yields

$$\mu^{NIHT_\kappa}(\delta, \rho) := 4 \max \left\{ \frac{1 + \mathcal{U}(\delta, 2\rho)}{1 - \mathcal{L}(\delta, \rho)} - 1, 1 - \frac{1 - \mathcal{L}(\delta, 2\rho)}{\kappa[1 + \mathcal{U}(\delta, 2\rho)]} \right\}.$$

The resulting phase transition, with $\kappa := 1.1$, is plotted alongside the one determined for NIHT

in this thesis in Figure 2.3(b). We again observe that the phase transition obtained from the present analysis is higher. However, a comparison with Figure 2.3(a) shows that the best phase transition for NIHT is considerably lower than its counterpart for IHT.

2.5 Support sizes of RIP constants

The observant reader will have noticed that the form of $\mu_k^{NIHT_\kappa}$ for the analysis in [27] presented above is similar to that for the analysis considered in this chapter, except that RIP constants of order $3k$ are replaced by RIP constants of order $2k$.[2] This change was sought deliberately by the authors in order to obtain a condition in terms of RIP constants of lower order. Indeed, this is one of several examples in the recent literature [127, 31, 30, 92] where reductions in support sizes of RIP recovery conditions have been intentionally sought. One motivation for this practice is the greater aesthetic appeal of conditions with smaller support sizes. In addition, there seems to be something natural about seeking results with support size $2k$ in the light of the $L_{2k} < 1$ condition for a sparse solution to be the most sparse (see Lemma 1.7). A particularly simple example is found in [92], in which the condition $R_{2k} < 1/4$ is deduced from the condition $R_{3k} < 1/2$. In this case, it is not clear upon initial inspection which condition is weaker, since both the bound and the support size have been reduced.

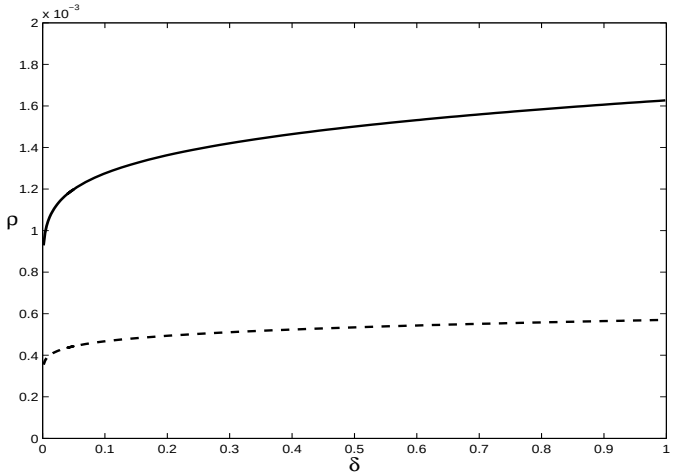


Figure 2.4: An example of support set splitting from [92]: phase transitions for Gaussian matrices for (a) $R_{3k} < 1/2$ (unbroken); (b) $R_{2k} < 1/4$ (dashed).

The phase transition framework for Gaussian matrices allows such support size reductions to be examined for efficacy. Figure 2.4 displays lower bounds on the phase transition for Gaussian matrices for each of the two conditions in [92]. We observe that the original condition $R_{3k} < 1/2$ is in fact the weaker condition, at least for Gaussian matrices. The point was made in [19, 20] that it is not always quantitatively advantageous to seek RIP conditions with smaller

[2]Material in this section extends the observations published in [19, 20], which were joint authorships with J. Blanchard.

support sizes, and several examples were given to illustrate this. In particular, it was noted that conditions involving smaller support sizes are often obtained simply by splitting support sets, and it was argued in [19, 20] that, since this approach involves the invoking of an inequality which is unlikely to be sharp, it is unlikely to offer any quantitative improvement. Foucart also obtains the result $R_{2k} < 1/4$ by splitting support sets, and therefore it is no surprise that the resulting condition is stricter.

In this chapter, we have sought to obtain the best possible phase transitions for Gaussian matrices by drawing upon current state-of-the-art RIP recovery analysis. The resulting phase transitions are still pessimistic compared to average-case behaviour of IHT algorithms (see Section 1.7), and there is a clear need for an alternative method of analysis leading to average-case recovery guarantees. We turn our efforts in this direction in the next chapter, where we present our new approach to recovery analysis for IHT algorithms.

Chapter 3

A new recovery analysis of IHT algorithms

In this chapter[1], we present a new recovery analysis of IHT algorithms which takes a markedly different approach to the analysis in Chapter 2. In the context of IHT, whereas previous analysis takes the direct approach of bounding the approximation error from iteration to iteration, we take a two-part approach in which we analyse the fixed points of the algorithm. Firstly, we prove necessary conditions for there to be a fixed point on a given support, in order to determine bounds on the approximation error of any fixed point. Secondly, we give a condition which guarantees the convergence of IHT to one of its fixed points. Our recovery results then follow by combining the two parts of the analysis. By extending the notion of a fixed point to the concept of an $\underline{\alpha}$-*stable point*, we follow broadly the same approach to obtain recovery results for NIHT. By way of outline, in Section 3.1, we introduce the concept of an $\underline{\alpha}$-*stable point* of generic IHT, and prove a necessary condition for there to be an $\underline{\alpha}$-stable point on a given support. In Section 3.2, we provide RIP conditions guaranteeing convergence of both the IHT and NIHT stepsize variants to an $\underline{\alpha}$-stable point. We prove our results in the context of Problem 1.3, namely in the most general case of k-compressible signals and noisy measurements. The reason for this approach is that, in the following two chapters, we will analyse the conditions derived in this chapter by making average-case assumptions, and therefore we wish to dispense with the RIP wherever possible. Considering Problem 1.3 directly means we avoid the need to revert to the RIP and Lemma 2.2 in order to extend results for k-sparse recovery to results for k-compressible recovery.

In later chapters, we will use average-case assumptions to analyse the stable point conditions derived in Section 3.1 for Gaussian matrices. However, the RIP can also be used to analyse these conditions, and we perform this analysis in Section 3.3 for comparison purposes. By combining these results with the RIP-based convergence analysis from Section 3.2, we deduce RIP conditions for IHT and NIHT which guarantee robust signal recovery. Since we perform worst-case RIP analysis alone in Section 3.3, we revert to Problem 1.2 in Section 3.3, relying

[1]Material in the following three chapters is in preparation for submission in [42], which is a joint authorship with C. Cartis whose permission has been obtained for the inclusion of the material.

upon Lemma 2.2 for the extension to k-compressible signals. We first introduce some notation.

Notation. Recalling from (1.13) that x_k^* is the best k-sparse approximation to x^*, we define the support set Λ to be

$$\Lambda := \text{supp}(x_k^*). \tag{3.1}$$

Note that Λ may also be viewed as the set of indices corresponding to the largest k coefficients of x^*. We will assume that x^* itself is at least k-sparse (it may not be sparse at all), so that x_k^* is exactly k-sparse and Λ has cardinality $|\Lambda| = k$. Recalling (2.6), the combined noise E due both to compression error and noise in the measurements may be written as

$$E := A(x^* - x_k^*) + e = A_{\Lambda^C} x_{\Lambda^C}^* + e. \tag{3.2}$$

Given some index set Γ of cardinality k, we also define

$$\tilde{e}_\Gamma := e + A_{(\Gamma\cup\Lambda)^C} x_{(\Gamma\cup\Lambda)^C}^*. \tag{3.3}$$

Note that the subscript in $\tilde{e}_\Gamma \in \mathbb{R}^n$ is used to highlight the dependence on the choice of support set Γ, but does not denote a restriction as is often the case for subscripts in this thesis.

3.1 Analysis of the stable point condition

Let us begin our considerations with IHT, the constant stepsize variant. Recalling the algorithm summary in Algorithm 2.1, let us write

$$\phi(x) := \mathcal{H}_k\{x + \alpha A^T(b - Ax)\}, \tag{3.4}$$

so that the IHT iteration can be expressed as $x^{m+1} = \phi(x^m)$. Then a *fixed point* of IHT is defined as any $\bar{x} \in \mathbb{R}^N$ such that

$$\phi(\bar{x}) = \bar{x}. \tag{3.5}$$

The following necessary conditions for some $\bar{x}$ to be a fixed point of IHT are a reformulation of those originally given for IHT with unit stepsize by Blumensath and Davies in [26].

Lemma 3.1 (IHT fixed point necessary condition [26]). *Suppose $\bar{x} \in \mathbb{R}^N$ is a fixed point of IHT with stepsize $\alpha > 0$. Then there exists some set Γ with supp$(\bar{x}) \subseteq \Gamma$ and $|\Gamma| = k$ such that the following two conditions hold:*

$$\left\{A^T(b - A\bar{x})\right\}_\Gamma = 0; \tag{3.6}$$

$$\min_{i\in\Gamma} |\bar{x}_i| \geq \alpha \max_{j\in\Gamma^C} |\left\{A^T(b - A\bar{x})\right\}_j|. \tag{3.7}$$

Proof: Suppose $\bar{x}$ is a fixed point of IHT with stepsize $\alpha > 0$ and let Γ be such that $\text{supp}(\bar{x}) \subseteq \Gamma$ and $|\Gamma| = k$. We denote

$$\bar{a} := \bar{x} + \alpha A^T(b - A\bar{x}) \tag{3.8}$$

so that

$$\phi(\bar{x}) = \mathcal{H}_k(\bar{a}). \tag{3.9}$$

We distinguish two cases: first suppose that $\text{supp}(\bar{x}) = \Gamma$. Then by (3.5), $\text{supp}(\phi(\bar{x})) = \Gamma$, which along with (3.9), implies that

$$\bar{a}_\Gamma = \bar{x}_\Gamma, \tag{3.10}$$

which in turn combines with (3.8) to give (3.6). Now (3.9) also combines with (1.26) to give

$$|\bar{a}_i| \geq |\bar{a}_j| \quad \text{for all } \ i \in \Gamma \ \text{ and all } \ j \in \Gamma^C. \tag{3.11}$$

Since $\text{supp}(\bar{x}) = \Gamma$, (3.8) implies that $\bar{a}_{\Gamma^C} = \alpha A^T_{\Gamma^C}(b - A\bar{x})$ which, together with (3.10) and (3.11), implies (3.7). Now suppose that $\text{supp}(\bar{x}) \subset \Gamma$. Then by (3.5), $\text{supp}(\phi(\bar{x})) = \text{supp}(\bar{x})$, which along with (3.9), implies that

$$\bar{a}_{\text{supp}(\bar{x})} = \bar{x}_{\text{supp}(\bar{x})},$$

which in turn combines with (3.8) to give

$$\left\{A^T(b - A\bar{x})\right\}_{\text{supp}(\bar{x})} = 0. \tag{3.12}$$

Since $|\text{supp}(\bar{x})| < k$, we also have $\bar{x}_i = 0$ for some $i \in \Gamma$, and it follows from (3.9) and (1.26) that $\bar{a}_{\Gamma^C} = 0$, and since $\text{supp}(\bar{x}) \subset \Gamma$, this implies

$$A^T_{\text{supp}(\bar{x})^C}(b - A\bar{x}) = 0. \tag{3.13}$$

We may now deduce (3.6) by combining (3.12) and (3.13), while (3.7) also follows trivially from (3.13) since $\Gamma^C \subset \text{supp}(\bar{x})^C$. □

Conditions (3.14) and (3.15) have a simple intuitive interpretation: if a further iteration of IHT is applied at a fixed point $\bar{x}$, there is no change in the support set, for which we require the gradient terms on the complement of the support of $\bar{x}$ to be suitably small, namely (3.7) must hold. Meanwhile, the coefficients on the support of $\bar{x}$ remain unchanged, for which we require the gradient on the support of $\bar{x}$ to be zero, namely (3.6) must hold.

If the inequality in (3.7) is replaced by a strict inequality, then it can be shown that the conditions in Lemma 3.1 are also sufficient for $\bar{x}$ to be a fixed point of IHT [26].

However, when it comes to NIHT, if (3.6) holds, the exact linesearch stepsize choice (2.2) with which the NIHT stepsize is initialized is not well-defined since the numerator and denominator are both zero. To address this issue, we introduce the concept of an $\underline{\alpha}$-*stable point*.

Definition 3.2 (**$\underline{\alpha}$-stable points of generic IHT**). *Given $\underline{\alpha} > 0$ and an index set Γ with $|\Gamma| = k$, we say $\bar{x} \in \mathbb{R}^N$ is an $\underline{\alpha}$-stable point of generic IHT on Γ if supp$(\bar{x}) \subseteq \Gamma$ and*

$$\left\{A^T(b - A\bar{x})\right\}_\Gamma = 0 \quad \textit{and} \tag{3.14}$$

$$\min_{i\in\Gamma} |\bar{x}_i| \geq \underline{\alpha} \max_{j\in\Gamma^C} |\left\{A^T(b - A\bar{x})\right\}_j|. \tag{3.15}$$

Note that Γ is defined to be a set of cardinality k and so is not necessarily the support of $\bar{x}$; rather the support of $\bar{x}$ is contained within Γ. We focus entirely on generic IHT in the present chapter and in Chapter 5, and so for brevity's sake we will generally drop the 'of generic IHT' label. Often, we will not need to specify a set Γ, and we will simply refer to an $\bar{x}$ satisfying Definition 3.2 as an $\underline{\alpha}$-stable point. For IHT, we can view an $\underline{\alpha}$-stable point as a generalization of the notion of a fixed point. In particular, note that, by Lemma 3.1, any fixed point of IHT with stepsize $\alpha > 0$ is an α-stable point of generic IHT.

In general, we will be interested in values of $\underline{\alpha}$ which lower bound the stepsize α^m of generic IHT. We next show that any $\underline{\alpha}$-stable point may be characterized as a minimum-norm solution on some k-subspace.

Lemma 3.3. *Suppose Assumption 1 holds and suppose $\bar{x}$ is an $\underline{\alpha}$-stable point of generic IHT on Γ for some $\underline{\alpha} > 0$. Then $\bar{x}_\Gamma = A_\Gamma^\dagger b$, where the Moore-Penrose pseudoinverse $A_\Gamma^\dagger$ is defined in (2.7).*

Proof: It follows from (3.14) that $A_\Gamma^T(b - A_\Gamma \bar{x}_\Gamma) = 0$ where supp$(\bar{x}) \subseteq \Gamma$ and $|\Gamma| = k$. Under Assumption 1, the pseudoinverse $A_\Gamma^\dagger$ in (2.7) is well-defined and we may rearrange to give $\bar{x}_\Gamma = A_\Gamma^\dagger b$. □

While the previous lemma tells us that any stable point is necessarily a minimum-norm solution on some k-subspace, the converse may not hold. In the case of Problem 1.3, we next deduce a necessary condition for a stable point on a given support which is only in terms of x^*, A and e and their restrictions to certain support sets.

Theorem 3.4 (**Stable point condition**). *Consider Problem 1.3. Suppose Assumption 1 holds and suppose there exists an $\underline{\alpha}$-stable point on some Γ such that $\Gamma \neq \Lambda$. Then*

$$\left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| + \left\|x^*_{\Gamma\setminus\Lambda}\right\| + \left\|A_\Gamma^\dagger \tilde{e}_\Gamma\right\| \geq \underline{\alpha}\left\{\left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma}\right\| - \left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)\tilde{e}_\Gamma\right\|\right\}, \tag{3.16}$$

where Λ is defined in (3.1) and $\tilde{e}_\Gamma$ is defined in (3.3).

Proof: Assume $\bar{x}$ is an $\underline{\alpha}$-stable point on Γ. Since $\Gamma \setminus \Lambda \subseteq \Gamma$ and $\Lambda \setminus \Gamma \subseteq \Gamma^C$, (3.15) implies that

$$\min_{i\in\Gamma\setminus\Lambda} |\bar{x}_i| \geq \underline{\alpha} \max_{j\in\Lambda\setminus\Gamma} |\left\{A^T(b - A\bar{x})\right\}_j|. \tag{3.17}$$

Definition 3.2 implies that $|\Gamma| = |\Lambda|$, and therefore $|\Gamma \setminus \Lambda| = |\Lambda \setminus \Gamma|$. This, properties of the Euclidean norm and (3.17) provide

$$\|\bar{x}_{\Gamma\setminus\Lambda}\|^2 \geq |\Gamma \setminus \Lambda| \left\{\min_{i\in\Gamma\setminus\Lambda} |\bar{x}_i|\right\}^2 \geq |\Lambda \setminus \Gamma| \left\{\underline{\alpha} \max_{j\in\Lambda\setminus\Gamma} |\left\{A^T(b - A\bar{x})\right\}_j|\right\}^2 \geq \underline{\alpha}^2 \|A^T_{\Lambda\setminus\Gamma}(b - A\bar{x})\|^2. \tag{3.18}$$

Problem 1.3 and (3.3) imply

$$b = Ax^* + e = A_\Gamma x^*_\Gamma + A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + A_{(\Lambda\cup\Gamma)^C}x^*_{(\Lambda\cup\Gamma)^C} + e = A_\Gamma x^*_\Gamma + A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + \tilde{e}_\Gamma. \tag{3.19}$$

This and Lemma 3.3 now provide, under Assumption 1,

$$\bar{x}_\Gamma = A_\Gamma^\dagger b = x^*_\Gamma + A_\Gamma^\dagger A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + A_\Gamma^\dagger \tilde{e}_\Gamma,$$

where in the last equality, we used $A_\Gamma^\dagger A_\Gamma = I$. Therefore,

$$\left\|\bar{x}_{\Gamma\setminus\Lambda}\right\| = \left\|\left(x^*_\Gamma + A_\Gamma^\dagger A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + A_\Gamma^\dagger\tilde{e}_\Gamma\right)_{\Gamma\setminus\Lambda}\right\| \leq \left\|x^*_{\Gamma\setminus\Lambda}\right\| + \left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma}\right\| + \left\|A_\Gamma^\dagger\tilde{e}_\Gamma\right\|, \tag{3.20}$$

which upper bounds the left-hand side of (3.18). Under Assumption 1, we may next use Lemma 3.3 and (3.19) to express the right-hand side of (3.18) independently of $\bar{x}$, as

$$A^T_{\Lambda\setminus\Gamma}(b - A\bar{x}) = A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)b = A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)(A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + \tilde{e}_\Gamma),$$

where in the last equality, we used $A_\Gamma^\dagger A_\Gamma = I$. We therefore may deduce

$$\left\|A^T_{\Lambda\setminus\Gamma}(b - A\bar{x})\right\| \geq \left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma}\right\| - \left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)\tilde{e}_\Gamma\right\|. \tag{3.21}$$

Substituting (3.20) and (3.21) into (3.18), we arrive at (3.16). □

In the case of Problem 1.2, Theorem 3.4 simplifies to the following corollary.

Corollary 3.5 (**Exactly sparse signals; noisy measurements**). *Consider Problem 1.2. Suppose Assumption 1 holds and suppose there exists an $\underline{\alpha}$-stable point of generic IHT on some Γ such that $\Gamma \neq \Lambda$. Then*

$$\left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| + \left\|A_\Gamma^\dagger e\right\| \geq \underline{\alpha}\left\{\left\|A_{\Lambda\setminus\Gamma}^T(I - A_\Gamma A_\Gamma^\dagger)A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| - \left\|A_{\Lambda\setminus\Gamma}^T(I - A_\Gamma A_\Gamma^\dagger)e\right\|\right\}, \quad (3.22)$$

where Λ is defined in (3.1).

Proof: Since $x^*_{\Lambda^C} = 0$, we also have $\tilde{e}_\Gamma = e$ and $x^*_{\Gamma\setminus\Lambda} = 0$, and making both these substitutions in (3.16) yields the required result. □

Corollary 3.5 simplifies further in the case of Problem 1.1.

Corollary 3.6 (**Noiseless case**). *Consider Problem 1.1. Suppose Assumption 1 holds and suppose there exists an $\underline{\alpha}$-stable point of generic IHT on some Γ such that $\Gamma \neq \Lambda$. Then*

$$\left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| \geq \underline{\alpha}\left\|A_{\Lambda\setminus\Gamma}^T(I - A_\Gamma A_\Gamma^\dagger)A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\|, \quad (3.23)$$

where Λ is defined in (3.1).

Proof: The result follows immediately upon setting $e := 0$ in (3.22). □

3.2 Convergence analysis

In this section, we derive convergence conditions for both IHT and NIHT. We first introduce some more notation. Recalling (1.24), we let

$$g^m := \nabla\Psi(x^m) \text{ and } \Gamma^m := \operatorname{supp}(x^m) \text{ for all } m \geq 0, \quad (3.24)$$

where $\{x^m\}$ are the iterates of generic IHT.

Some useful properties of the iterates of the generic IHT algorithm are given in the next lemma.

Lemma 3.7. *The iterates of generic IHT satisfy, for all $m \geq 0$,*

$$\|x^{m+1} - x^m\|^2 + 2\alpha^m (g^m)^T (x^{m+1} - x^m) \leq 0 \quad (3.25)$$

and

$$\Psi(x^{m+1}) - \Psi(x^m) = (g^m)^T (x^{m+1} - x^m) + \frac{1}{2}\left\|A(x^{m+1} - x^m)\right\|^2. \quad (3.26)$$

Proof: Using (1.25) and (3.24), we may rewrite the generic IHT iteration (1.27) as

$$x^{m+1} = \underset{\|z\|_0 \leq k}{\arg\min} \|z - \{x^m - \alpha^m g^m\}\|^2,$$

from which we may deduce

$$\|x^{m+1} - (x^m - \alpha^m g^m)\|^2 \leq \|x^m - (x^m - \alpha^m g^m)\|^2 = (\alpha^m)^2 \|g^m\|^2,$$

which expands to give

$$\|x^{m+1} - x^m\|^2 + 2\alpha^m (g^m)^T (x^{m+1} - x^m) + (\alpha^m)^2 \|g^m\|^2 \leq (\alpha^m)^2 \|g^m\|^2,$$

and so (3.25) holds. Since $\Psi(x)$ in (1.24) is a quadratic function, we have no remainder in the following second-order Taylor expansion

$$\begin{aligned} \Psi(x^{m+1}) - \Psi(x^m) &= [\nabla \Psi(x^m)]^T (x^{m+1} - x^m) + \tfrac{1}{2}(x^{m+1} - x^m)^T \left[\nabla^2 \Psi\right] (x^{m+1} - x^m) \\ &= (g^m)^T (x^{m+1} - x^m) + \tfrac{1}{2}(x^{m+1} - x^m)^T A^T A (x^{m+1} - x^m), \end{aligned}$$

and so (3.26) follows. □

A sufficient condition for convergence of generic IHT is given next.

Lemma 3.8 (**Sufficient condition for convergence**). *Consider Problem 1.3. Suppose Assumption 1 holds, and suppose the iterates of generic IHT satisfy*

$$\|x^{m+1} - x^m\|^2 \leq d \left[\Psi(x^m) - \Psi(x^{m+1})\right] \quad \text{for all } m \geq 0, \tag{3.27}$$

for some $d > 0$ which does not depend upon m, where $\Psi(x)$ is defined in (1.24). Assume that there exist $\overline{\alpha} \geq \underline{\alpha} > 0$ such that

$$\overline{\alpha} \geq \alpha^m \geq \underline{\alpha} \quad \text{for all } m \geq 0. \tag{3.28}$$

Then $x^m \to \bar{x}$ as $m \to \infty$, where $\bar{x}$ is an $\underline{\alpha}$-stable point of generic IHT.

Proof: We deduce from (3.27) that

$$\sum_{m=0}^{\infty} \|x^{m+1} - x^m\|^2 \leq d \sum_{m=0}^{\infty} \left[\Psi(x^m) - \Psi(x^{m+1})\right] \leq d\Psi(x^0),$$

where to obtain the last inequality we used $\Psi(x^m) \geq 0$. Thus convergent series properties provide

$$\|x^{m+1} - x^m\| \to 0 \quad \text{as } m \to \infty. \tag{3.29}$$

From (1.27) and (3.24), we deduce

$$x^{m+1}_{\Gamma^{m+1}} = x^m_{\Gamma^{m+1}} - \alpha^m g^m_{\Gamma^{m+1}} \quad \text{and} \quad x^{m+1}_{(\Gamma^{m+1})^C} = 0.$$

Thus restricting (3.29) to Γ^{m+1} and using (3.28) provide

$$\|g^m_{\Gamma^{m+1}}\| \to 0 \quad \text{as } m \to \infty, \tag{3.30}$$

while restricting (3.29) to $\Gamma^m \setminus \Gamma^{m+1}$ yields

$$\|x^m_{\Gamma^m \setminus \Gamma^{m+1}}\| \to 0. \tag{3.31}$$

For $m \geq 0$, let y^m denote the minimum-norm solution on Γ^m, namely,

$$y^m_{\Gamma^m} := A^\dagger_{\Gamma^m} b \quad \text{and} \quad y^m_{(\Gamma^m)^C} := 0, \tag{3.32}$$

which is well-defined due to Assumption 1. Then (3.32) and $x^m_{(\Gamma^m)^C} = 0$ provide

$$\begin{aligned} \|y^{m+1} - x^m\| &\leq \|y^{m+1}_{\Gamma^{m+1}} - x^m_{\Gamma^{m+1}}\| + \|x^m_{(\Gamma^{m+1})^C}\| = \|A^\dagger_{\Gamma^{m+1}} b - x^m_{\Gamma^{m+1}}\| + \|x^m_{\Gamma^m \setminus \Gamma^{m+1}}\| \\ &= \|(A^T_{\Gamma^{m+1}} A_{\Gamma^{m+1}})^{-1} A^T_{\Gamma^{m+1}} (b - A_{\Gamma^{m+1}} x^m_{\Gamma^{m+1}})\| + \|x^m_{\Gamma^m \setminus \Gamma^{m+1}}\| \\ &= \|(A^T_{\Gamma^{m+1}} A_{\Gamma^{m+1}})^{-1} g^m_{\Gamma^{m+1}}\| + \|x^m_{\Gamma^m \setminus \Gamma^{m+1}}\| \to 0 \quad \text{as } m \to \infty, \end{aligned}$$

where the limit follows from (3.30), (3.31), Assumption 1 and the fact that there are finitely many distinct support sets Γ^m, $m \geq 0$. This and (3.29) further give

$$\|y^m - x^m\| \to 0 \quad \text{as } m \to \infty, \tag{3.33}$$

and so for any $\epsilon > 0$, there exists $m_0 \geq 0$ such that

$$\|y^m - x^m\| \leq \epsilon \quad \text{for all } m \geq m_0. \tag{3.34}$$

We denote the index set of changing minimum-norm solutions by $\mathcal{S} := \{m \geq m_0 : y^{m+1} \neq y^m\}$, and we will show that $\mathcal{S}$ is finite. Now set

$$\epsilon := \frac{1}{4} \min_{m \in \mathcal{S}} \|y^{m+1} - y^m\|. \tag{3.35}$$

Note that $\epsilon > 0$ since there are finitely many distinct support sets Γ^m, $m \geq 0$. Then, the triangle inequality, (3.34) and (3.35) yield

$$\|x^{m+1} - x^m\| \geq \|y^{m+1} - y^m\| - \|y^{m+1} - x^{m+1}\| - \|y^m - x^m\| \geq 4\epsilon - \epsilon - \epsilon > \epsilon \quad \text{for all } m \in \mathcal{S}.$$

This and (3.29) imply that $\mathcal{S}$ must be finite and so there exists $m_1 \geq m_0$ such that $y^{m+1} = y^m = \bar{x}$ for all $m \geq m_1$, where $\bar{x}_\Gamma = A_\Gamma^\dagger b$ and $\bar{x}_{\Gamma^C} = 0$, for some Γ with $|\Gamma| = k$. This and (3.33) give

$$x^m \to \bar{x} \quad \text{as } m \to \infty. \tag{3.36}$$

Clearly, (3.14) holds for the limit point $\bar{x}$ of the iterates $\{x^m\}$. To complete the proof, it remains to establish (3.15). The thresholding operation that defines x^{m+1} in generic IHT gives that

$$\min_{i \in \Gamma^{m+1}} |x_i^{m+1}| \geq \max_{j \in (\Gamma^{m+1})^C} |\{x^m - \alpha^m g^m\}_j| \quad \text{for all } m \geq 0, \tag{3.37}$$

and (3.28) implies that there exists a convergent subsequence of stepsizes,

$$\alpha^{m_r} \to \tilde{\alpha} \geq \underline{\alpha} \quad \text{as } r \to \infty. \tag{3.38}$$

Letting $\epsilon := \frac{1}{2} \min_{i \in \text{supp}(\bar{x})} \bar{x}_i$, (3.36) implies that $\|x^m - \bar{x}\| \leq \epsilon$, and so

$$\text{supp}(\bar{x}) \subseteq \Gamma^m, \quad \text{for all } m \text{ sufficiently large.} \tag{3.39}$$

Firstly, assume that $\text{supp}(\bar{x}) = \Gamma$. Then, since $|\Gamma| = |\Gamma^m| = k$, (3.39) implies that $\Gamma^m = \Gamma$ for all m sufficiently large which, together with (3.37), provides

$$\min_{i \in \Gamma} |x_i^{m+1}| \geq \max_{j \in \Gamma^C} |\{x^m - \alpha^m g^m\}_j| \quad \text{for all } m \text{ sufficiently large.} \tag{3.40}$$

Passing to the limit in (3.40) on the subsequence m_r for which (3.38) holds, using (3.36), $\bar{x}_{\Gamma^C} = 0$ and the right-hand side of (3.28) imply (3.15) holds in this case. It remains to consider the case when $\text{supp}(\bar{x}) \subset \Gamma$. Then $\min_{i \in \Gamma} |\bar{x}_i| = 0$ and so (3.36) further provides

$$\min_{i \in \Gamma^{m+1}} |x_i^{m+1}| \to 0 \quad \text{as } m \to \infty. \tag{3.41}$$

Now (3.39) and again (3.36) provide

$$x^m_{\Gamma^{m+1}} \to 0 \quad \text{as } m \to \infty. \tag{3.42}$$

Passing to the limit in (3.37) on the subsequence m_r for which (3.38) holds, and using (3.41) and (3.42), we obtain that $g^m_{(\Gamma^{m+1})^C} \to 0$ as $m \to \infty$. This and (3.30) now give that $g^m = A^T(Ax^m - b) \to 0$, which due to (3.36), implies that $A^T(b - A\bar{x}) = 0$ and so (3.15) trivially holds in this case. □

In [26], Blumensath and Davies prove convergence of the iterates of IHT to a single limit point under the assumption that $\alpha\|A\|_2 < 1$. Largely following their method of proof, we now

show that the latter requirement on the IHT stepsize can be weakened to a condition involving the RIP constant U_{2k} of A.

Theorem 3.9 (**IHT convergence**). *Consider Problem 1.3. Suppose that Assumption 1 holds, and suppose that the IHT stepsize satisfies*

$$\alpha < \frac{1}{1+U_{2k}}. \tag{3.43}$$

Then IHT with stepsize α converges to an α-stable point $\bar{x}$ of generic IHT.

Proof: Let $m \geq 0$. Since the support size of the change to the iterates $x^{m+1} - x^m$ is at most $2k$, (1.15) with $s = 2k$ provides $\|A(x^{m+1} - x^m)\|^2 \leq (1 + U_{2k})\|x^{m+1} - x^m\|^2$. Using this bound, and (3.25) with the choice (3.43), in (3.26), we obtain

$$\Psi(x^{m+1})-\Psi(x^m) \leq -\frac{1}{2\alpha}\|x^{m+1}-x^m\|^2+\frac{1}{2}(1+U_{2k})\|x^{m+1}-x^m\|^2 = \frac{\alpha(1+U_{2k})-1}{2\alpha}\|x^{m+1}-x^m\|^2,$$

which, due to (3.43), implies that (3.27) holds with $d := 2\alpha/[1 - \alpha(1 + U_{2k})]$. Due to (6.39), (3.28) trivially holds with $\overline{\alpha} = \underline{\alpha} = \alpha$. Thus Lemma 3.8 applies, and the IHT iterates x^m converge to an α-stable point. □

A similar result has also recently appeared in [93]. We next obtain a convergence result for NIHT. In this case, there is no explicit requirement for an RIP condition to be satisfied; however, the RIP this time appears in the choice of $\underline{\alpha}$.

Theorem 3.10 (**NIHT convergence**). *Suppose Assumption 1 holds. Then NIHT with shrinkage parameter κ converges to a $[\kappa(1 + U_{2k})]^{-1}$-stable point $\bar{x}$ of generic IHT.*

Proof: Firstly, we consider the case when the exact linesearch choice is accepted, so that α^m is given by (2.2). Then (3.24) implies $\Gamma^{m+1} = \Gamma^m$, and (1.27) implies

$$x^{m+1}_{\Gamma^m} = x^m_{\Gamma^m} - \alpha^m g^m_{\Gamma^m}. \tag{3.44}$$

Using (3.44), (2.2) becomes

$$\alpha^m = \frac{\|g^m_{\Gamma^m}\|^2}{\|A_{\Gamma^m} g^m_{\Gamma^m}\|^2} = \frac{\|x^{m+1} - x^m\|^2}{\|A(x^{m+1} - x^m)\|^2}. \tag{3.45}$$

Using that $x^{m+1} - x^m$ is supported on Γ^m, expressing $g^m_{\Gamma^m}$ from (3.44) and substituting into

(3.26), we deduce that

$$\begin{aligned}\Psi(x^{m+1}) - \Psi(x^m) &= -\frac{1}{\alpha^m}(x^{m+1}_{\Gamma^m} - x^m_{\Gamma^m})^T(x^{m+1}_{\Gamma} - x^m_{\Gamma}) + \frac{1}{2}\|A(x^{m+1} - x^m)\|^2 \\ &= -\frac{1}{\alpha^m}\|x^{m+1} - x^m\|^2 + \frac{1}{2\alpha^m}\|x^{m+1} - x^m\|^2 \\ &= -\frac{1}{2\alpha^m}\|x^{m+1} - x^m\|^2, \end{aligned} \tag{3.46}$$

where to obtain the second equality, we also used (3.45). Alternatively, when α^m is computed by shrinkage, we deduce that

$$\|A(x^{m+1} - x^m)\|^2 \leq \frac{1-c}{2\alpha^m}\|x^{m+1} - x^m\|^2.$$

Substituting this and (3.25) into (3.26), we obtain

$$\Psi(x^{m+1}) - \Psi(x^m) \leq -\frac{1}{2\alpha^m}\|x^{m+1} - x^m\|^2 + \frac{1-c}{2\alpha^m}\|x^{m+1} - x^m\|^2 = -\frac{c}{2\alpha^m}\|x^{m+1} - x^m\|^2. \tag{3.47}$$

Thus (3.46), (3.47) and $c \in (0, 1)$ imply that, for all $m \geq 0$,

$$\|x^{m+1} - x^m\|^2 \leq \frac{2\alpha^m}{c}[\Psi(x^m) - \Psi(x^{m+1})] \leq \frac{2(1-c)}{c(1-L_{2k})}[\Psi(x^m) - \Psi(x^{m+1})],$$

due to (2.3). Hence (3.27) holds with $d := 2(1-c)/[c(1-L_{2k})]$, and so does (3.28) due to (2.3). Lemma 3.8 applies and, together with (2.3), provides the required conclusion. □

3.3 RIP conditions based on the new recovery analysis

In this section, we use the RIP to translate the results in Sections 3.1 and 3.2 into new RIP-based recovery conditions for both IHT and NIHT. Since we choose to use the RIP, Lemma 2.2 and the comments thereafter apply, which show that any result for Problem 1.2 can be extended to a result for Problem 1.3. Hence, we may restrict our focus to Problem 1.2, noting that in this simplified case we have $E := e$. The next pivotal result proves that, under a certain RIP condition, any $\underline{\alpha}$-stable point of generic IHT approximates the original k-sparse signal x^* to within a fixed multiple of $\|e\|$.

Lemma 3.11 (Stable point error bounds). *Consider Problem 1.2. Suppose Assumption 1 holds and let $\bar{x}$ be an $\underline{\alpha}$-stable point of generic IHT. Define*

$$\pi_k^{\underline{\alpha}} := \underline{\alpha}\left[(1 - L_k) - \frac{(L_{2k} + U_{2k})^2}{4(1 - L_k)}\right] - \frac{(L_{2k} + U_{2k})}{2(1 - L_k)} \tag{3.48}$$

and

$$\xi_k^{\underline{\alpha}} := \frac{1}{\sqrt{1 - L_k}} + \underline{\alpha}\sqrt{1 + U_k}. \tag{3.49}$$

Provided $\pi_k^{\alpha} > 0$, $\bar{x}$ satisfies

$$\|\bar{x} - x^*\| \leq \sqrt{\left(\frac{\xi_k^{\alpha}}{\pi_k^{\alpha}}\right)^2 + \left[\frac{\xi_k^{\alpha}(L_{2k} + U_{2k})}{2\pi_k^{\alpha}(1 - L_k)} + \frac{1}{\sqrt{1 - L_k}}\right]^2} \|e\|. \tag{3.50}$$

Proof: Assume $\bar{x}$ is an $\underline{\alpha}$-stable point on Γ. Then, under Assumption 1, we may apply Corollary 3.5 to deduce (3.22). Next we use the RIP to upper bound the left-hand side of (3.22). By (2.10) and (2.12) of Lemma 2.3, we have

$$\left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| = \left\|(A_\Gamma^T A_\Gamma)^{-1} A_\Gamma^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| \leq \frac{1}{1 - L_k} \left\|A_\Gamma^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| \leq \frac{L_{2k} + U_{2k}}{2(1 - L_k)} \left\|x^*_{\Lambda\setminus\Gamma}\right\|, \tag{3.51}$$

and, by (2.11) of Lemma 2.3, we have

$$\left\|A_\Gamma^\dagger e\right\| \leq \frac{1}{\sqrt{1 - L_k}} \|e\|. \tag{3.52}$$

Similarly, we use the RIP to lower bound the right-hand side of (3.22). By (2.9), (2.12) and (2.10) of Lemma 2.3, we have

$$\begin{aligned}
\left\|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| &\geq \left\|A_{\Lambda\setminus\Gamma}^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| - \left\|A_{\Lambda\setminus\Gamma}^T A_\Gamma (A_\Gamma^T A_\Gamma)^{-1} A_\Gamma^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| \\
&\geq (1 - L_k) \left\|x^*_{\Lambda\setminus\Gamma}\right\| - \frac{(L_{2k} + U_{2k})^2}{4(1 - L_k)} \left\|x^*_{\Lambda\setminus\Gamma}\right\| \\
&= \left[(1 - L_k) - \frac{(L_{2k} + U_{2k})^2}{4(1 - L_k)}\right] \left\|x^*_{\Lambda\setminus\Gamma}\right\|,
\end{aligned} \tag{3.53}$$

and by (2.8) of Lemma 2.3, and since $(I - A_\Gamma A_\Gamma^\dagger)$ is a contraction, we have

$$\left\|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) e\right\| \leq \sqrt{1 + U_k} \left\|(I - A_\Gamma A_\Gamma^\dagger) e\right\| \leq \sqrt{1 + U_k} \|e\|. \tag{3.54}$$

Combining (3.22) with (3.51), (3.52), (3.53) and (3.54), we obtain

$$\frac{L_{2k} + U_{2k}}{2(1 - L_k)} \left\|x^*_{\Lambda\setminus\Gamma}\right\| + \frac{1}{\sqrt{1 - L_k}} \|e\| \geq \underline{\alpha} \left\{ \left[(1 - L_k) - \frac{(L_{2k} + U_{2k})^2}{4(1 - L_k)}\right] \left\|x^*_{\Lambda\setminus\Gamma}\right\| - \sqrt{1 + U_k} \|e\| \right\},$$

which may be rearranged to give

$$\pi_k^{\alpha} \left\|x^*_{\Lambda\setminus\Gamma}\right\| \leq \xi_k^{\alpha} \|e\|. \tag{3.55}$$

Under the assumption $\pi_k^{\alpha} > 0$, we may conclude

$$\left\|x^*_{\Lambda\setminus\Gamma}\right\| \leq \frac{\xi_k^{\alpha}}{\pi_k^{\alpha}} \|e\|. \tag{3.56}$$

We now bound the approximation error of $\bar{x}$. We may apply (3.56) to deduce

$$\left\|(\bar{x}-x^*)_{\Lambda\setminus\Gamma}\right\| = \left\|x^*_{\Lambda\setminus\Gamma}\right\| \leq \frac{\xi_k^\alpha}{\pi_k^\alpha}\|e\|. \tag{3.57}$$

Meanwhile, applying Lemma 3.3, and using $A_\Gamma^\dagger A_\Gamma = I$, we have

$$(\bar{x}-x^*)_\Gamma = A_\Gamma^\dagger b - x^*_\Gamma = A_\Gamma^\dagger A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + x^*_\Gamma + A_\Gamma^\dagger e - x^*_\Gamma = A_\Gamma^\dagger A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma} + A_\Gamma^\dagger e,$$

and we may therefore use the triangle inequality to bound

$$\|(\bar{x}-x^*)_\Gamma\| \leq \left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma}x^*_{\Lambda\setminus\Gamma}\right\| + \left\|A_\Gamma^\dagger e\right\|. \tag{3.58}$$

Applying (3.51) and (3.52) to (3.58), along with a further application of (3.56), leads to

$$\|(\bar{x}-x^*)_\Gamma\| \leq \frac{L_{2k}+U_{2k}}{2(1-L_k)}\left\|x^*_{\Lambda\setminus\Gamma}\right\| + \frac{1}{\sqrt{1-L_k}}\|e\| \leq \left[\frac{\xi_k^\alpha(L_{2k}+U_{2k})}{2\pi_k^\alpha(1-L_k)} + \frac{1}{\sqrt{1-L_k}}\right]\|e\|. \tag{3.59}$$

Since $(\bar{x}-x^*)$ is supported on $\Gamma\cup\Lambda$, we may write

$$\|\bar{x}-x^*\|^2 = \left\|(\bar{x}-x^*)_{\Lambda\setminus\Gamma}\right\|^2 + \|(\bar{x}-x^*)_\Gamma\|^2,$$

from which (3.50) now follows upon substitution of (3.57) and (3.59). □

We may use Lemma 3.11 to deduce the following RIP-based recovery result for IHT.

Theorem 3.12 (IHT). *Suppose Assumption 1 holds, and define*

$$\alpha^{min} := \frac{2(L_{2k}+U_{2k})}{4(1-L_k)^2-(L_{2k}+U_{2k})^2} \quad \textit{and} \quad \alpha^{max} := \frac{1}{1+U_{2k}}. \tag{3.60}$$

Provided

$$\alpha^{min} < \alpha^{max}, \tag{3.61}$$

and provided α satisfies

$$\alpha \in (\alpha^{min}, \alpha^{max}), \tag{3.62}$$

IHT with stepsize α converges to $\bar{x}$ such that (3.50) holds with $\underline{\alpha} := \alpha$.

Proof: By (3.60) and (3.61), any α chosen according to (3.62) satisfies (3.43). Therefore, under Assumption 1, we may apply Theorem 3.9 and conclude that IHT converges to some α-stable point $\bar{x}$. On the other hand, simple rearrangement shows that, if α is chosen according to (3.62), then $\pi_k^\alpha > 0$, where π_k^α is defined in (3.48). Hence, under Assumption 1, we may apply Lemma 3.11 with $\underline{\alpha} := \alpha$, giving the desired result. □

Theorem 3.12 guarantees recovery provided the stepsize α falls within some admissible interval, whenever the interval is well-defined. In fact, the stable point condition yields a lower bound on the stepsize, while the convergence condition enforces an upper bound on the stepsize. It has been observed empirically [72] that care must be taken to ensure that the IHT stepsize is neither too small or too large. Our analysis suggests a simple interpretation of this observation: the stepsize must be small enough to ensure that the algorithm converges, but large enough to ensure that it does not converge to fixed points other than the original signal.

It is common in the CS literature to give RIP-based recovery guarantees in terms of a single symmetric RIP constant (see Definition 1.2). In this regard, it is straightforward to deduce from Theorem 3.12 that, provided

$$R_{2k} < \frac{\sqrt{13}-3}{2} \approx 0.303,$$

stable recovery of IHT is guaranteed for any α falling in the interval

$$\frac{R_{2k}}{1-2R_{2k}} < \alpha < \frac{1}{1+R_{2k}}.$$

However, weaker recovery conditions in terms of R_{2k} already exist for IHT: the current best known such condition is $R_{2k} < 1/3$, obtained by Garg and Khandekar [96] for the stepsize choice $\alpha := 1/(1+R_{2k})$.

Next we give a recovery result for NIHT.

Theorem 3.13 (**NIHT**). *Suppose Assumption 1 holds, and suppose that*

$$(L_{2k}+U_{2k})\big[L_{2k}+U_{2k}+2\kappa(1+U_{2k})\big] < 4(1-L_k)^2. \tag{3.63}$$

Then NIHT with shrinkage parameter κ converges to $\bar{x}$ such that (3.50) holds with $\underline{\alpha} := [\kappa(1+U_{2k})]^{-1}$.

Proof: Since Assumption 1 holds, we may apply Theorem 3.10 and conclude that NIHT converges to some $\underline{\alpha}$-stable point $\bar{x}$, where

$$\underline{\alpha} := [\kappa(1+U_{2k})]^{-1}. \tag{3.64}$$

On the other hand, if $\underline{\alpha}$ is given by (3.64), simple rearrangement shows that (3.63) is equivalent to $\pi_k^{\underline{\alpha}} > 0$, where $\pi_k^{\underline{\alpha}}$ is defined in (3.48). Hence, under Assumption 1, we may apply Lemma 3.11, concluding that $\bar{x}$ satisfies (3.50) for $\underline{\alpha}$ given by (3.64). □

The result in Theorem 3.13 degrades as the shrinkage parameter κ is increased. As $\kappa \to 1$, the recovery condition for NIHT in Theorem 3.13 approaches the recovery condition for IHT in

Theorem 3.12.

In the context of symmetric RIP, it is straightforward to deduce from Theorem 3.13 that robust recovery is guaranteed for NIHT with stepsize κ provided

$$R_{2k} < \frac{\sqrt{\kappa^2 + 8\kappa + 4} - (\kappa + 2)}{2\kappa},$$

which gives, for example, $R_{2k} < (\sqrt{1401} - 31)/22 \approx 0.292$ for $\kappa := 1.1$, and which represents an improvement upon the previous best known recovery condition for NIHT in terms of R_{2k}, namely $R_{2k} < 1/9$ [27], which holds for NIHT with no backtracking.

Without quantifying the results in this section, it is unclear how they compare with the results in Chapter 2. We will present such a comparison in the phase transition framework for Gaussian matrices at the start of Chapter 5, where we will argue for the need for average case analysis in order to obtain significantly improved phase transitions for IHT algorithms.

Chapter 4

Large deviations tools for Gaussian matrices

In Chapter 2, we obtained phase transitions for IHT algorithms in the case of Gaussian matrices by making use of numerically computable upper bounds on RIP constants [4]. These RIP bounds hold with overwhelming probability in the proportional-dimensional asymptotic defined in Definition 1.5. In this chapter, we develop asymptotic bounds in the proportional-dimensional framework for other quantities related to Gaussian matrices. First, in Section 4.1, we analyse the stable point condition from Theorem 3.4, assuming Gaussian measurement noise, and assuming independence of the original signal, measurement matrix and measurement noise. By making use of standard results concerning Rayleigh quotients of Gaussian-related matrices, we derive the precise distribution of bounds for each term in (3.16) in terms of the χ^2 and F distributions.

There are combinatorially many possible support sets for a stable point, and each gives rise to instances of the distributions arising from (3.16). In order to bound over all support sets, the remainder of the chapter is devoted to proving large deviations bounds in the proportional-dimensional asymptotic for a combinatorial number of χ^2- and F-distributed random variables. In Section 4.2, we first establish results concerning the asymptotic behaviour of the regularized incomplete gamma and beta functions when their parameters grow proportionally. Since these functions are the distribution functions for the χ^2 and F distributions respectively, we are then able to draw upon these results in Section 4.3 to prove our large deviations bounds. Furthermore, the tail bounds we derive for the χ^2 distribution may be viewed as a special version of the RIP in which the matrix and vector are assumed to be independent, and Section 4.3 also includes a discussion comparing the two notions.

4.1 Distribution results for Rayleigh quotients

The aim of this section is to derive distribution results in the context of Gaussian measurement matrices for each of the terms in the stable point condition (3.16) of Theorem 3.4. We first give some definitions of Gaussian and Gaussian-related matrix variate distributions, along with some fundamental results concerning their Rayleigh quotients when applied to independent vectors.

4.1.1 Definitions and preliminary lemmas

We will consider a particular kind of matrix variate Gaussian distribution in which all entries are i.i.d. Gaussian random variables.

Definition 4.1 (**Matrix variate Gaussian distribution**). *We say that an $s \times t$ matrix B follows the matrix variate Gaussian distribution $B \sim \mathcal{N}_{s,t}(\mu, \sigma^2)$, if each entry of B independently follows the (univariate) Gaussian distribution $B_{ij} \sim \mathcal{N}(\mu, \sigma^2)$.*

The matrix variate distribution is usually defined in the more general setting where entries are not necessarily independent, see for example [103], but we choose to restrict to the case of independent entries in this thesis. We will in fact further restrict our attention to the *central* matrix variate Gaussian distribution, in which $\mu = 0$. In this case, the probability density function (pdf) of the matrix variate Gaussian distribution defined here has the following elegant form.

Lemma 4.2 (**Matrix variate Gaussian pdf**). *Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$. Then the pdf of B is*

$$f_B(M) = \frac{1}{(2\pi\sigma^2)^{\frac{st}{2}}} e^{-\frac{\|M\|_F^2}{\sigma^2}}, \tag{4.1}$$

where

$$\|M\|_F := \sqrt{\sum_{i=1}^{s}\sum_{j=1}^{t} M_{ij}^2}$$

is the Frobenius norm.

Proof: Since each entry B_{ij} of B is i.i.d. $\mathcal{N}(0, \sigma^2)$, we have

$$f_B(M) = \prod_{i=1}^{s}\prod_{j=1}^{t} f_{B_{ij}}(M_{ij}) = \prod_{i=1}^{s}\prod_{j=1}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{M_{ij}^2}{\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{\frac{st}{2}}} e^{\left\{-\frac{1}{\sigma^2}\sum_{i=1}^{s}\sum_{j=1}^{t} M_{ij}^2\right\}},$$

and the result now follows. □

Crucial to our argument will be the well-known result that the central matrix variate Gaussian distribution defined in Definition 4.1 is invariant under transformation by an independent orthogonal matrix.

Lemma 4.3 (**Orthogonal invariance**). *Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$ and let $Z_1 \in \mathbb{R}^{s\times s}$ and $Z_2 \in \mathbb{R}^{t\times t}$ be orthogonal and independent of B. Then*

$$Z_1 B \sim \mathcal{N}_{s,t}(0, \sigma^2), \quad \textit{independently of } Z_1, \tag{4.2}$$

and

$$B Z_2 \sim \mathcal{N}_{s,t}(0, \sigma^2), \quad \textit{independently of } Z_2. \tag{4.3}$$

Proof: Let us write $Y := Z_1 B$, and first consider the case where Z_1 is a fixed orthogonal matrix. Then $\|Y\|_F = \|Z_1 B\|_F = \|B\|_F$ by the orthogonal invariance of the Frobenius norm, and since the Jacobian of the transformation $Z \to Y$ is equal to 1, and using Lemma 4.2, it then follows that the pdf of Y satisfies

$$f_Y(M) = \frac{1}{(2\pi\sigma^2)^{\frac{st}{2}}} e^{-\frac{\|M\|_F^2}{\sigma^2}} = f_B(M),$$

and therefore that $Y \sim \mathcal{N}_{s,t}(0, \sigma^2)$, independently of Z_1. Now suppose Z_1 is random but independent of B. Then, by the preceding argument, the conditional distribution $Y|Z_1$ satisfies $Y|Z_1 \sim \mathcal{N}_{s,t}(0, \sigma^2)$, independently of Z_1, yielding (4.2). The proof of (4.3) is analogous. □

We can further deduce from Lemma 4.3 the following result.

Lemma 4.4. *Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$ and let $z \in \mathbb{R}^t$ be a fixed vector. Then $Bz \sim \mathcal{N}_{s,1}(0, \sigma^2\|z\|^2)$.*

Proof: Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$ and let $Z \in \mathbb{R}^{t\times t}$ be an orthogonal matrix such that its first column is z normalized, so that

$$Z = \left[\frac{z}{\|z\|}\middle| Z_2\right].$$

Then by Lemma 4.3, $BZ \sim \mathcal{N}_{s,t}(0, \sigma^2)$. Since $\frac{Bz}{\|z\|}$ is a submatrix of BZ, we therefore have $\frac{Bz}{\|z\|} \sim \mathcal{N}_{s,1}(0, \sigma^2)$, from which the result follows. □

Let us also define the matrix variate Wishart distribution, which is closely related to the matrix variate Gaussian distribution.

Definition 4.5 (**Matrix variate Wishart distribution**). *Let $B \sim \mathcal{N}_{s,t}(\mu, \sigma^2)$ such that $s \geq t$. Then we say that $B^T B$ follows a matrix variate Wishart distribution $\mathcal{W}_t(s; \mu, \sigma^2)$ with s degrees of freedom, mean μ and variance σ^2.*

We will state distribution results in terms of the (univariate) χ^2 and F distributions, which we first formally define.

Definition 4.6 (χ^2 **and** F **distributions [1, pp.940,946]**). *Given a positive integer s, let $Z_i \sim \mathcal{N}(0, 1)$ for $1 \leq i \leq s$. Then we say $P = Z_1^2 + Z_2^2 + \ldots + Z_s^2$ follows a chi-squared distribution with s degrees of freedom, and we write $P \sim \chi_s^2$. Furthermore, given positive integers s and t, if $P \sim \frac{1}{s}\chi_s^2$ and $Q \sim \frac{1}{t}\chi_t^2$ are independent random variables, we say that P/Q follows the F-distribution, and we write $P/Q \sim \mathcal{F}(s, t)$.*

Useful results concerning the distributions of Rayleigh quotients related to Gaussian and Wishart matrices are given in the next lemma.

Lemma 4.7 (**Distributions of Rayleigh quotients**). *Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$ with $s \geq t$. Let $z \in \mathbb{R}^t$ be independent of B, and such that $\mathbb{P}(z \neq 0) = 1$. Then*

$$\frac{z^T B^T B z}{z^T z} \sim \sigma^2 \chi^2_s \quad \textit{and is independent of } z; \tag{4.4}$$

$$\frac{z^T z}{z^T (B^T B)^{-1} z} \sim \sigma^2 \chi^2_{s-t+1} \quad \textit{and is independent of } z; \tag{4.5}$$

$$\frac{z^T (B^T B)^2 z}{z^T z} \quad \textit{has the same distribution as } \left\{(B^T B)^2\right\}_{11}. \tag{4.6}$$

Proof: Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$ with $s \geq t$. [103, Theorem 3.3.12] gives a more general result than (4.4) for when the entries of B are not necessarily independent. The present result follows by setting $\Sigma = \sigma^2 I$ for the covariance matrix. Similarly, (4.5) follows by setting $\Sigma = \sigma^2 I$ in [103, Corollary 3.3.14.1]. To prove (4.6), let $S = B^T B$ so that $S \sim \mathcal{W}_t(s; 0, \sigma^2)$ and let $Z \in \mathbb{R}^{t \times t}$ be any orthogonal matrix which is independent of B. Lemma 4.3 yields $BZ \sim \mathcal{N}_{s,t}(0, \sigma^2)$ independently of Z, and, writing $T := Z^T S Z$, we therefore have

$$T = Z^T S Z = Z^T B^T B Z = (BZ)^T BZ \sim \mathcal{W}_t(s; 0, \sigma^2), \tag{4.7}$$

independently of Z. In particular, let us fix the first column of Z as z normalized so that

$$Z = \left[\frac{z}{\|z\|} \middle| Z_2 \right],$$

which leads to

$$\begin{aligned} z^T S^2 z &= z^T (ZTZ^T)^2 z = z^T ZTZ^T ZTZ^T z = z^T ZT^2 Z^T z \\ &= z^T \left[\frac{z}{\|z\|} \middle| Z_2 \right] T^2 \left[\frac{\frac{z^T}{\|z\|}}{Z_2^T} \right] z = [\, \|z\| \mid 0 \,] T^2 \left[\frac{\|z\|}{0} \right] = (T^2)_{11} \|z\|^2. \end{aligned}$$

Dividing by $\|z\|^2$ and using (4.7) then gives the desired result. □

Equipped with the necessary tools, we proceed in the next section to obtain our desired distribution results.

4.1.2 Analysis of the stable point condition for Gaussian matrices

We now make the assumption that the measurement matrix is drawn from the (central) matrix variate Gaussian distribution with appropriate normalization.

Assumption 2 (**Independent Gaussian measurements**). *The measurement matrix A has i.i.d. $\mathcal{N}(0, 1/n)$ entries, so that $A \sim \mathcal{N}_{n,N}(0, 1/n)$. Furthermore, A is independent of x^*.*

We are thus considering precisely the same family of matrices for which RIP-based recovery

phase transitions were obtained for IHT algorithms in Chapter 2. Furthermore, given Assumption 2, the following result from [103] allows us to dispense with Assumption 1.

Lemma 4.8 (**[103, Theorem 3.2.1]**). *Let $B \sim \mathcal{N}_{s,t}(0, \sigma^2)$ where $s \geq t$. Then $B^T B$ is positive definite with probability* 1.

Provided $2k \leq n$, Assumption 2 now implies Assumption 1, since Theorem 4.8 holds for all the $2k$-column submatrices of A.

As in Chapters 2 and 3, we will assume measurements to be noisy, but from now on we will impose the further assumption that measurement noise is itself Gaussian and independent of the original signal.

Assumption 3 (**Independent Gaussian noise**). *The noise vector e has i.i.d. Gaussian entries $e_i \sim N(0, \sigma^2/n)$, independently of both A and x.*

Note that, under Assumption 3, $\mathbb{E}\|e\|^2 = \sigma^2$, so that $\|e\| \approx \sigma$.

We saw in Section 2.2 that any result for exactly sparse signals (Problem 1.2) can be extended by means of the RIP to a result for compressible signals (Problem 1.3). However, our aim in the rest of this thesis is to perform average case analysis, and therefore we choose to dispense with worst-case tools such as the RIP wherever possible. We will instead directly prove results applicable to noisy measurements and compressible signals. Unless explicitly stated otherwise, in all that follows in this and the next chapter, we are considering solving Problem 1.3. Recalling (1.13) and (3.1), we define the *unrecoverable energy* Σ of the problem, due both to noise in the measurements and to the signal being only approximately sparse, as follows.

Definition 4.9 (**Unrecoverable energy**). *Given some index set Γ of cardinality k, define*

$$\tilde{\sigma}_\Gamma := \sqrt{\sigma^2 + \|x^*_{(\Gamma \cup \Lambda)^C}\|^2}, \tag{4.8}$$

and define

$$\Sigma := \sigma + \|x^*_{\Lambda^C}\|. \tag{4.9}$$

We now give the main result of this section, in which we derive precise distributions for various expressions which make up the stable point condition (3.16) of Theorem 3.4, in terms of the χ^2 and F distributions.

Lemma 4.10 (**Distribution results for the stable point condition**). *Suppose Assumptions 2 and 3 hold, and let Γ and Λ be index sets of cardinality k, where $k < n$, such that $\Gamma \neq \Lambda$. Then*

$$\frac{\|A^\dagger_\Gamma A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\|}{\|x^*_{\Lambda\setminus\Gamma}\|} = \sqrt{F_\Gamma}, \quad \textit{where} \quad F_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1); \tag{4.10}$$

$$\frac{\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A^\dagger_\Gamma) A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\|}{\|x^*_{\Lambda\setminus\Gamma}\|} \geq \left(\frac{n-k}{n}\right) \cdot R_\Gamma, \quad \textit{where} \quad R_\Gamma \sim \frac{1}{n-k} \chi^2_{n-k}. \tag{4.11}$$

$$\|A_\Gamma^\dagger \tilde{e}_\Gamma\| \le \tilde{\sigma}_\Gamma \cdot \sqrt{G_\Gamma}, \quad \textit{where} \quad G_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1); \tag{4.12}$$

$$\|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger)\tilde{e}_\Gamma\| \le \tilde{\sigma}_\Gamma \sqrt{\frac{k(n-k)}{n^2} \cdot (S_\Gamma)(T_\Gamma)}, \quad \textit{where} \quad S_\Gamma \sim \frac{1}{n-k}\chi^2_{n-k}, \quad T_\Gamma \sim \frac{1}{k}\chi^2_k, \tag{4.13}$$

where $\tilde{e}_\Gamma$ is defined in (3.3) and $\tilde{\sigma}_\Gamma$ is defined in (4.8).

Proof of (4.10): Let A_Γ have the singular value decomposition

$$A_\Gamma := U[D \mid 0]V^T = U_1 D V^T, \tag{4.14}$$

where $D \in \mathbb{R}^{k\times k}$ is diagonal, and where $V \in \mathbb{R}^{k\times k}$ and $U = [U_1 \mid U_2] \in \mathbb{R}^{n\times n}$ are orthogonal, with $U_1 \in \mathbb{R}^{n\times k}$. By Assumption 2, $A_\Gamma^\dagger$ is well-defined and we have the standard result

$$A_\Gamma^\dagger = V D^{-1} U_1^T, \tag{4.15}$$

and since $(A_\Gamma^T A_\Gamma)^{-1} = V D^{-2} V^T$, it follows by rearrangement that

$$D^{-2} = V^T (A_\Gamma^T A_\Gamma)^{-1} V. \tag{4.16}$$

Using (4.15) and (4.16), we have

$$\begin{aligned}
\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\|^2 &= (x^*_{\Lambda\setminus\Gamma})^T A_{\Lambda\setminus\Gamma}^T (A_\Gamma^\dagger)^T (A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma} \\
&= (x^*_{\Lambda\setminus\Gamma})^T A_{\Lambda\setminus\Gamma}^T U_1 D^{-1} V^T V D^{-1} U_1^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma} \\
&= (x^*_{\Lambda\setminus\Gamma})^T A_{\Lambda\setminus\Gamma}^T U_1 D^{-2} U_1^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma} \\
&= (x^*_{\Lambda\setminus\Gamma})^T A_{\Lambda\setminus\Gamma}^T U_1 V^T (A_\Gamma^T A_\Gamma)^{-1} V U_1^T A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}.
\end{aligned} \tag{4.17}$$

By Lemma 4.3, we have $U^T A_{\Lambda\setminus\Gamma} \sim \mathcal{N}_{n,r}(0, 1/n)$, independently of U, where $r := |\Lambda \setminus \Gamma|$. Since $U_1^T A_{\Lambda\setminus\Gamma}$ is a submatrix of $U^T A_{\Lambda\setminus\Gamma}$, it follows that $U_1^T A_{\Lambda\setminus\Gamma} \sim \mathcal{N}_{k,r}(0, 1/n)$, independently of U. Writing $C := V U_1^T A_{\Lambda\setminus\Gamma} \in \mathbb{R}^{k\times r}$, we also have by Lemma 4.3 that $C \sim \mathcal{N}_{k,r}(0, 1/n)$, independently of both U and V, and therefore independently of A_Γ. Substituting for C in (4.17), we have

$$\begin{aligned}
\frac{\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\|^2}{\|x^*_{\Lambda\setminus\Gamma}\|^2} &= \frac{(x^*_{\Lambda\setminus\Gamma})^T C^T (A_\Gamma^T A_\Gamma)^{-1} C x^*_{\Lambda\setminus\Gamma}}{(x^*_{\Lambda\setminus\Gamma})^T x^*_{\Lambda\setminus\Gamma}} \\
&= \frac{(x^*_{\Lambda\setminus\Gamma})^T C^T (A_\Gamma^T A_\Gamma)^{-1} C x^*_{\Lambda\setminus\Gamma}}{(x^*_{\Lambda\setminus\Gamma})^T C^T C x^*_{\Lambda\setminus\Gamma}} \cdot \frac{(x^*_{\Lambda\setminus\Gamma})^T C^T C x^*_{\Lambda\setminus\Gamma}}{(x^*_{\Lambda\setminus\Gamma})^T x^*_{\Lambda\setminus\Gamma}},
\end{aligned} \tag{4.18}$$

where x^*, C and A_Γ are all independent. Now it follows from Lemma 4.7 that

$$\frac{(x^*_{\Lambda\setminus\Gamma})^T C^T C x^*_{\Lambda\setminus\Gamma}}{(x^*_{\Lambda\setminus\Gamma})^T x^*_{\Lambda\setminus\Gamma}} \sim \frac{1}{n}\,\chi^2_k \quad \text{and} \quad \frac{(x^*_{\Lambda\setminus\Gamma})^T C^T C x^*_{\Lambda\setminus\Gamma}}{(x^*_{\Lambda\setminus\Gamma})^T C^T (A_\Gamma^T A_\Gamma)^{-1} C x^*_{\Lambda\setminus\Gamma}} \sim \frac{1}{n}\,\chi^2_{n-k+1}, \tag{4.19}$$

where both distributions are independent of each other. Combining (4.18) and (4.19) leads us to conclude

$$\frac{\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\|^2}{\|x^*_{\Lambda\setminus\Gamma}\|^2} \sim \frac{\chi_k^2}{\chi_{n-k+1}^2} = \frac{k}{n-k+1}\mathcal{F}(k, n-k+1),$$

where in the last step we use the fact that the two distributions are independent, which proves (4.10).

Proof of (4.11): Using (4.14) and (4.15), we have

$$A_\Gamma A_\Gamma^\dagger = U_1 D V^T V D^{-1} U_1^T = U_1 U_1^T = U \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} U^T,$$

and writing $I = UU^T$,

$$I - A_\Gamma A_\Gamma^\dagger = U \left\{ \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \right\} U^T = U \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} U^T = U_2 U_2^T, \tag{4.20}$$

which in turn gives

$$A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} = A_{\Lambda\setminus\Gamma}^T U_2 U_2^T A_{\Lambda\setminus\Gamma}. \tag{4.21}$$

Writing $F := U_2^T A_{\Lambda\setminus\Gamma}$, we have $U^T A_{\Lambda\setminus\Gamma} \sim \mathcal{N}_{n,r}(0, 1/n)$ by Lemma 4.3, and since $U_2^T A_{\Lambda\setminus\Gamma} \in \mathbb{R}^{(n-k)\times r}$ is a submatrix of $U^T A_{\Lambda\setminus\Gamma}$, it follows that

$$F \sim \mathcal{N}_{(n-k),r}(0, 1/n). \tag{4.22}$$

Substituting for F in (4.21) gives

$$A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} = F^T F. \tag{4.23}$$

Now, writing $M := F^T F$, and using (4.23) and (4.6) of Lemma 4.7, we deduce

$$\frac{\|A_{\Lambda\setminus\Gamma}^T (I - A_\Gamma A_\Gamma^\dagger) A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\|^2}{\|x^*_{\Lambda\setminus\Gamma}\|^2} = \frac{\|F^T F x_{\Lambda\setminus\Gamma}\|^2}{\|x^*_{\Lambda\setminus\Gamma}\|^2} = \frac{(x^*_{\Lambda\setminus\Gamma})^T (F^T F)^2 x^*_{\Lambda\setminus\Gamma}}{(x^*_{\Lambda\setminus\Gamma})^T x^*_{\Lambda\setminus\Gamma}} \sim (M^2)_{11}. \tag{4.24}$$

To obtain a lower bound in terms of the chi-squared distribution, note that

$$(M^2)_{11} = \sum_{i=1}^{r} M_{i1}^2 = M_{11}^2 + \sum_{i=2}^{r} M_{i1}^2 \geq M_{11}^2. \tag{4.25}$$

Meanwhile it follows from (4.22) and (4.6) that

$$M_{11} = \sum_{i=1}^{n-k} F_{i1}^2 \sim \frac{1}{n}\chi_{n-k}^2,$$

which combines with (4.24) and (4.25) to give (4.11).

Proof of (4.12): Since $A_{(\Gamma\cup\Lambda)^C} \in \mathcal{N}_{n,N-2k+r}(0,1/n)$ independently of $x^*_{(\Gamma\cup\Lambda)^C}$, we may apply Lemma 4.4 to deduce

$$\left\{A_{(\Gamma\cup\Lambda)^C} x^*_{(\Gamma\cup\Lambda)^C}\right\}_i \sim \mathcal{N}\left(0, \frac{\|(x^*_{\Gamma\cup\Lambda)^C}\|^2}{n}\right),$$

noting also that $\left\{A_{(\Gamma\cup\Lambda)^C} x^*_{(\Gamma\cup\Lambda)^C}\right\}_i$ is independent of $A_{\Gamma\cup\Lambda}$, which combines with Assumption 3 to give

$$\{\tilde{e}_\Gamma\}_i \sim \mathcal{N}\left(0, \frac{\tilde{\sigma}^2_\Gamma}{n}\right), \tag{4.26}$$

independently of $A_{\Gamma\cup\Lambda}$. By (4.15), we have

$$A^\dagger_\Gamma \tilde{e}_\Gamma = VD^{-1}U_1^T\tilde{e}_\Gamma = VD^{-1}p, \tag{4.27}$$

where $p := U_1^T\tilde{e}_\Gamma \in \mathbb{R}^{n-k}$. By (4.26), we may view $\tilde{e}_\Gamma$ as a one-column Gaussian matrix, such that $\tilde{e}_\Gamma \sim \mathcal{N}_{n,1}(0,\tilde{\sigma}^2_\Gamma/n)$, it follows from Lemma 4.3 that

$$p \sim \mathcal{N}_{k,1}(0,\tilde{\sigma}^2_\Gamma/n), \tag{4.28}$$

independently of U and therefore independently of A_Γ. Substituting (4.16) into (4.27) then gives

$$\|A^\dagger_\Gamma\tilde{e}_\Gamma\|^2 = \|VD^{-1}p\|^2 = \|D^{-1}p\|^2 = p^TD^{-2}p = p^TV^T(A_\Gamma^TA_\Gamma)^{-1}Vp = q^T(A_\Gamma^TA_\Gamma)^{-1}q, \tag{4.29}$$

where $q := Vp \in \mathbb{R}^k$. It now follows from (4.28) and Lemma 4.3 that $q \sim \mathcal{N}_{k,1}(0,\tilde{\sigma}^2_\Gamma/n)$, independently of V and therefore independently of A_Γ, and consequently that

$$q^Tq \sim \tilde{\sigma}^2_\Gamma\chi^2_k. \tag{4.30}$$

By (4.5) of Lemma 4.7,

$$\frac{q^Tq}{q^T(A_\Gamma^TA_\Gamma)^{-1}q} \sim \frac{1}{n}\chi^2_{n-k+1}. \tag{4.31}$$

Since q and A_Γ are independent, we may combine (4.29), (4.30) and (4.31) to give

$$\|A^\dagger_\Gamma\tilde{e}_\Gamma\| \sim \tilde{\sigma}_\Gamma\sqrt{G_\Gamma}, \quad \text{where} \quad G_\Gamma \sim \frac{k}{n-k+1}F(k,n-k+1), \tag{4.32}$$

and (3.20) now follows.

Proof of (4.13): Using (4.20), we have

$$A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A^\dagger_\Gamma)\tilde{e}_\Gamma = A^T_{\Lambda\setminus\Gamma}U_2U_2^T\tilde{e}_\Gamma = A^T_{\Lambda\setminus\Gamma}U_2f = B^Tf, \tag{4.33}$$

where $B := U_2^T A_{\Lambda\setminus\Gamma} \sim \mathcal{N}_{n-k,r}(0, 1/n)$ by Lemma 4.3, and where

$$f := U_2^T \tilde{e}_\Gamma \sim \mathcal{N}_{n-k,1}(0, \tilde{\sigma}_\Gamma^2/n) \tag{4.34}$$

by (4.26) and Lemma 4.3. Now let B have singular value decomposition

$$W[F \mid 0]Y^T = W_1 F Y^T, \tag{4.35}$$

where $F \in \mathbb{R}^{r\times r}$ is diagonal, and where $Y \in \mathbb{R}^{r\times r}$ and $W = [W_1 \mid W_2] \in \mathbb{R}^{(n-k)\times(n-k)}$ are orthogonal, noting that $W_1 \in \mathbb{R}^{(n-k)\times r}$. We have

$$g := W_1^T f \sim \mathcal{N}_{r,1}(0, \tilde{\sigma}_\Gamma^2/n) \tag{4.36}$$

by (4.34) and Lemma 4.3, and we may apply (4.33) to give

$$\begin{aligned} &\|A_{\Lambda\setminus\Gamma}^T(I - A_\Gamma A_\Gamma^\dagger)\tilde{e}_\Gamma\|^2 \le \|B^T f\|^2 = \|YFW_1^T f\|^2 \\ &= \|Fg\|^2 = g^T F^2 g = g^T Y^T (B^T B) Y g = h^T (B^T B) h, \end{aligned} \tag{4.37}$$

where $h := Yg \in \mathbb{R}^k$. Since $h \sim \mathcal{N}_{r,1}(0, \tilde{\sigma}_\Gamma^2/n)$ by (4.36) and Lemma 4.3, it follows that

$$h^T h \sim \tilde{\sigma}_\Gamma^2 \chi_r^2 \le \tilde{\sigma}_\Gamma^2 \chi_k^2, \tag{4.38}$$

since a χ_r^2 random variate may be viewed as a truncation of its extension to a χ_k^2 random variate. By (4.4) of Lemma 4.7,

$$\frac{h^T B^T B h}{h^T h} \sim \frac{1}{n}\chi_{n-k}^2. \tag{4.39}$$

Combining (4.37), (4.38) and (4.39) then proves (4.13). □

4.2 Proportional-dimensional asymptotic results for the regularized incomplete gamma and beta functions

In this section, we obtain proportional-dimensional asymptotic results for the distribution functions of the χ^2 and F distributions, namely the regularized incomplete gamma and beta functions.

4.2.1 Definitions and preliminary results

The gamma function, and its incomplete and regularized extensions, can be defined as integrals.

Definition 4.11 (**Gamma functions [1, Sections 6.1 and 6.5]**). *Given $p, q > 0$, define the*

gamma function as

$$\Gamma(p) := \int_0^\infty t^{p-1} e^{-t} \, \mathrm{d}t, \tag{4.40}$$

define the lower and upper incomplete gamma functions, respectively, as

$$\gamma(p,q) := \int_0^q t^{p-1} e^{-t} \, \mathrm{d}t \quad \textit{and} \quad \Gamma(p,q) := \int_q^\infty t^{p-1} e^{-t} \, \mathrm{d}t, \tag{4.41}$$

and define the lower and upper regularized incomplete gamma functions to be, respectively,

$$P(p,q) := \frac{\gamma(p,q)}{\Gamma(p)} \quad \textit{and} \quad Q(p,q) := \frac{\Gamma(p,q)}{\Gamma(p)}. \tag{4.42}$$

The gamma function can also be viewed as an extension of the factorial function to the positive real line. Unsurprisingly, then, the gamma function has asymptotic behaviour akin to Stirling's formula. The next result describes this asymptotic behaviour in the context of the logarithm of the gamma function.

Lemma 4.12 (**Log-gamma asymptotic [136, Theorem 12 and (4)]**). *For $p > 0$,*

$$\ln \Gamma(p) = \left(p - \frac{1}{2}\right) \ln p - p + \frac{1}{2} \ln(2\pi) + o(1), \tag{4.43}$$

where we adopt the usual convention of writing $f(p) = o(1)$ if $f(p) \to 0$ as $p \to \infty$.

For notational convenience, for $p > 0$, let us define

$$\Phi(p) := p \ln p - p - \ln \Gamma(p). \tag{4.44}$$

Using Lemma 4.12, we may deduce the following limiting result for $\Phi(p)$.

Lemma 4.13 (**Log-gamma limit**). *The following limit holds.*

$$\lim_{p \to \infty} \frac{1}{p} \Phi(p) = 0, \tag{4.45}$$

where $\Phi(p)$ is defined in (4.44).

Proof: Dividing (4.43) throughout by p gives

$$\frac{1}{p} \ln \Gamma(p) = \ln p - 1 + o(1),$$

and dividing (4.44) throughout by p then gives

$$\frac{1}{p} \Phi(p) = \ln p - 1 - [\ln p - 1 + o(1)] = o(1),$$

which proves the result. □

The beta function, and its incomplete and regularized variants, can also be defined as integrals.

Definition 4.14 (**Beta functions [1, Sections 6.2 and 6.6]**). *Given $p, q > 0$ and $x \in [0, 1]$, define the beta function to be*

$$B(p, q) := \int_0^1 t^{p-1}(1-t)^{q-1}\,\mathrm{d}t, \tag{4.46}$$

define the incomplete beta function to be

$$B_x(p, q) := \int_0^x t^{p-1}(1-t)^{q-1}\,\mathrm{d}t, \tag{4.47}$$

and define the regularized incomplete beta function to be

$$I_x(p, q) := \frac{B_x(p, q)}{B(p, q)}. \tag{4.48}$$

The next result shows that the standard gamma and beta functions defined in (4.40) and (4.46) are closely related.

Lemma 4.15 (**Beta-gamma identity [136, Theorem 7]**). *For any $p, q > 0$,*

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \tag{4.49}$$

Preliminaries addressed, we proceed to the main results of this section.

4.2.2 Asymptotics for the regularized incomplete gamma function

By analogy with the proportional-dimensional asymptotic for CS defined in (1.5), we consider the case in which the parameters of the regularized incomplete gamma function grow proportionally, with their ratio tending to some fixed limit. The functions $P(p, q)$ and $Q(p, q)$ tend to a spike at $q = p$ in the limit as $(p, q) \to \infty$, which indicates that exponential decay can only occur for $P(p, q)$ if $p > q$, and for $Q(p, q)$ if $q > p$. We therefore derive a separate result for both the lower and upper variants of the function. We first give the result for $P(p, q)$. In the following two results, we will consider the change of variables

$$u(t) := \sqrt{2\left[\frac{t}{p} - 1 - \ln\left(\frac{t}{p}\right)\right]}, \tag{4.50}$$

which is strictly decreasing for $0 < t < p$ and strictly increasing for $t > p$, and we will further define

$$u^*(p, q) := \sqrt{2\left[\frac{q}{p} - 1 - \ln\left(\frac{q}{p}\right)\right]}. \tag{4.51}$$

Theorem 4.16 (**Lower regularized incomplete gamma function**). *Let $p > q > 0$ be such that $\frac{q}{p} \to \gamma$ as $(p,q) \to \infty$ with $0 < \gamma < 1$. Then*

$$\lim_{p\to\infty} \frac{1}{p} \ln P(p,q) = -(\gamma - 1 - \ln \gamma), \tag{4.52}$$

where $P(p,q)$ is defined in (4.42).

Proof: By (4.41) and simple rearrangement, we have

$$\gamma(p,q) = \int_0^q t^{p-1} e^{-t} \, \mathrm{d}t = \left(\frac{p}{e}\right)^p \int_0^q \frac{1}{t} e^{-p\left[\frac{t}{p} - 1 - \ln\left(\frac{t}{p}\right)\right]} \, \mathrm{d}t,$$

to which we may apply the change of variables (4.50) to obtain

$$\gamma(p,q) = \left(\frac{p}{e}\right)^p \mathcal{I}_{p,q}, \tag{4.53}$$

where

$$\mathcal{I}_{p,q} := \int_{u^*(p,q)}^{\infty} u e^{-\frac{1}{2}pu^2} \cdot \frac{1}{\left(1 - \frac{t}{p}\right)} \, \mathrm{d}u, \tag{4.54}$$

and where $u^*(p,q)$ is defined in (4.51). Note that, since $q/p \to \gamma$,

$$\lim_{p\to\infty} u^*(p,q) = \sqrt{2(\gamma - 1 - \ln\gamma)}. \tag{4.55}$$

Since $t \le q$, we may upper bound (4.54) as

$$\mathcal{I}_{p,q} \le \frac{1}{\left(1 - \frac{q}{p}\right)} \int_{u^*(p,q)}^{\infty} u e^{-\frac{1}{2}pu^2} = \frac{1}{p\left(1 - \frac{q}{p}\right)} e^{-\frac{1}{2}p[u^*(p,q)]^2}. \tag{4.56}$$

On the other hand, since $t \le q$ and $1 - q/p \le 1$, we may lower bound (4.54) as

$$\mathcal{I}_{p,q} \ge \int_{u^*(p,q)}^{\infty} u e^{-\frac{1}{2}pu^2} \, \mathrm{d}u = \frac{1}{p} e^{-\frac{1}{2}p[u^*(p,q)]^2},$$

which, together with (4.56) and taking logarithms, gives

$$\frac{1}{p} \ln \left\{ \frac{1}{p} e^{-\frac{1}{2}p[u^*(p,q)]^2} \right\} \le \frac{1}{p} \ln \mathcal{I}_{p,q} \le \frac{1}{p} \ln \left\{ \frac{1}{p\left(1 - \frac{q}{p}\right)} e^{-\frac{1}{2}p[u^*(p,q)]^2} \right\}. \tag{4.57}$$

Using (4.55), we pass to the limit of (4.57) and deduce

$$\lim_{p\to\infty} \frac{1}{p} \ln \mathcal{I}_{p,q} = -\frac{1}{2} \lim_{p\to\infty} [u^*(p,q)]^2 = -(\gamma - 1 - \ln\gamma). \tag{4.58}$$

Combining (4.42) and (4.53), we have

$$P(p,q) = \frac{1}{\Gamma(p)} \left(\frac{p}{e}\right)^p \mathcal{I}_{p,q},$$

and taking logarithms we have

$$\frac{1}{p} \ln P(p,q) = \frac{1}{p}\Phi(p) + \frac{1}{p} \ln \mathcal{I}_{p,q}, \tag{4.59}$$

where $\Phi(p)$ is defined in (4.44), and the result now follows by using (4.58) and Lemma 4.12 to take limits of (4.59). □

The corresponding result for $Q(p,q)$ is given next.

Theorem 4.17 (**Upper regularized incomplete gamma function**). *Let $0 < p < q$ be such that $\frac{q}{p} \to \gamma$ as $(p,q) \to \infty$ with $\gamma > 1$. Then*

$$\lim_{p\to\infty} \frac{1}{p} \ln Q(p,q) = -(\gamma - 1 - \ln \gamma), \tag{4.60}$$

where $Q(p,q)$ is defined in (4.42).

Proof: By (4.41) and simple rearrangement, we have

$$\Gamma(p,q) = \int_q^\infty t^{p-1} e^{-t} \, \mathrm{d}t = \left(\frac{p}{e}\right)^p \int_q^\infty \frac{1}{t} e^{-p\left[\frac{t}{p} - 1 - \ln\left(\frac{t}{p}\right)\right]} \, \mathrm{d}t,$$

to which we may apply the change of variables (4.50) to obtain

$$\gamma(p,q) = \left(\frac{p}{e}\right)^p \mathcal{J}_{p,q}, \tag{4.61}$$

where

$$\mathcal{J}_{p,q} := \int_{u^*(p,q)}^\infty u e^{-\frac{1}{2}pu^2} \cdot \frac{1}{\left(\frac{t}{p} - 1\right)} \, \mathrm{d}u, \tag{4.62}$$

and where $u^*(p,q)$ is given by (4.51). Note that, since $q/p \to \gamma$,

$$\lim_{p\to\infty} u^*(p,q) = \sqrt{2(\gamma - 1 - \ln\gamma)}. \tag{4.63}$$

Since $t \geq q$, we may upper bound (4.62) as

$$\mathcal{J}_{p,q} \leq \frac{1}{\left(\frac{q}{p} - 1\right)} \int_{u^*(p,q)}^\infty u e^{-\frac{1}{2}pu^2} = \frac{1}{p\left(\frac{q}{p} - 1\right)} e^{-\frac{1}{2}p[u^*(p,q)]^2}. \tag{4.64}$$

In order to lower bound $\mathcal{J}_{p,q}$, we need a more involved argument. Let us first show that

$$\frac{t}{p} \leq u^2 + 4. \tag{4.65}$$

First note that (4.65) holds trivially if $t/p < 4$. Supposing $t/p \geq 4$, basic calculus then gives $t/p - 1 \geq 2\ln(t/p)$, which upon substitution into (4.50) gives $t/p \leq u^2 + 1$, which thus proves (4.65) in general. We may use (4.65) to lower bound (4.62) as

$$\begin{aligned}
\mathcal{J}_{p,q} &\geq \int_{u^*(p,q)}^{\infty} \frac{u e^{-\frac{1}{2}pu^2}}{u^2+3}\,\mathrm{d}u \\
&= \left[\frac{-1}{p(u^2+3)} e^{-\frac{1}{2}pu^2}\right]_{u^*(p,q)}^{\infty} - \int_{u^*(p,q)}^{\infty} \frac{2u}{p(u^2+3)^2} e^{-\frac{1}{2}pu^2}\,\mathrm{d}u \\
&\geq \frac{e^{-\frac{1}{2}p[u^*(p,q)]^2}}{p\{[u^*(p,q)]^2+3\}} - \frac{2}{3p}\mathcal{J}_{p,q},
\end{aligned} \tag{4.66}$$

where in the last step we use $u^2 + 3 \geq 3$ and $\frac{t}{p} - 1 \leq u^2 + 3$. Rearranging (4.66) yields

$$\mathcal{J}_{p,q} \geq \frac{1}{\left(p+\frac{2}{3}\right)\{[u^*(p,q)]^2+3\}} e^{-\frac{1}{2}p[u^*(p,q)]^2},$$

which, together with (4.64) and taking logarithms, gives

$$\frac{1}{p}\ln\left\{\frac{1}{\left(p+\frac{2}{3}\right)\{[u^*(p,q)]^2+3\}} e^{-\frac{1}{2}p[u^*(p,q)]^2}\right\} \leq \frac{1}{p}\ln\mathcal{J}_{p,q} \leq \frac{1}{p}\ln\left\{\frac{1}{p\left(\frac{q}{p}-1\right)} e^{-\frac{1}{2}p[u^*(p,q)]^2}\right\}. \tag{4.67}$$

We now use (4.63) to pass to the limit of (4.67), deducing

$$\lim_{p\to\infty}\frac{1}{p}\ln\mathcal{J}_{p,q} = -\frac{1}{2}\lim_{p\to\infty}[u^*(p,q)]^2 = -(\gamma - 1 - \ln\gamma). \tag{4.68}$$

Combining (4.61) and (4.42), we have

$$Q(p,q) = \frac{1}{\Gamma(p)}\left(\frac{p}{e}\right)^p \mathcal{J}_{p,q},$$

and taking logarithms we have

$$\frac{1}{p}\ln Q(p,q) = \frac{1}{p}\Phi(p) + \frac{1}{p}\ln\mathcal{J}_{p,q}, \tag{4.69}$$

where $\Phi(p)$ is defined in (4.44), and the result now follows by using (4.68) and Lemma 4.12 to take limits of (4.69). □

4.2.3 Asymptotics for the regularized incomplete beta function

Next, we obtain a proportional-dimensional asymptotic result for the regularized incomplete beta function $I_x(p,q)$. As $(p,q) \to \infty$, $I_x(p,q)$ tends to a spike at $x = p/(p+q)$, from which we see that exponential decay can only occur if we are in the left tail of the function, *i.e.* $x < p/(p+q)$. We have the following result.

Theorem 4.18 (**Regularized incomplete beta function**). *Fix* $0 < x < 1$ *and let* $p, q > 0$ *satisfy*

$$\frac{p}{p+q} > x. \tag{4.70}$$

Let $(p,q) \to \infty$ *such that*

$$\lim_{p\to\infty} \frac{p}{p+q} = \beta, \tag{4.71}$$

where $\beta > x$. *Then*

$$\lim_{p\to\infty} \frac{1}{p+q} \ln I_x(p,q) = -\left[\beta \ln\left(\frac{\beta}{x}\right) + (1-\beta)\ln\left(\frac{1-\beta}{1-x}\right)\right]. \tag{4.72}$$

Proof: By (4.47) and simple rearrangement, we have

$$B_x(p,q) = \int_0^x t^{p-1}(1-t)^{q-1}\,\mathrm{d}t = \frac{p^p q^q}{(p+q)^{p+q}} \int_0^x \frac{1}{t(1-t)} e^{-\left\{p\ln\left[\frac{p}{t(p+q)}\right] + q\ln\left[\frac{q}{(1-t)(p+q)}\right]\right\}}\,\mathrm{d}t,$$

to which we may apply the change of variables

$$u = \sqrt{\frac{2}{p+q}\left\{p\ln\left[\frac{p}{t(p+q)}\right] + q\ln\left[\frac{q}{(1-t)(p+q)}\right]\right\}}, \tag{4.73}$$

and obtain

$$B_x(p,q) = \frac{p^p q^q}{(p+q)^{p+q}} \mathcal{K}_{p,q}, \tag{4.74}$$

where

$$\mathcal{K}_{p,q} := \int_{u^*(p,q)}^{\infty} u e^{-\frac{1}{2}(p+q)u^2} \cdot \frac{p+q}{p - t(p+q)}\,\mathrm{d}u \tag{4.75}$$

and

$$u^*(p,q) := \sqrt{\frac{2}{p+q}\left\{p\ln\left[\frac{p}{x(p+q)}\right] + q\ln\left[\frac{q}{(1-x)(p+q)}\right]\right\}}. \tag{4.76}$$

Note that, by (4.71),

$$\lim_{p\to\infty} u^*(p,q) = \sqrt{2\left[\beta\ln\left(\frac{\beta}{x}\right) + (1-\beta)\ln\left(\frac{1-\beta}{1-x}\right)\right]}. \tag{4.77}$$

Since $t \le x$, we may upper bound (4.75) as

$$\mathcal{K}_{p,q} \le \frac{p+q}{p - x(p+q)} \int_{u^*(p,q)}^{\infty} u e^{-\frac{1}{2}(p+q)u^2}\,\mathrm{d}u = \frac{1}{p - x(p+q)} e^{-\frac{1}{2}(p+q)[u^*(p,q)]^2}. \tag{4.78}$$

On the other hand, since $p - t(p+q) \leq p+q$, we may lower bound (4.75) as

$$\mathcal{K}_{p,q} \geq \int_{u^*(p,q)}^{\infty} u e^{-\frac{1}{2}(p+q)u^2} = \frac{1}{p+q} e^{-\frac{1}{2}(p+q)[u^*(p,q)]^2},$$

which, on combination with (4.78) and taking logarithms, gives

$$\frac{1}{p+q} \ln \left\{ \frac{1}{p+q} e^{-\frac{1}{2}(p+q)[u^*(p,q)]^2} \right\} \leq \frac{1}{p+q} \ln \mathcal{K}_{p,q} \leq \frac{1}{p+q} \ln \left\{ \frac{1}{p - x(p+q)} e^{-\frac{1}{2}(p+q)[u^*(p,q)]^2} \right\}. \tag{4.79}$$

We may now use (4.71) and (4.77) to take limits of (4.79), deducing

$$\lim_{p \to \infty} \frac{1}{p+q} \ln \mathcal{K}_{p,q} = -\frac{1}{2} \cdot \lim_{p \to \infty} [u^*(p,q)]^2 = -\left[\beta \ln \left(\frac{\beta}{x} \right) + (1-\beta) \ln \left(\frac{1-\beta}{1-x} \right) \right]. \tag{4.80}$$

Combining (4.74) and (4.48), we have

$$I_x(p,q) = \frac{1}{B(p,q)} \cdot \frac{p^p q^q}{(p+q)^{p+q}} \cdot \mathcal{K}_{p,q} = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \cdot \frac{p^p q^q}{(p+q)^{p+q}} \cdot \mathcal{K}_{p,q},$$

where the last step follows by Lemma 4.15. Taking logarithms, we have

$$\frac{1}{p+q} \ln I_x(p,q) = \frac{1}{p+q}\Phi(p) + \frac{1}{p+q}\Phi(q) - \frac{1}{p+q}\Phi(p+q) + \frac{1}{p+q} \ln \mathcal{K}_{p,q}, \tag{4.81}$$

where $\Phi(p)$ is defined in (4.44), and the result now follows by using (4.80) and Lemma 4.12 to take limits of (4.81). □

4.3 Large deviations for matrix-vector independence

4.3.1 Results for the χ^2 distribution

The next lemma gives a standard result concerning the distribution function of the χ^2-distribution in terms of the regularized incomplete gamma function.

Lemma 4.19 (χ^2 **distribution function [1, 26.4.19]**). *Let $X \sim \chi^2_s$. Then*

$$\mathbb{P}(X \leq x) = P\left(\frac{s}{2}, \frac{x}{2}\right) \quad \text{and} \quad \mathbb{P}(X \geq x) = Q\left(\frac{s}{2}, \frac{x}{2}\right), \tag{4.82}$$

where $P(p,q)$ and $Q(p,q)$ are defined in (4.42).

By drawing upon the results in Theorems 4.16 and 4.17, we obtain the following asymptotic tail bound result for the normalized χ^2 distribution.

Lemma 4.20. *Let $0 < l \leq n$ and let the random variable $X_l \sim \frac{1}{l}\chi_l^2$. Let $l/n \to \lambda \in (0,1]$ as $n \to \infty$. Then, for any $\nu > 0$,*

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(X_l \geq 1+\nu) = -\frac{\lambda}{2}[\nu - \ln(1+\nu)] \tag{4.83}$$

and, for any $\nu \in (0,1)$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(X_l \leq 1-\nu) = -\frac{\lambda}{2}[-\nu - \ln(1-\nu)]. \tag{4.84}$$

Proof: We first show (4.83). We have

$$\mathbb{P}(X_l \geq 1+\nu) = \mathbb{P}[\chi_l^2 \geq l(1+\nu)] = Q\left[\frac{l}{2}, \frac{l(1+\nu)}{2}\right], \tag{4.85}$$

where the first step follows from the definition of X_l, and the second step follows from (4.82). Setting $p = l/2$ and $q = l(1+\nu)/2$ in Theorem 4.17, we have

$$\lim_{p\to\infty} \frac{q}{p} = 1 + \nu > 1,$$

and therefore we may apply Theorem 4.17 with these choices for p and q, obtaining

$$\lim_{n\to\infty} \frac{1}{n} \ln Q\left[\frac{l}{2}, \frac{l(1+\nu)}{2}\right] = \frac{\lambda}{2} \cdot \lim_{n\to\infty} \frac{2}{l} \ln Q\left[\frac{l}{2}, \frac{l(1+\nu)}{2}\right] = -\frac{\lambda}{2}[\nu - \ln(1+\nu)],$$

which combines with (4.85) to yield (4.83). To show (4.84), we can note that

$$\mathbb{P}(X_l \leq 1-\nu) = \mathbb{P}[\chi_l^2 \leq l(1-\nu)] = P\left[\frac{l}{2}, \frac{l(1-\nu)}{2}\right], \tag{4.86}$$

where the first step follows from the definition of X_l, and the second step follows from (4.82). Setting $p = l/2$ and $q = l(1-\nu)/2$ in Theorem 4.16, we have

$$\lim_{p\to\infty} \frac{q}{p} = 1 - \nu < 1,$$

and therefore we may apply Theorem 4.16 with these choices for p and q, obtaining

$$\lim_{n\to\infty} \frac{1}{n} \ln P\left[\frac{l}{2}, \frac{l(1-\nu)}{2}\right] = \frac{\lambda}{2} \cdot \lim_{n\to\infty} \frac{2}{l} \ln P\left[\frac{l}{2}, \frac{l(1-\nu)}{2}\right] = -\frac{\lambda}{2}[-\nu - \ln(1-\nu)],$$

which combines with (4.86) to yield (4.84). □

We now derive tail bounds which hold with overwhelming probability asymptotically for a number of χ^2 distributions which grows exponentially in n. We introduce a further definition.

Definition 4.21. *In the proportional-growth asymptotic, given a sequence of index sets* $\{S_n\}$*, let*

$$\mathcal{S}(\delta,\rho) := \lim_{n\to\infty} \frac{1}{n} \ln |S_n|, \tag{4.87}$$

provided the limit is well-defined.

As an example, one choice of $\{S_n\}$ which will be of particular interest to us is to let S_n be, for each (k,n,N), the set of all possible support sets of cardinality k, so that $|S_n| = \binom{N}{k}$. In this case, it is straightforward to show that $\mathcal{S}(\delta,\rho) = H(\delta\rho)/\delta$, where $H(\cdot)$ is defined in (2.34), see Lemma 5.1 for further elucidation. However, in the tree-based model, the number of permissible support sets is much reduced, and so we will also consider other choices for $\{S_n\}$ and $\mathcal{S}(\delta,\rho)$ in this thesis.

We define the following tail bound functions.

Definition 4.22 (χ^2 **tail bounds**). *Let* $\delta \in (0,1]$*,* $\rho \in (0,1)$ *and* $\lambda \in (0,1]$*. Let* $\mathcal{IU}_{\mathcal{S}}(\delta,\rho,\lambda)$ *be the unique solution to*

$$\nu - \ln(1+\nu) = \frac{2\mathcal{S}(\delta,\rho)}{\lambda} \quad \textit{for} \quad \nu > 0, \tag{4.88}$$

and let $\mathcal{IL}_{\mathcal{S}}(\delta,\rho,\lambda)$ *be the unique solution to*

$$-\nu - \ln(1-\nu) = \frac{2\mathcal{S}(\delta,\rho)}{\lambda} \quad \textit{for} \quad \nu \in (0,1), \tag{4.89}$$

where $\mathcal{S}(\delta,\rho)$ *is defined in (4.87).*

That $\mathcal{IU}$ is well-defined follows since the left-hand side of (4.88) is zero at $\nu = 0$, tends to infinity as $\nu \to \infty$, and is strictly increasing on $\nu > 0$. Similarly, $\mathcal{IL}$ is well-defined since the left-hand side of (4.89) is zero at $\nu = 0$, tends to infinity as $\nu \to 1$, and is strictly increasing on $\nu \in (0,1)$. Our main result for the χ^2 distribution follows.

Lemma 4.23 (**Large deviations result for** χ^2). *Let* $l \in \{1,\dots,n\}$ *and let the random variables* $X_l^i \sim \frac{1}{l}\chi_l^2$ *for all* $i \in S_n$*, and let* $\epsilon > 0$*. In the proportional growth asymptotic, let* $l/n \to \lambda \in (0,1]$*. Then*

$$\mathbb{P}\left\{\cup_{i\in S_n}[X_l^i \geq 1 + \mathcal{IU}_{\mathcal{S}}(\delta,\rho,\lambda) + \epsilon]\right\} \to 0 \tag{4.90}$$

and

$$\mathbb{P}\left\{\cup_{i\in S_n}[X_l^i \leq 1 - \mathcal{IL}_{\mathcal{S}}(\delta,\rho,\lambda) - \epsilon]\right\} \to 0, \tag{4.91}$$

exponentially in n*, where* $\mathcal{IU}_{\mathcal{S}}(\delta,\rho,\lambda)$ *and* $\mathcal{IL}_{\mathcal{S}}(\delta,\rho,\lambda)$ *are defined in (4.88) and (4.89) respectively.*

Proof: Union bounding $\mathbb{P}\left(X_l^i \geq 1+\nu\right)$ over all $i \in S_n$ gives

$$\mathbb{P}\left\{\cup_{i\in S_n}(X_l^i \geq 1+\nu)\right\} \leq \sum_{i\in S_n} \mathbb{P}\left(X_l^i \geq 1+\nu\right) = |S_n| \cdot \mathbb{P}(X_l^1 \geq 1+\nu). \tag{4.92}$$

Taking logarithms and limits of the right-hand side of (4.92), using (4.83) and (4.87), we have

$$\lim_{n\to\infty} \frac{1}{n} \ln\left[|S_n| \cdot \mathbb{P}(X_l^1 \geq 1+\nu)\right] = \mathcal{S}(\delta,\rho) - \frac{\lambda}{2}[\nu - \ln(1+\nu)],$$

and so (4.92) implies that, for any $\eta > 0$,

$$\frac{1}{n} \ln \mathbb{P}\left\{\cup_{i\in S_n}(X_l^i \geq 1+\nu)\right\} \leq \mathcal{S}(\delta,\rho) - \frac{\lambda}{2}[\nu - \ln(1+\nu)] + \eta, \tag{4.93}$$

for all n sufficiently large. By the definition of $\mathcal{IU}_{\mathcal{S}}(\delta,\rho,\lambda)$ in (4.88), and since $[\nu - \ln(1+\nu)]$ is strictly increasing on $\nu > 0$, then, for any $\epsilon > 0$, setting $\nu := \nu^* = \mathcal{IU}_{\mathcal{S}}(\delta,\rho,\lambda) + \epsilon$ and choosing η sufficiently small in (4.93) ensures

$$\frac{1}{n} \ln \mathbb{P}\left\{\cup_{i\in S_n}(X_l^i \geq 1+\nu^*)\right\} \leq -c_Q \quad \text{for all } n \text{ sufficiently large,}$$

where c_Q is some positive constant, from which it follows that

$$\mathbb{P}\left\{\cup_{i\in S_n}(X_l^i \geq 1+\nu^*)\right\} \leq e^{-c_Q \cdot n} \quad \text{for all } n \text{ sufficiently large,}$$

and (4.90) follows. Combining the same union bound argument with the lower tail result of Lemma 4.20 shows that, if we take $\nu^* = \mathcal{IL}_{\mathcal{S}}(\delta,\rho,\lambda) + \epsilon$ for some $\epsilon > 0$, then

$$\frac{1}{n} \ln \mathbb{P}\left\{\cup_{i\in S_n}(X_l^i \leq 1-\nu^*)\right\} \leq -c_P \quad \text{for all } n \text{ sufficiently large,}$$

where c_P is some positive constant, and (4.91) follows similarly to (4.90). □

'Independent RIP'. We previously noted that the particular choice of $\mathcal{S}(\delta,\rho) := H(\delta\rho)/\delta$, where $H(\cdot)$ is defined in (2.34), corresponds to counting over all possible support sets of cardinality k. Making this choice and setting $\lambda := 1$ in $\mathcal{IU}_{\mathcal{S}}(\delta,\rho,\lambda)$ and $\mathcal{IL}_{\mathcal{S}}(\delta,\rho,\lambda)$, an interesting comparison can be made with the RIP bounds used in Chapter 2 [4]. Suppose $B \sim \mathcal{N}_{n,N}(0,1/n)$ and let Γ be an index set of cardinality k. Then, by Lemma 4.7, given a vector $y \in \mathbb{R}^k$ independent of B,

$$\frac{\|B_\Gamma y\|^2}{\|y\|^2} \sim \frac{1}{n}\chi_n^2, \tag{4.94}$$

which is precisely the Rayleigh quotient in the definition of the RIP in (1.6). The only difference now is that we are assuming in addition that B and y are independent. Since we are bounding over all $\binom{N}{k}$ possible support sets, it follows that $\mathcal{IU}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1)$ and $\mathcal{IL}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1)$ may be

viewed as upper bounds on 'independent RIP' constants for Gaussian matrices.

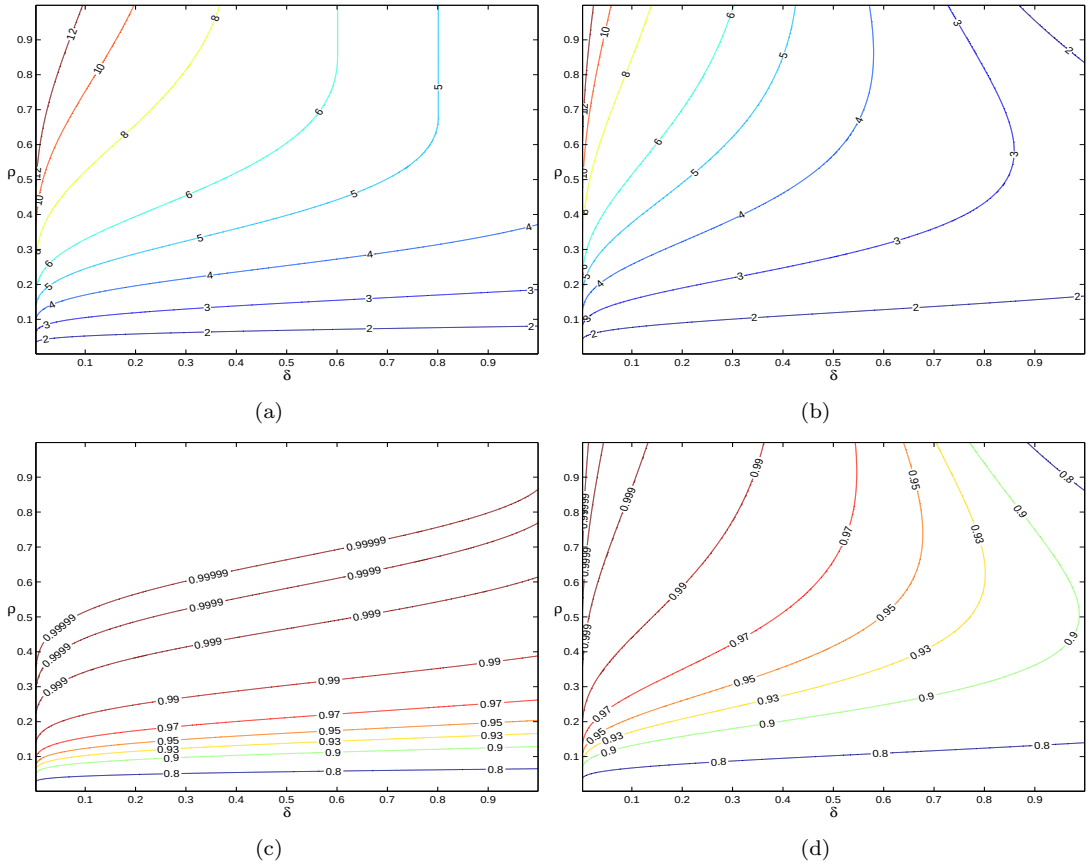


Figure 4.1: A comparison of standard RIP bounds and 'independent RIP' bounds: (a) $\mathcal{U}(\delta,\rho)$ (b) $\mathcal{IU}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1)$ (c) $\mathcal{L}(\delta,\rho)$ (d) $\mathcal{IL}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1)$.

Figure 4.1 gives plots of the 'independent RIP' bounds for Gaussian matrices $\mathcal{IU}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1)$ and $\mathcal{IL}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1)$ derived here, along with plots of the standard RIP bounds for Gaussian matrices obtained in [4]. One observes empirically the inequalities

$$\mathcal{IU}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1) < \mathcal{U}(\delta,\rho) \quad \text{and} \quad \mathcal{IL}_{\frac{H(\delta\rho)}{\delta}}(\delta,\rho,1) < \mathcal{L}(\delta,\rho).$$

A simple interpretation is that the additional information that the matrix and vector are independent allows us to tighten the bounds on (4.94). This consideration accounts for a large part of the quantitative improvement that is obtained in this thesis over existing recovery results for IHT algorithms which rely solely upon the RIP. Of course, such an analysis is only possible if matrix-vector independence can be used, which is the case for the stable point condition (3.16).

4.3.2 Results for the F distribution

The distribution function for the F distribution may be expressed in terms of the regularized incomplete beta function.

Lemma 4.24 (F **complementary distribution function [1, 26.6.2]**). *Let* $X \sim F(s,t)$. *Then*

$$\mathbb{P}(X \geq x) = I_{(\frac{s}{s+tx})}\left(\frac{t}{2}, \frac{s}{2}\right). \tag{4.95}$$

Having established the link with Section 4.2, we may derive the following proportional-dimensional asymptotic result for the F distribution.

Lemma 4.25. *Let the random variable* $X_n \sim \dfrac{k}{n-k+1}\,\mathcal{F}(k, n-k+1)$. *Provided*

$$f > \frac{\rho}{1-\rho}, \tag{4.96}$$

in the proportional-growth asymptotic,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(X_n \geq f) = -\frac{1}{2}\left[\ln(1+f) - \rho \ln f - H(\rho)\right], \tag{4.97}$$

where $H(\cdot)$ *is defined in (2.34).*

Proof: We have

$$\mathbb{P}(X_n \geq f) = \mathbb{P}\left[\mathcal{F}(k, n-k+1) \geq \frac{(n-k+1)f}{k}\right] = I_{\frac{1}{1+f}}\left(\frac{n-k+1}{2}, \frac{k}{2}\right), \tag{4.98}$$

where the first step follows from the definition of X_n, and the second step follows from (4.95). Setting $p = (n-k+1)/2$, $q = k/2$ and $x = 1/(1+f)$ in Theorem 4.18, we have $\beta = 1-\rho$, and we may use (4.96) to deduce

$$x = \frac{1}{1+f} < \frac{1}{1+\frac{\rho}{1-\rho}} = 1-\rho = \beta. \tag{4.99}$$

Meanwhile, since $p/(p+q) = (n-k+1)/(n+1) \to 1-\rho$, it also follows from (4.99) that there exists some $\tilde{p}$ such that $x < p/(p+q)$ for all $p \geq \tilde{p}$. Hence, using these choices of p, q and x, and ignoring small (p,q) for which $p < \tilde{p}$, we may apply Theorem 4.18 to obtain

$$\lim_{n\to\infty} \frac{2}{n+1} \ln I_{\frac{1}{1+f}}\left(\frac{n-k+1}{2}, \frac{k}{2}\right) = -\left\{(1-\rho)\ln\left[(1-\rho)(1+f)\right] + \rho \ln\left[\frac{\rho(1+f)}{f}\right]\right\},$$

from which it follows that

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \ln I_{\frac{1}{1+f}}\left(\frac{n-k+1}{2}, \frac{k}{2}\right) &= -\frac{1}{2}\left\{(1-\rho)\ln\left[(1-\rho)(1+f)\right] + \rho \ln\left[\frac{\rho(1+f)}{f}\right]\right\} \\ &= -\frac{1}{2}\left[\ln(1+f) - \rho\ln f - H(\rho)\right],\end{aligned}$$

where in the last line we used (2.34). This, together with (4.98), yields (4.97). □

We define the following tail bound function.

Definition 4.26 (*F* **tail bound**). *Let* $\delta \in (0,1]$ *and* $\rho \in (0,1/2]$. *Let* $\mathcal{IF}_{\mathcal{S}}(\delta,\rho)$ *be the unique solution in* f *to*

$$\ln(1+f) - \rho \ln f = 2\mathcal{S}(\delta,\rho) + H(\rho) \quad \textit{for} \quad f > \frac{\rho}{1-\rho}, \tag{4.100}$$

where $\mathcal{S}(\delta,\rho)$ *is defined in (4.87) and* $H(\cdot)$ *is defined in (2.34).*

That $\mathcal{IF}$ is well-defined follows since the left-hand side of (4.100) is equal to $H(\rho)$ at $f = \rho/(1-\rho)$, tends to infinity as $f \to \infty$, and is strictly increasing on $f > \rho/(1-\rho)$. Our main large deviation result, giving asymptotic tail bounds for a number of F distributed random variables which grows exponetially in n, is given next.

Lemma 4.27 (**Large deviations result for** F). *Let the random variables* $X_n^i \sim \frac{k}{n-k+1}\,\mathcal{F}(k, n-k+1)$ *for all* $i \in S_n$, *and let* $\epsilon > 0$. *In the proportional growth asymptotic,*

$$\mathbb{P}\left\{\cup_{i\in S_n}[X_n^i \geq \mathcal{IF}_{\mathcal{S}}(\delta,\rho) + \epsilon]\right\} \to 0, \tag{4.101}$$

exponentially in n, *where* $\mathcal{IF}_{\mathcal{S}}(\delta,\rho)$ *is defined in (4.100).*

Proof: Union bounding $\mathbb{P}(X_n^i \geq 1+f)$ over all $i \in S_n$ gives

$$\mathbb{P}\left\{\bigcup_{i\in S_n}(X_n^i \geq f)\right\} \leq \sum_{i\in S_n}\mathbb{P}\left(X_n^i \geq f\right) = |S_n| \cdot \mathbb{P}(X_n^1 \geq f), \tag{4.102}$$

Taking logarithms and limits of the right-hand side of (4.102), using (4.97) and (4.87), we have

$$\lim_{n\to\infty}\frac{1}{n}\ln\left[|S_n| \cdot \mathbb{P}(X_n^1 \geq f)\right] = \mathcal{S}(\delta,\rho) - \frac{1}{2}\left[\ln(1+f) - \rho\ln f - H(\rho)\right],$$

which combines with (4.102) to imply that, for any $\eta > 0$,

$$\frac{1}{n}\ln\mathbb{P}\left\{\cup_{i\in S_n}(X_n^i \geq f)\right\} \leq \mathcal{S}(\delta,\rho) - \frac{1}{2}\left[\ln(1+f) - \rho\ln f - H(\rho)\right] + \eta, \tag{4.103}$$

for all n sufficiently large. By the definition of $\mathcal{IF}_{\mathcal{S}}(\delta,\rho)$ in (4.100), and since the left-hand side of (4.100) on $f > \dfrac{\rho}{1-\rho}$ is strictly increasing in f, then, for any $\epsilon > 0$, setting $f := f^* = \mathcal{IF}_{\mathcal{S}}(\delta,\rho) + \epsilon$ and choosing η sufficiently small in (4.103) ensures

$$\frac{1}{n}\ln\mathbb{P}\left\{\cup_{i\in S_n}(X_n^i \geq f^*)\right\} \leq -c_I \quad \text{for all } n \text{ sufficiently large,}$$

where c_I is some positive constant, from which it follows that

$$\mathbb{P}\left\{\cup_{i\in S_n}(X_n^i \geq f^*)\right\} \leq e^{-c_I \cdot n} \quad \text{for all } n \text{ sufficiently large,}$$

and (4.101) now follows. □

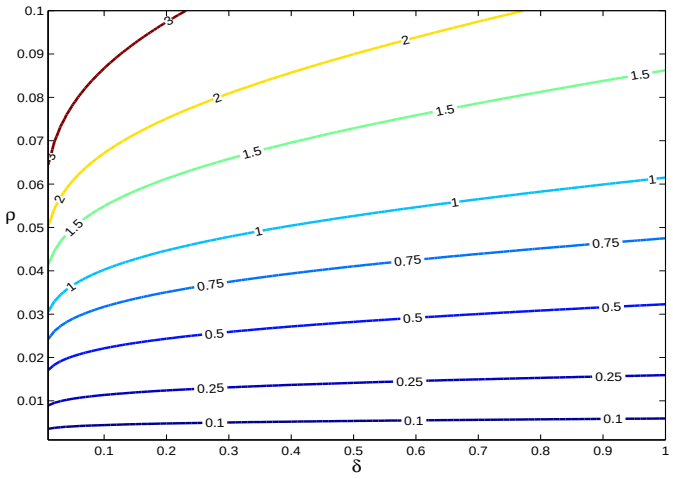


Figure 4.2: A plot of $\mathcal{IF}(\delta, \rho)$ for $\rho \in (0, 0.1)$.

In Chapter 5, we will require bounds which hold over $S_n := \binom{N}{k}$ support sets, which gives $\mathcal{S}(\delta, \rho) = H(\delta\rho)/\delta$. As an illustration, a plot of the tail bound function $\mathcal{IF}_{\mathcal{S}}(\delta, \rho)$ is given in Figure 4.2 for this choice of $\mathcal{S}(\delta, \rho)$. Note that $\mathcal{IF}_{\mathcal{S}}(\delta, \rho)$ grows rapidly for larger values of ρ.

Chapter 5

Phase transitions for Gaussian matrices

5.1 Introduction and motivation for average case analysis

In Chapter 3, we presented a new recovery analysis which considers the fixed or stable points of IHT algorithms. By using the RIP, we obtained worst-case recovery conditions which guarantee that the given measurement matrix allows recovery of any k-compressible signal from noisy measurements. We begin the present chapter by quantifying these new RIP conditions in the phase transition framework for Gaussian matrices, to enable a comparison with the current state-of-the-art RIP results presented in Chapter 2.

Figure 5.1 compares the recovery phase transition resulting from the RIP-based stable point analysis in Section 3.3 with the phase transitions derived in Chapter 2 based upon Foucart's RIP analysis [93]. The plots have precisely the same interpretation as in Chapter 2: in the case of zero measurement noise, recovery of all k-sparse vectors is asymptotically guaranteed in the proportional-dimensional framework for a (δ, ρ) pair falling below the curve. For IHT, the stable point phase transition holds over a range of stepsizes which shrinks in size as the phase transition is approached, while the optimum phase transition based upon Foucart's analysis holds only for a single stepsize choice, deteriorating for other stepsize choices. For NIHT, a shrinkage parameter of $\kappa := 1.1$ is selected in both cases.

We observe that the stable point approach yields a lower phase transition for IHT, and therefore gives no improvement upon the current best result. On the other hand, we do observe an improvement in the phase transition by roughly a factor of 2 in the case of NIHT, which therefore gives the highest worst-case phase transition for NIHT yet derived. However, in both cases, the observation of Sections 1.5 and 2.4 still applies, namely that these worst-case results are pessimistic compared to the empirical performance of IHT algorithms. We argued in Chapter 1 that it would be expected that IHT algorithms, like l_1-minimization, would exhibit two different phase transitions: a strong phase transition capturing worst-case behaviour (all k-sparse signals), and a weak phase transition capturing average-case behaviour (most k-sparse

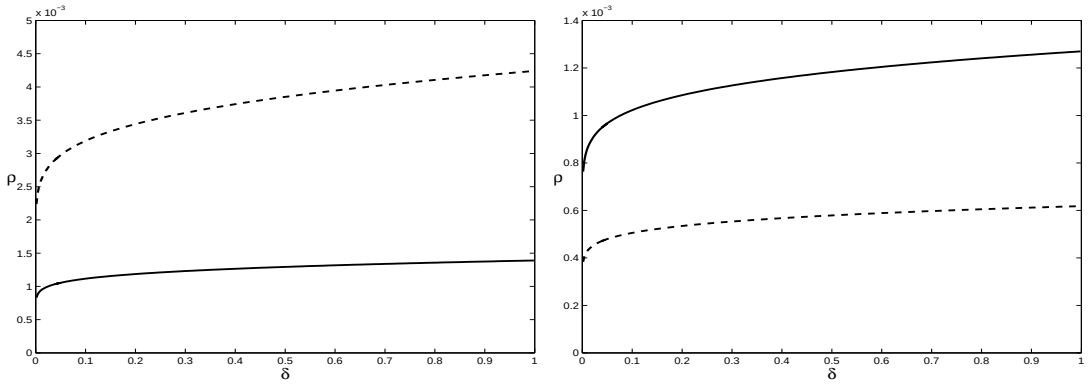


Figure 5.1: Phase transitions from the stable point RIP analysis of Chapter 3 (unbroken), along with current best RIP-based phase transitions from Chapter 2 [93] (dashed); (a) IHT (b) NIHT.

signals). Furthermore, it was argued that it is the weak phase transition that one would expect to observe in practice. It is natural, therefore, to expect that improved recovery guarantees could be obtained by relaxing worst-case assumptions and switching to average-case analysis.

Average-case analysis of l_0-based algorithms is generally considered to be challenging, or even impossible, due to the difficult probabilistic dependencies that are repeatedly introduced at each iteration of an algorithm such as IHT. However, our stable point analysis in Chapter 3 departs, at least in part, from the usual approach of tracking the algorithm from one iteration to the next, allowing necessary conditions for the existence of a stable point to be expressed independently of the iterates of the algorithm, thereby removing the offending dependency. Furthermore, we saw in Chapter 4 that the stable point condition is particularly amenable to analysis in the case of Gaussian matrices: precise distributions of its constituent terms can be derived, and large deviations results can be obtained to derive bounds on these terms over all permissible support sets in the proportional-dimensional asymptotic. Underlying this approach is the average-case assumption that the original signal, measurement matrix and measurement noise are all independent of each other, a realistic assumption in CS.

In this chapter, we combine the deterministic stable point recovery analysis of Sections 3.1 and 3.2 with the distribution results and large deviations results of Chapter 4. In so doing, we obtain average-case recovery phase transitions for both IHT and NIHT in the context of Gaussian matrices, thereby giving lower bounds on a particular kind of weak phase transition for IHT algorithms. We prove our results in the context of Problem 1.3, namely in the most general case of k-compressible signals and noisy measurements. As explained in Chapter 3, this direct approach means we avoid the need to revert to the RIP and Lemma 2.2 in order to extend results for k-sparse recovery to results for k-compressible recovery. In fact, we must take a different approach out of necessity, since the independence assumption we will impose upon the measurement noise means that Lemma 2.2 will not apply in our case. While results for

inexactly sparse signals and noisy measurements are more likely to be practically relevant, the most mathematical insight is gained by considering the special case of k-sparse signals and zero measurement noise (Problem 1.1). For this reason, we will also give corollaries of our general results in the context of Problem 1.1.

The outline of the chapter is as follows. In Section 5.2, we define our recovery phase transitions and establish that they are well-defined. We prove our recovery results in Section 5.3, first for IHT and then for NIHT. We illustrate our results in Section 5.4, and present a discussion, firstly for the noiseless case of Problem 1.1, and secondly for the general case of Problem 1.3.

5.2 Definitions of recovery phase transitions

We will use the tail bound functions defined in (4.88), (4.89) and 4.100 to perform union bounds over all possible support sets of cardinality k, of which there are $\binom{N}{k}$. Setting $|S_n| := \binom{N}{k}$, we have the following expression for $\mathcal{S}(\delta, \rho)$.

Lemma 5.1 (**Combinatorial limit**). *Let the sequence of sets S_n be such that $|S_n| = \binom{N}{k}$. Let $\mathcal{S}(\delta, \rho)$ be defined in terms of S_n as in (4.87). Then*

$$\mathcal{S}(\delta, \rho) = \frac{H(\delta\rho)}{\delta}, \tag{5.1}$$

where $H(\cdot)$ is defined in (2.34).

Proof: In the proportional-dimensional asymptotic,

$$\mathcal{S}(\delta, \rho) = \lim_{n\to\infty} \frac{1}{n} \ln |S_n| = \lim_{n\to\infty} \frac{N}{n} \cdot \frac{1}{N} \ln \binom{N}{k} = \frac{1}{\delta} \cdot H(\delta\rho),$$

where the last step follows from Stirling's formula. □

Since we will use this expression for $\mathcal{S}(\delta, \rho)$ throughout this chapter, we will simplify notation by omitting the $\mathcal{S}$ subscript from the tail bound functions.

Definition 5.2. *Let $\delta \in (0, 1]$, $\rho \in (0, 1)$ and $\lambda \in (0, 1]$. Define*

$$\mathit{IU}(\delta, \rho, \lambda) := \mathit{IU}_{\frac{H(\delta\rho)}{\delta}}(\delta, \rho, \lambda) \quad \text{and} \quad \mathit{IL}(\delta, \rho, \lambda) := \mathit{IL}_{\frac{H(\delta\rho)}{\delta}}(\delta, \rho, \lambda). \tag{5.2}$$

Definition 5.3. *Let $\delta \in (0, 1]$ and $\rho \in (0, 1/2]$. Define*

$$\mathit{IF}(\delta, \rho) := \mathit{IF}_{\frac{H(\delta\rho)}{\delta}}(\delta, \rho). \tag{5.3}$$

We define the following two asymptotic recovery phase transitions for IHT algorithms, in terms of these large deviations bounds.

Definition 5.4 (**Recovery phase transitions**). *Given $\delta \in (0,1]$, define $\hat{\rho}_{SP}^{IHT}(\delta)$ to be the unique solution to*

$$\frac{\sqrt{\mathcal{IF}(\delta,\rho)}}{(1-\rho)\left[1-\mathcal{IL}(\delta,\rho,1-\rho)\right]} = \frac{1}{1+\mathcal{U}(\delta,2\rho)} \quad \textit{for} \quad \rho \in (0,1/2], \tag{5.4}$$

and define $\hat{\rho}_{SP}^{NIHT_\kappa}(\delta)$ to be the unique solution to

$$\frac{\sqrt{\mathcal{IF}(\delta,\rho)}}{(1-\rho)\left[1-\mathcal{IL}(\delta,\rho,1-\rho)\right]} = \frac{1}{\kappa[1+\mathcal{U}(\delta,2\rho)]} \quad \textit{for} \quad \rho \in (0,1/2], \tag{5.5}$$

where $\mathcal{IF}$ is defined in (5.3), $\mathcal{IL}$ is defined in (5.2) and $\mathcal{U}$ is defined in Definition 2.11.

We need to establish that $\hat{\rho}_{SP}^{IHT}(\delta)$ and $\hat{\rho}_{SP}^{NIHT_\kappa}(\delta)$ are well-defined, to which we devote the rest of this section. The proof relies upon three lemmas, which we give next. The first lemma shows that each of the tail bound functions in Definition 5.4 is strictly increasing in $\rho \in (0,1/2]$.

Lemma 5.5. *Given $\delta \in (0,1]$, $\mathcal{IF}(\delta,\rho)$, $\mathcal{IL}(\delta,\rho,1-\rho)$ and $\mathcal{U}(\delta,\rho)$ are each strictly increasing in $\rho \in (0,1/2]$.*

Proof: Writing $f := \mathcal{IF}(\delta,\rho)$, differentiating (4.100) with respect to ρ gives

$$\frac{\partial f}{\partial \rho}\left(\frac{1}{1+f}-\frac{\rho}{f}\right) = \ln\left[\frac{f(1-\rho)(1-\delta\rho)^2}{\delta^2\rho^3}\right]. \tag{5.6}$$

Now $f > \rho/(1-\rho)$ by (4.100), which implies that

$$\frac{f(1-\rho)(1-\delta\rho)^2}{\delta^2\rho^3} > \frac{(1-\delta\rho)^2}{(\delta\rho)^2}.$$

The function $(1-\delta\rho)/(\delta\rho)$ is strictly decreasing in $\delta\rho$, and so takes its minimum value when $\delta\rho = 1/2$, from which it follows that the logarithm in (5.6) is strictly positive. Since $f > \rho/(1-\rho)$, we also have

$$\frac{1}{1+f}-\frac{\rho}{f} > 0,$$

and the result now follows for $\mathcal{IF}(\delta,\rho)$. Similarly, writing $\nu := \mathcal{IL}(\delta,\rho,1-\rho)$ and differentiating (4.89) with $\gamma = 1-\rho$ gives

$$(1-\rho)\left(\frac{1}{1-\nu}-1\right)\frac{\partial \nu}{\partial \rho} - [-\nu-\ln(1-\nu)] = 2\ln\left(\frac{1-\delta\rho}{\delta\rho}\right).$$

A further application of (4.89) followed by some rearrangement gives

$$\frac{\partial \nu}{\partial \rho}\left(\frac{\nu(1-\rho)}{1-\nu}\right) = 2\ln\left(\frac{1-\delta\rho}{\delta\rho}\right) + \frac{2H(\delta\rho)}{\delta(1-\rho)}. \tag{5.7}$$

The term on the left-hand side of (5.7) is strictly positive, as is the term involving the Shannon entropy. The function $(1-\delta\rho)/(\delta\rho)$ is strictly decreasing in $\delta\rho$, and so takes its minimum value

when $\delta\rho = 1/2$, which implies that

$$\ln\left(\frac{1-\delta\rho}{\delta\rho}\right) \geq 0,$$

from which the result for $\mathcal{IL}(\delta, \rho, 1-\rho)$ follows. Finally, writing $\lambda := \lambda^{max}(\delta, \rho; \gamma)$, Definition 2.11 gives λ as the solution to

$$H(\delta\rho) - \delta\gamma H\left(\frac{\rho}{\gamma}\right) + \frac{\delta}{2}\left[(1+\gamma)\ln\lambda - \gamma\ln\gamma + 1 + \gamma - \lambda\right] = 0,$$

which differentiates to give

$$\frac{\partial\lambda}{\partial\rho}\left[1 - \frac{\delta(1+\gamma)}{2(1+\lambda)}\right] = \delta\ln\left[\left(\frac{1-\delta\rho}{\delta\rho}\right)\left(1 - \frac{\rho}{\gamma-\rho}\right)\right]. \tag{5.8}$$

Since $\lambda \geq \gamma$, the term on the left-hand side of (5.8) is strictly positive, and by using $\gamma \leq \delta^{-1}$ it is straightforward to show that the right-hand side of (5.8) is also strictly positive for $\rho \leq 1/2$, from which it follows that $\lambda^{max}(\delta, \rho; \gamma)$ is strictly increasing in $\rho \in (0, 1/2]$ for fixed γ. Now supposing $0 < \rho_1 < \rho_2 \leq 1/2$, we have for all $\gamma \in [\rho, \delta^{-1}]$,

$$\lambda^{max}(\delta, \rho_1) \leq \lambda^{max}(\delta, \rho_1; \gamma) < \lambda^{max}(\delta, \rho_2; \gamma),$$

and therefore

$$\lambda^{max}(\delta, \rho_1) < \lambda^{max}(\delta, \rho_2),$$

and the result now follows. □

The second lemma shows that $\mathcal{IF}(\delta, \rho)$ grows to be much greater than 1 at $\rho = 1/2$.

Lemma 5.6. *For any $\delta \in (0, 1]$,*

$$\mathcal{IF}(\delta, 1/2) \geq 31 + 8\sqrt{15}.$$

Proof: Substituting $\rho = 1/2$ into (4.100) implies that $\mathcal{IF}(\delta, 1/2)$ solves for f the equation

$$\ln(1+f) - \frac{1}{2}\ln f = \frac{2}{\delta}\cdot H\left(\frac{\delta}{2}\right) + H\left(\frac{1}{2}\right) \quad \text{for} \quad f > 1.$$

Now $H(1/2) = \ln 2$, and we may lower bound $H(p)/p$ by $2\ln 2$ since $H(p)/p$ is strictly decreasing in $p \in (0, 1/2]$, which implies that

$$\ln(1+f) - \frac{1}{2}\ln f \geq 3\ln 2,$$

which may be rearranged to give

$$f^2 - 62f + 1 \geq 0,$$

which, together with $f > 1$, yields the required result. □

The third lemma proves that each of the tail bound functions in Definition 5.4 tends to zero as $\rho \to 0$.

Lemma 5.7. *Given $\delta \in (0, 1]$, the following limiting results hold:*

$$\lim_{\rho \to 0} \mathcal{IF}(\delta, \rho) = 0; \qquad \lim_{\rho \to 0} \mathcal{IL}(\delta, \rho, 1-\rho) = 0; \qquad \lim_{\rho \to 0} \mathcal{U}(\delta, \rho) = 0.$$

Proof: Recalling (2.34), we have

$$\lim_{p \to 0} H(p) = 0, \tag{5.9}$$

which, together with (5.1), shows that the right-hand side of (4.100) tends to zero as $\rho \to 0$. Writing $f = \mathcal{IF}(\delta, \rho)$, it now follows from (4.100) that

$$\lim_{\rho \to 0} [\ln(1+f) - \rho \ln f] \leq 0. \tag{5.10}$$

However, Lemma 5.6 and (4.100) imply that

$$-\rho \ln(31 + 8\sqrt{15}) \leq -\rho \ln f \leq -\rho \ln \left(\frac{\rho}{1-\rho} \right),$$

from which it follows that $\rho \ln f \to 0$ as $\rho \to 0$, which combines with (5.10) to give $\ln(1+f) \to 0$ as $\rho \to 0$, which yields the result for $\mathcal{IF}(\delta, \rho)$. It also follows from (5.9) and (5.1) that the right-hand side of (4.89) with $\gamma = 1 - \rho$ tends to zero as $\rho \to 0$, which in turn implies the result for $\mathcal{IL}(\delta, \rho)$. To prove the result for $\mathcal{U}(\delta, \rho)$, note that, in Lemma 2.11, if instead of optimizing over $\gamma \in [\rho, \delta^{-1}]$, we instead fix $\gamma = \rho$, we obtain a function $\tilde{\mathcal{U}}(\delta, \rho)$ as the solution in ν of

$$\frac{\delta}{2} \left[(1+\rho) \ln(1+\nu) - \nu + \rho - \rho \ln \rho \right] + H(\delta\rho) = 0 \quad \text{for} \quad \nu \geq \rho,$$

to which we may apply (5.9) to deduce that $\tilde{\mathcal{U}}(\delta, \rho) \to 0$ as $\rho \to 0$. Since, in obtaining $\tilde{\mathcal{U}}(\delta, \rho)$, we have not optimized over γ, we have $\mathcal{U}(\delta, \rho) \leq \tilde{\mathcal{U}}(\delta, \rho)$, and the result for $\mathcal{U}(\delta, \rho)$ now follows.□

That $\hat{\rho}_{SP}^{IHT}(\delta)$ is well-defined can now be shown as follows. Let us define $\psi^{IHT}(\delta, \rho)$ to be

$$\psi^{IHT}(\delta, \rho) := \frac{\sqrt{\mathcal{IF}(\delta, \rho)}}{(1-\rho)\left[1 - \mathcal{IL}(\delta, \rho, 1-\rho)\right]} - \frac{1}{1 + \mathcal{U}(\delta, 2\rho)}.$$

By Lemma 5.7, we have

$$\lim_{\rho \to 0} \psi^{IHT}(\delta, \rho) = -1 < 0.$$

By Lemma 5.5, $\psi^{IHT}(\delta, \rho)$ is strictly increasing in ρ for $\rho \in (0, 1/2]$. Also, Lemma 5.6, $\mathcal{U} \geq 0$ and $\mathcal{IL} \geq 0$ together imply that $\psi^{IHT}(\delta, 1/2) > 0$. It therefore follows that there exists a unique $\rho \in (0, 1/2]$ for which $\psi^{IHT}(\delta, \rho) = 0$ and the definition of $\hat{\rho}_{SP}^{IHT}(\delta)$ is valid. A similar argument applies for $\hat{\rho}_{SP}^{NIHT_\kappa}(\delta)$.

5.3 Novel recovery results for IHT algorithms

We prove our main recovery results for IHT algorithms in this section, beginning with constant stepsize IHT.

5.3.1 Recovery results for IHT

We begin by introducing several definitions.

Definition 5.8 (**Stability factor for IHT**). *Consider Problem 1.3. Given $\delta \in (0, 1]$, $\rho \in (0, 1/2]$ and $\alpha > 0$, provided*

$$\alpha > \frac{\sqrt{\mathcal{IF}(\delta, \rho)}}{(1-\rho)[1 - \mathcal{IL}(\delta, \rho, 1-\rho)]}, \tag{5.11}$$

define

$$a(\delta, \rho) := \frac{1 + \sqrt{\mathcal{IF}(\delta, \rho)} + \alpha\sqrt{\rho(1-\rho)[1 + \mathcal{IU}(\delta, \rho, 1-\rho)][1 + \mathcal{IU}(\delta, \rho, \rho)]}}{\alpha(1-\rho)[1 - \mathcal{IL}(\delta, \rho, 1-\rho)] - \sqrt{\mathcal{IF}(\delta, \rho)}}, \tag{5.12}$$

and

$$\xi(\delta, \rho) := \sqrt{\mathcal{IF}(\delta, \rho)\left[1 + a(\delta, \rho)\right]^2 + 1 + \left[a(\delta, \rho)\right]^2}, \tag{5.13}$$

where $\mathcal{IF}$ is defined in (5.3), and where $\mathcal{IU}$ and $\mathcal{IL}$ are defined in (5.2).

Note that (5.11) ensures that the denominator in (5.12) is strictly positive and that $a(\delta, \rho)$ is therefore well-defined. The function $\xi(\delta, \rho)$ will represent a stability factor in our results, bounding the approximation error of the output of IHT as a multiple of the unrecoverable energy Σ.

Definition 5.9 (**Support set partition for IHT**). *Consider Problem 1.3 and suppose $\delta \in (0, 1]$, $\rho \in (0, 1/2]$ and $\alpha > 0$. Given $\zeta > 0$, let us write*

$$a^*(\delta, \rho; \zeta) := a(\delta, \rho) + \zeta, \tag{5.14}$$

where $a(\delta, \rho)$ is defined in (5.8), let us write $\{\Gamma_i : i \in S_n\}$ for the set of all possible support sets

of cardinality k, and let us disjointly partition $S_n := \Theta_n^1 \cup \Theta_n^2$ such that

$$\Theta_n^1 := \left\{ i \in S_n \ : \ \|x^*_{\Lambda\setminus\Gamma_i}\| > \Sigma \cdot a^*(\delta,\rho;\zeta) \right\}; \quad \Theta_n^2 := \left\{ i \in S_n \ : \ \|x^*_{\Lambda\setminus\Gamma_i}\| \leq \Sigma \cdot a^*(\delta,\rho;\zeta) \right\}, \tag{5.15}$$

where Σ is defined in (4.9), and where Λ is defined in (3.1).

We recall that Λ is defined to be the support of x_k^*, the best k-sparse approximation to x^*. Note that the partition $S_n := \Theta_n^1 \cup \Theta_n^2$ defined in (5.15) also depends on ζ, though we omit this dependency from our notation for the sake of brevity. Note also that if $\Gamma_i = \Lambda$, then $\|x^*_{\Lambda\setminus\Gamma_i}\| = 0$ and $i \in \Theta_n^2$. In other words, the index corresponding to the support set of x_k^* is contained in Θ_n^2.

Let us outline how our argument will proceed. The partition in (5.15) has been defined in such a way that, provided (5.11) holds, an analysis of the stable point condition (3.16) shows that there are asymptotically no α-stable points on any Γ_i such that $i \in \Theta_n^1$, and this is proved in Lemma 5.10. On the other hand, it is also possible to use the large deviations results of Chapter 4 to bound the error in approximating x^* by any α-stable point on Γ_i such that $i \in \Theta_n^2$, which is achieved by Lemma 5.12. It follows that, for any $\alpha > 0$, all α-stable points have bounded approximation error. Finally, Lemma 5.13 builds on the convergence result in Theorem 3.9 and gives a condition on the stepsize α which asymptotically guarantees convergence of IHT to some α-stable point. Combining all three results, we have convergence to some α-stable point with guaranteed approximation error, provided the conditions in each lemma hold; combining the conditions leads to the phase transition defined in (5.4).

We first show that, asymptotically, there are no α-stable points on any Γ_i such that $i \in \Theta_n^1$, and we write NSP_α for this event.

Lemma 5.10. *Consider Problem 1.3 and choose $\zeta > 0$. Suppose Assumptions 2 and 3 hold, and suppose that (5.11) holds. Then, in the proportional-growth asymptotic, there are no α-stable points on any Γ_i such that $i \in \Theta_n^1$, with probability tending to 1 exponentially in n.*

Proof: For any Γ_i such that $i \in \Theta_n^1$, we have $\Gamma_i \neq \Lambda$, and we may therefore use Theorem 3.4 and Lemma 4.10 with $\Gamma := \Gamma_i$ to deduce that a necessary condition for there to be an α-stable point on Γ_i is

$$\|x^*_{\Lambda\setminus\Gamma_i}\|\cdot\sqrt{F_{\Gamma_i}}+\|x^*_{\Gamma_i\setminus\Lambda}\|+\tilde{\sigma}_{\Gamma_i}\cdot\sqrt{G_{\Gamma_i}} \geq \alpha \left[\left(\frac{n-k}{n} \right) \|x^*_{\Lambda\setminus\Gamma_i}\| \cdot R_{\Gamma_i} - \tilde{\sigma}_{\Gamma_i} \cdot \sqrt{\frac{k(n-k)}{n^2} \cdot S_{\Gamma_i} \cdot T_{\Gamma_i}} \right], \tag{5.16}$$

where

$$F_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1); \quad G_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1);$$
$$R_{\Gamma_i} \sim \frac{1}{n-k} \chi^2_{n-k}; \quad S_{\Gamma_i} \sim \frac{1}{n-k} \chi^2_{n-k}; \quad T_{\Gamma_i} \sim \frac{1}{k} \chi^2_k.$$

But (4.9) implies $\tilde{\sigma}_{\Gamma_i} \leq \Sigma$, while $\|x^*_{\Gamma_i \setminus \Lambda}\| \leq \|x^*_{\Lambda^C}\| \leq \Sigma$, and substituting both into (5.16) gives

$$\|x^*_{\Lambda\setminus\Gamma_i}\| \cdot \sqrt{F_{\Gamma_i}} + \Sigma\left(1+\sqrt{G_{\Gamma_i}}\right) \geq \alpha\left[\left(\frac{n-k}{n}\right)\|x^*_{\Lambda\setminus\Gamma_i}\| \cdot R_{\Gamma_i} - \Sigma\sqrt{\frac{k(n-k)}{n^2}\cdot S_{\Gamma_i}\cdot T_{\Gamma_i}}\right]. \tag{5.17}$$

We also have, by (5.15),

$$\Sigma \leq \frac{\|x^*_{\Lambda\setminus\Gamma_i}\|}{a^*(\delta,\rho;\zeta)} \tag{5.18}$$

for any Γ_i such that $i \in \Theta^1_n$. Since $\Gamma_i \neq \Lambda$, $\|x^*_{\Lambda\setminus\Gamma}\| > 0$, and substitution of (5.18) into (5.17), rearrangement and division by $\|x^*_{\Lambda\setminus\Gamma_i}\|$ yields

$$a^*(\delta,\rho;\zeta)\left[\alpha\left(\frac{n-k}{n}\right)\cdot R_{\Gamma_i} - \sqrt{F_{\Gamma_i}}\right] \leq 1 + \sqrt{G_{\Gamma_i}} + \alpha\sqrt{\frac{k(n-k)}{n^2}\cdot S_{\Gamma_i}\cdot T_{\Gamma_i}}.$$

Consequently,

$$\begin{aligned}
& \mathbb{P}(\overline{NSP_\alpha}) \\
= \;& \mathbb{P}\left\{\cup_{i\in\Theta^1_n}(\exists \text{ an } \alpha\text{-stable point supported on } \Gamma_i)\right\} \\
\leq \;& \mathbb{P}\left\{\bigcup_{i\in\Theta^1_n}\left[a^*(\delta,\rho;\zeta)\left[\alpha\,(1-\rho_n)\cdot R_{\Gamma_i} - \sqrt{F_{\Gamma_i}}\right] \leq 1 + \sqrt{G_{\Gamma_i}} + \alpha\sqrt{\rho_n(1-\rho_n)\cdot S_{\Gamma_i}\cdot T_{\Gamma_i}}\right]\right\},
\end{aligned} \tag{5.19}$$

where we write ρ_n for the sequence of values of the ratio k/n. For brevity's sake, let us define

$$\Phi[\rho,F,G,R,S,T] := 1 + \sqrt{G} + \alpha\sqrt{\rho(1-\rho)(S)(T)} - a^*(\delta,\rho;\zeta)\cdot\left[\alpha(1-\rho)\cdot R - \sqrt{F}\right], \tag{5.20}$$

so that (5.19) may be equivalently written as

$$\mathbb{P}(\overline{NSP_\alpha}) \leq \mathbb{P}\left\{\cup_{i\in\Theta^1_n}\left(\Phi[\rho_n, F_{\Gamma_i}, G_{\Gamma_i}, R_{\Gamma_i}, S_{\Gamma_i}, T_{\Gamma_i}] \geq 0\right)\right\}. \tag{5.21}$$

Given some $\epsilon > 0$, we now define

$$\begin{aligned}
F^* = G^* &:= \mathcal{IF}(\delta,\rho) + \epsilon; \qquad R^* := 1 - \mathcal{IL}(\delta,\rho,1-\rho) - \epsilon; \\
S^* &:= 1 + \mathcal{IU}(\delta,\rho,1-\rho) + \epsilon; \qquad T^* := 1 + \mathcal{IU}(\delta,\rho,\rho) + \epsilon.
\end{aligned} \tag{5.22}$$

Using (5.22), we deduce from (5.21) that

$$\begin{aligned}
& \mathbb{P}(\overline{NSP_\alpha}) \\
\leq \;& \mathbb{P}\left\{\cup_{i\in\Theta^1_n}\left(\Phi[\rho_n, F_{\Gamma_i}, G_{\Gamma_i}, R_{\Gamma_i}, S_{\Gamma_i}, T_{\Gamma_i}] \geq \Phi[\rho_n, F^*, G^*, R^*, S^*, T^*]\right)\right\} && (5.23) \\
+ \;& \mathbb{P}\left\{\Phi[\rho_n, F^*, G^*, R^*, S^*, T^*] \geq \Phi[\rho, F^*, G^*, R^*, S^*, T^*] + \epsilon\right\} && (5.24) \\
+ \;& \mathbb{P}\left\{\Phi[\rho, F^*, G^*, R^*, S^*, T^*] + \epsilon \geq 0\right\}, && (5.25)
\end{aligned}$$

since the event in the right-hand side of (5.21) lies in the union of the three events in (5.23), (5.24) and (5.25). Now (5.25) is a deterministic event, and $a^*(\delta, \rho; \zeta)$ has been defined in such a way that, for any $\zeta > 0$, provided ϵ is taken sufficiently small, the event has probability 0. This follows from (5.11), (5.12), (5.14), and by the continuity of Φ. The event (5.24) is also deterministic, and by continuity and since $\rho_n \to \rho$, it follows that there exists some $\tilde{n}$ such that

$$\mathbb{P}\left\{\Phi[\rho_n, F^*, G^*, R^*, S^*, T^*] \geq \Phi[\rho, F^*, G^*, R^*, S^*, T^*] + \epsilon\right\} = 0 \quad \text{for all } \; n \geq \tilde{n}.$$

Taking limits as $n \to \infty$, the terms (5.24) and (5.25) are zero, leaving only (5.23), and we have

$$\begin{aligned}
& \lim_{n\to\infty} \mathbb{P}(\overline{NSP_\alpha}) \\
\leq \; & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1} \left(\Phi[\rho_n, F_{\Gamma_i}, G_{\Gamma_i}, R_{\Gamma_i}, S_{\Gamma_i}, T_{\Gamma_i}] \geq \Phi[\rho_n, F^*, G^*, R^*, S^*, T^*]\right)\right\} \\
\leq \; & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1} (F_{\Gamma_i} \geq F^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1} (G_{\Gamma_i} \geq G^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1} (R_{\Gamma_i} \leq R^*)\right\} \\
+ \; & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1} (S_{\Gamma_i} \geq S^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1} (T_{\Gamma_i} \geq T^*)\right\}, \qquad (5.26)
\end{aligned}$$

where the last line follows from the monotonicity of Φ with respect to F, G, R, S and T. Since $|\Theta_n^1| \leq \binom{N}{k}$, we may apply Lemmas 4.23 and 4.27 with $|S_n| := \binom{N}{k}$ and $\mathcal{S}(\delta, \rho) = H(\delta\rho)/\delta$ to (5.26), and we deduce $\mathbb{P}(\overline{NSP_\alpha}) \to 0$ as $n \to \infty$, exponentially in n, as required. □

In the case of IHT applied to Problem 1.1, the above result has a remarkable corollary: a condition can be given which guarantees that, with overwhelming probability, the underlying k-sparse signal x^* is the algorithm's only fixed point. In other words, within some portion of phase space, there is only one possible solution to which the IHT algorithm can converge, namely the underlying signal x^*. This is remarkable since IHT is a gradient projection algorithm for the nonconvex problem (1.21) which can be shown to have a combinatorially large number of local minimizers. The conclusion is that the properties of Gaussian matrices ensure that, within this region of phase space, the IHT algorithm will never 'get stuck' at an unwanted local minimizer, thus exhibiting a behaviour one would usually only expect if a convex problem was being solved. The result follows.

Corollary 5.11 (**Single fixed point condition**). *Consider Problem 1.1. Suppose Assumption 2 holds, and suppose that (5.11) holds. Then, in the proportional-growth asymptotic, x^* is the only fixed point of IHT with stepsize α, with probability tending to* 1 *exponentially in n.*

Proof: Lemma 5.10 establishes that, if (5.11) holds, there are asymptotically no α-stable points on any Γ_i such that $i \in \Theta_n^1$. Setting $\Sigma := 0$ in (5.15), we have $i \in \Theta_n^2 \Rightarrow \Gamma_i = \Lambda$. Therefore any α-stable point is supported on Λ, and Lemma 3.3 implies that it must be x^*. However, any fixed point of IHT with stepsize α is necessarily an α-stable point, and therefore x^* is also the only fixed point of IHT with stepsize α. □

Next, we show that any α-stable points on Γ_i such that $i \in \Theta_n^2$ are 'close' to x^*.

Lemma 5.12. *Consider Problem 1.3. Suppose Assumptions 2 and 3 hold, and suppose that (5.11) holds. Then there exists ζ sufficiently small such that, in the proportional-growth asymptotic, any α-stable point $\bar{x}$ on Γ_i such that $i \in \Theta_n^2$ satisfies*

$$\|\bar{x} - x^*\| \leq \xi(\delta, \rho) \cdot \Sigma, \tag{5.27}$$

with probability tending to 1 exponentially in n, where $\xi(\delta, \rho)$ is defined in (5.13) and Σ is defined in (4.9).

Proof: Suppose $\bar{x}$ is a minimum-norm solution on Γ, so that $\bar{x}_\Gamma = A_\Gamma^\dagger b$. Then, using $A_\Gamma^\dagger A_\Gamma = I$, we have

$$\begin{aligned}
(\bar{x} - x^*)_\Gamma &= A_\Gamma^\dagger(A_\Gamma x_\Gamma^* + A_{\Gamma^C} x_{\Gamma^C}^* + e) - x_\Gamma^* \\
&= x_\Gamma^* + A_\Gamma^\dagger(A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}^* + A_{(\Lambda\cup\Gamma)^C} x_{(\Lambda\cup\Gamma)^C}^* + e) - x_\Gamma^* \\
&= A_\Gamma^\dagger(A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}^* + \tilde{e}_\Gamma) + x_\Gamma^* - x_\Gamma^* \\
&= A_\Gamma^\dagger(A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}^* + \tilde{e}_\Gamma),
\end{aligned} \tag{5.28}$$

while

$$(\bar{x} - x^*)_{\Gamma^C} = -x_{\Gamma^C}^*. \tag{5.29}$$

Combining (5.28) and (5.29) using the triangle inequality, we may bound

$$\begin{aligned}
\|\bar{x} - x^*\|^2 &= \|(\bar{x} - x^*)_\Gamma\|^2 + \|(\bar{x} - x^*)_{\Gamma^C}\|^2 \\
&= \|A_\Gamma^\dagger(A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}^* + \tilde{e}_\Gamma)\|^2 + \|x_{\Gamma^C}^*\|^2 \\
&\leq \left[\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}^*\| + \|A_\Gamma^\dagger \tilde{e}_\Gamma\|\right]^2 + \|x_{\Lambda\setminus\Gamma}^*\|^2 + \|x_{(\Lambda\cup\Gamma)^C}^*\|^2
\end{aligned} \tag{5.30}$$

We may deduce, by (4.10) of Lemma 4.10,

$$\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x_{\Lambda\setminus\Gamma}^*\|^2 = \|x_{\Lambda\setminus\Gamma}^*\|^2 \cdot P_\Gamma, \quad \text{where} \quad P_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1), \tag{5.31}$$

and by (4.12) of Lemma 4.10,

$$\|A_\Gamma^\dagger \tilde{e}_\Gamma\|^2 = \tilde{\sigma}_\Gamma^2 \cdot Q_\Gamma \leq \Sigma^2 \cdot Q_\Gamma, \quad \text{where} \quad Q_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1). \tag{5.32}$$

Substituting (5.31) and (5.32) into (5.30), and using $\|x_{(\Lambda\cup\Gamma)^C}^*\| \leq \|x_{\Lambda^C}^*\| \leq \Sigma$, we have

$$\|\bar{x} - x^*\|^2 \leq \left[\|x_{\Lambda\setminus\Gamma}^*\| \cdot \sqrt{P_\Gamma} + \Sigma \cdot \sqrt{Q_\Gamma}\right]^2 + \|x_{\Lambda\setminus\Gamma}^*\|^2 + \Sigma^2, \tag{5.33}$$

and we may use (5.15) to further deduce

$$\begin{aligned}\|\bar{x}-x^*\|^2 &\leq \Sigma^2\left[a^*(\delta,\rho;\zeta)\cdot\sqrt{P_\Gamma}+\sqrt{Q_\Gamma}\right]^2+[a^*(\delta,\rho;\zeta)]^2\cdot\Sigma^2+\Sigma^2\\ &= \Sigma^2\left\{\left[a^*(\delta,\rho;\zeta)\cdot\sqrt{P_\Gamma}+\sqrt{Q_\Gamma}\right]^2+1+[a^*(\delta,\rho;\zeta)]^2\right\}. \qquad (5.34)\end{aligned}$$

For the sake of brevity, let us define

$$\Psi(P,Q):=\sqrt{\left(a^*(\delta,\rho;\zeta)\cdot\sqrt{P}+\sqrt{Q}\right)^2+1+a^*(\delta,\rho;\zeta)^2}, \qquad (5.35)$$

so that (5.34) may equivalently be written as

$$\|\bar{x}-x^*\|\leq\Sigma\cdot\Psi\left[P_\Gamma,Q_\Gamma\right]. \qquad (5.36)$$

Given $\zeta>0$, let us define

$$P^*=Q^*:=\mathcal{IF}(\delta,\rho)+\zeta. \qquad (5.37)$$

Now we use (5.36) to perform a union bound over all Γ_i such that $i\in\Theta_n^2$, writing $\bar{x}_i$ for the minimum-norm solution on Γ_i, giving

$$\begin{aligned}&\mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i\in\Theta_n^2 \text{ and } \|\bar{x}_i-x^*\|>\Sigma\cdot\Psi\left[P^*,Q^*\right]\right\}\\ =\;&\mathbb{P}\left\{\bigcup_{i\in\Theta_n^2}\left(\|\bar{x}_i-x^*\|>\Sigma\cdot\Psi\left[P^*,Q^*\right]\right)\right\} \qquad (5.38)\\ \leq\;&\mathbb{P}\left\{\bigcup_{i\in\Theta_n^2}\left(\|\bar{x}_i-x^*\|>\Sigma\cdot\Psi\left[P_{\Gamma_i},Q_{\Gamma_i}\right]\right)\right\} \qquad (5.39)\\ +\;&\mathbb{P}\left\{\bigcup_{i\in\Theta_n^2}\left(\Sigma\cdot\Psi\left[P_{\Gamma_i},Q_{\Gamma_i}\right]\geq\Sigma\cdot\Psi\left[P^*,Q^*\right]\right)\right\}, \qquad (5.40)\end{aligned}$$

since the event in (5.38) lies in the union of the two events in (5.39) and (5.40). It is an immediate consequence of (5.36) that the event in (5.39) has probability 0. Taking limits of (5.40) as $n\to\infty$, and cancelling Σ, we have

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i\in\Theta_n^2 \text{ and } \|\bar{x}_i-x^*\|>\Sigma\cdot\Psi\left[P^*,Q^*\right]\right\}\\ \leq\;&\lim_{n\to\infty}\mathbb{P}\left\{\bigcup_{i\in\Theta_n^2}\left(\Psi\left[P_{\Gamma_i},Q_{\Gamma_i}\right]\geq\Psi\left[P^*,Q^*\right]\right)\right\}\\ \leq\;&\lim_{n\to\infty}\mathbb{P}\left\{\cup_{i\in\Theta_n^2}(P_{\Gamma_i}\geq P^*)\right\}+\lim_{n\to\infty}\mathbb{P}\left\{\cup_{i\in\Theta_n^2}(Q_{\Gamma_i}\geq Q^*)\right\}, \qquad (5.41)\end{aligned}$$

where we used the monotonicity of Ψ with respect to P and Q in the last line. Since $|\Theta_n^2|\leq\binom{N}{k}$,

and using (5.31) and (5.32), we may apply Lemma 4.27 with $|S_n| := \binom{N}{k}$ and $\mathcal{S}(\delta,\rho) = H(\delta\rho)/\delta$ to (5.41), yielding that each of the limits in the right-hand side of (5.41) converges to zero exponentially in n, and so finally

$$\lim_{n\to\infty} \mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and } \|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[a^*(\delta,\rho;\zeta), P^*, Q^*\right]\right\} = 0,$$

exponentially in n. Since by Lemma 3.3, any stable point is necessarily a minimum-norm solution, and recalling the definition of $a^*(\delta,\rho;\zeta)$ in (5.14), $\Psi(a,P,Q)$ in (5.35), and the definitions of P^*, Q^* in (5.37), we have

$$\lim_{n\to\infty} \mathbb{P}\left\{ \begin{array}{l} \exists \text{ some } \alpha\text{-stable point } \bar{x}_i \text{ on } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and} \\ \|\bar{x}_i - x^*\| > \Sigma\sqrt{\mathcal{IF}(\delta,\rho)\left[1+a(\delta,\rho)+\zeta\right]^2 + 1 + \left[a(\delta,\rho)+\zeta\right]^2} \end{array} \right\} = 0, \tag{5.42}$$

with convergence exponential in n. Finally, by continuity,

$$\begin{aligned} &\|\bar{x}_i - x^*\| > \Sigma\sqrt{\mathcal{IF}(\delta,\rho)\left[1+a(\delta,\rho)\right]^2 + 1 + \left[a(\delta,\rho)\right]^2} \\ &\qquad \Longrightarrow \|\bar{x}_i - x^*\| > \Sigma\sqrt{\mathcal{IF}(\delta,\rho)\left[1+a(\delta,\rho)+\zeta\right]^2 + 1 + \left[a(\delta,\rho)+\zeta\right]^2}, \end{aligned}$$

for some ζ suitably small, and the result now follows from the definition of $\xi(\delta,\rho)$ in (5.13). □

In the context of IHT, we obtain the following convergence result in the proportional-dimensional asymptotic framework.

Lemma 5.13. *Consider Problem 1.3. Suppose Assumption 2 holds, suppose that the stepsize α of IHT is chosen to satisfy*

$$\alpha < \frac{1}{1+\mathcal{U}(\delta,2\rho)}. \tag{5.43}$$

Then, in the proportional-growth asymptotic, IHT converges to an α-stable point with probability tending to 1 exponentially in n.

Proof: Given (5.43), we may apply Theorem 2.12 with ϵ sufficiently small to deduce $\alpha(1+U_{2k}) < 1$, with probability tending to 1 exponentially in n. Under Assumption 2, we may then apply Theorem 3.9 and deduce convergence of IHT to an α-stable point. □

We now combine Lemmas 5.10, 5.12 and 5.13 and prove a recovery result for IHT, one of the main results of the thesis.

Theorem 5.14 (**Recovery result for IHT**). *Consider Problem 1.3. Suppose Assumptions 2 and 3 hold, suppose that*

$$\rho < \hat{\rho}_{SP}^{IHT}(\delta), \tag{5.44}$$

where $\hat{\rho}_{SP}^{IHT}(\delta)$ is defined in (5.4), and suppose that the IHT stepsize α satisfies

$$\frac{\sqrt{\mathcal{IF}(\delta,\rho)}}{(1-\rho)\left[1-\mathcal{IL}(\delta,\rho,1-\rho)\right]} < \alpha < \frac{1}{1+\mathcal{U}(\delta,2\rho)}. \tag{5.45}$$

Then, in the proportional-growth asymptotic, IHT converges to $\bar{x}$ such that (5.27) holds with probability tending to 1 exponentially in n.

Proof: First note that (5.44) implies that the interval in (5.45) is well-defined. Provided α is chosen to satisfy (5.45), (5.43) holds, and under Assumption 2, we may apply Lemma 5.13 to deduce convergence of IHT to an α-stable point. On the other hand, Lemma 5.10 establishes that there are asymptotically no α-stable points on any Γ_i such that $i \in \Theta_n^1$, while we may apply Lemma 5.12 to deduce that any α-stable points on any Γ_i such that $i \in \Theta_n^2$ satisfy (5.27).□

In the special case of Problem 1.1, the same phase transition guarantees exact recovery of the underlying signal x^*.

Corollary 5.15 (**Noiseless case**). *Consider Problem 1.1. Suppose Assumption 2 holds, suppose that (5.44) holds, and suppose that α satisfies (5.45). Then, in the proportional-growth asymptotic, IHT converges to x^* with probability tending to 1 exponentially in n.*

Proof: The result follows by setting $\Sigma := 0$ in Theorem 5.14. □

5.3.2 Recovery results for NIHT

In the case of NIHT, it is possible to prove convergence to an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point, where

$$\underline{\alpha}(\delta,\rho;\epsilon) := \{\kappa[1+\mathcal{U}(\delta,2\rho)+\epsilon]\}^{-1}, \tag{5.46}$$

for some $\epsilon > 0$. Due to the dependence of $\underline{\alpha}(\delta,\rho;\epsilon)$ upon (δ,ρ), we need new versions of Definitions 5.8 and 5.9 for NIHT.

Definition 5.16 (**Stability factor for NIHT**). *Consider Problem 1.3. Given $\delta \in (0,1]$ and $\rho \in (0,1/2]$, provided*

$$\rho < \hat{\rho}_{SP}^{NIHT_\kappa}(\delta) \tag{5.47}$$

define

$$a(\delta,\rho) := \frac{1+\sqrt{\mathcal{IF}(\delta,\rho)}+\{\kappa[1+\mathcal{U}(\delta,2\rho)]\}^{-1}\sqrt{\rho(1-\rho)[1+\mathcal{IU}(\delta,\rho,1-\rho)][1+\mathcal{IU}(\delta,\rho,\rho)]}}{(1-\rho)\{\kappa[1+\mathcal{U}(\delta,2\rho)]\}^{-1}[1-\mathcal{IL}(\delta,\rho,1-\rho)]-\sqrt{\mathcal{IF}(\delta,\rho)}}, \tag{5.48}$$

and

$$\xi(\delta,\rho) := \sqrt{\mathcal{IF}(\delta,\rho)\left[1+a(\delta,\rho)\right]^2 + 1 + \left[a(\delta,\rho)\right]^2}, \tag{5.49}$$

where $\mathcal{IF}$ is defined in (5.3), where $\mathcal{IU}$ and $\mathcal{IL}$ are defined in (5.2), and where $\mathcal{U}$ is defined in Definition 2.11.

Note that (5.47) ensures that the denominator in (5.48) is strictly positive and that $a(\delta,\rho)$ is therefore well-defined.

Definition 5.17 (**Support set partition for NIHT**). *Consider Problem 1.3 and suppose $\delta \in (0,1]$ and $\rho \in (0,1/2]$. Given $\zeta > 0$, let us write*

$$a^*(\delta,\rho;\zeta) := a(\delta,\rho) + \zeta, \tag{5.50}$$

where $a(\delta,\rho)$ is defined in (5.48), let us write $\{\Gamma_i : i \in S_n\}$ for the set of all possible support sets of cardinality k, and let us disjointly partition $S_n := \Theta_n^1 \cup \Theta_n^2$ such that

$$\Theta_n^1 := \left\{ i \in S_n \;\; : \;\; \|x^*_{\Lambda\setminus\Gamma_i}\| > \Sigma \cdot a^*(\delta,\rho;\zeta) \right\}; \quad \Theta_n^2 := \left\{ i \in S_n \;\; : \;\; \|x^*_{\Lambda\setminus\Gamma_i}\| \leq \Sigma \cdot a^*(\delta,\rho;\zeta) \right\}, \tag{5.51}$$

where Σ is defined in (4.9) and Λ is defined in (3.1).

The reader may verify by comparison with Definitions 5.8 and 5.9 that, essentially, the α terms have been replaced by the term $\{\kappa[1+\mathcal{U}(\delta,2\rho)]\}^{-1}$.

The proof of recovery results for NIHT corresponding to those for IHT in Section 5.3.1 takes broadly the same approach. However, in order to finally eliminate the dependence upon ϵ in $\underline{\alpha}(\delta,\rho;\epsilon)$, the results corresponding to Lemmas 5.10 and 5.13 for IHT need to be combined together. This is accomplished by Lemma 5.18, which establishes that, provided (5.47) holds and ϵ is taken sufficiently small, NIHT converges to an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point on some Γ_i such that $i \in \Theta_n^2$. Lemma 5.19 corresponds to Lemma 5.12 for IHT, giving bounds on the approximation error of an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point on some Γ_i such that $i \in \Theta_n^2$, for any $\epsilon > 0$. Combining the two lemmas leads us to conclude that NIHT converges to some limit point with bounded approximation error. We write $NSP_{\underline{\alpha}}$ for the event that there is no $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point on any Γ_i such that $i \in \Theta_n^1$.

Lemma 5.18. *Consider Problem 1.3 and choose $\zeta > 0$. Suppose Assumptions 2 and 3 hold, and suppose that (5.47) holds. Then there exists ϵ such that, in the proportional-growth asymptotic, NIHT with shrinkage parameter κ converges to an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point on some Γ_i such that $i \in \Theta_n^2$, with probability tending to 1 exponentially in n.*

Proof: Under Assumption 2, we have by Theorem 3.10 convergence of NIHT to a $[\kappa(1+U_{2k})]^{-1}$-stable point. By Definition 3.2, for any $\alpha_1 < \alpha_2$, the set of α_1-stable points includes the set of α_2-stable points, and this observation combines with Theorem 2.12 to imply convergence

to an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point, where $\underline{\alpha}(\delta,\rho;\epsilon)$ is defined in (5.46), with probability tending to 1 exponentially in n. We now rehearse the argument of Lemma 5.10 to show that, provided ϵ is taken sufficiently small, this stable point must be on Γ_i such that $i \in \Theta_n^2$. For any Γ_i such that $i \in \Theta_n^1$, we have $\Gamma_i \neq \Lambda$, and we may therefore use Theorem 3.4 and Lemma 4.10 with $\Gamma := \Gamma_i$ to deduce that, given some $\epsilon > 0$, a necessary condition for there to be an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point on Γ_i is

$$\begin{aligned} &\|x^*_{\Lambda\setminus\Gamma_i}\| \cdot \sqrt{F_{\Gamma_i}} + \|x^*_{\Gamma_i\setminus\Lambda}\| + \tilde{\sigma}_{\Gamma_i} \cdot \sqrt{G_{\Gamma_i}} \\ &\qquad \geq \underline{\alpha}(\delta,\rho;\epsilon) \left[\left(\tfrac{n-k}{n}\right) \|x^*_{\Lambda\setminus\Gamma_i}\| \cdot R_{\Gamma_i} - \tilde{\sigma}_{\Gamma_i} \cdot \sqrt{\tfrac{k(n-k)}{n^2} \cdot S_{\Gamma_i} \cdot T_{\Gamma_i}} \right], \end{aligned} \tag{5.52}$$

where

$$\begin{aligned} F_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1); \quad G_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1); \\ R_{\Gamma_i} \sim \frac{1}{n-k} \chi^2_{n-k}; \quad S_{\Gamma_i} \sim \frac{1}{n-k} \chi^2_{n-k}; \quad T_{\Gamma_i} \sim \frac{1}{k} \chi^2_k. \end{aligned}$$

But (4.9) implies $\tilde{\sigma}_{\Gamma_i} \leq \Sigma$, while $\|x^*_{\Gamma_i\setminus\Lambda}\| \leq \|x^*_{\Lambda^C}\| \leq \Sigma$, and substituting both into (5.52) gives

$$\|x^*_{\Lambda\setminus\Gamma_i}\| \cdot \sqrt{F_{\Gamma_i}} + \Sigma \left(1 + \sqrt{G_{\Gamma_i}}\right) \geq \underline{\alpha}(\delta,\rho;\epsilon) \left[\left(\frac{n-k}{n}\right) \|x^*_{\Lambda\setminus\Gamma_i}\| \cdot R_{\Gamma_i} - \Sigma \sqrt{\frac{k(n-k)}{n^2} \cdot S_{\Gamma_i} \cdot T_{\Gamma_i}} \right]. \tag{5.53}$$

We also have, by (5.51),

$$\Sigma \leq \frac{\|x^*_{\Lambda\setminus\Gamma_i}\|}{a^*(\delta,\rho;\zeta)} \tag{5.54}$$

for any Γ_i such that $i \in \Theta_n^1$. Since $\Gamma_i \neq \Lambda$, $\|x^*_{\Lambda\setminus\Gamma}\| > 0$, and substitution of (5.54) into (5.53), rearrangement and division by $\|x^*_{\Lambda\setminus\Gamma_i}\|$ yields

$$a^*(\delta,\rho;\zeta) \left[\underline{\alpha}(\delta,\rho;\epsilon) \left(\frac{n-k}{n}\right) \cdot R_{\Gamma_i} - \sqrt{F_{\Gamma_i}} \right] \leq 1 + \sqrt{G_{\Gamma_i}} + \underline{\alpha}(\delta,\rho;\epsilon) \sqrt{\frac{k(n-k)}{n^2} \cdot S_{\Gamma_i} \cdot T_{\Gamma_i}},$$

and consequently

$$\begin{aligned} \mathbb{P}(\overline{NSP_{\underline{\alpha}}}) &= \mathbb{P}\left\{ \cup_{i \in \Theta_n^1} (\exists \text{ an } \underline{\alpha}(\delta,\rho;\epsilon)\text{-stable point supported on } \Gamma_i) \right\} \\ &\leq \mathbb{P}\left\{ \cup_{i \in \Theta_n^1} (\Phi[\rho_n, F_{\Gamma_i}, G_{\Gamma_i}, R_{\Gamma_i}, S_{\Gamma_i}, T_{\Gamma_i}] \geq 0) \right\}, \end{aligned} \tag{5.55}$$

where we write ρ_n for the sequence of values of the ratio k/n, and where

$$\Phi[\rho, F, G, R, S, T] := 1 + \sqrt{G} + \underline{\alpha}(\delta,\rho;\epsilon) \sqrt{\rho(1-\rho)(S)(T)} - a^*(\delta,\rho;\zeta) \cdot \left[\underline{\alpha}(\delta,\rho;\epsilon)(1-\rho) \cdot R - \sqrt{F} \right]. \tag{5.56}$$

We now define

$$\begin{aligned} F^* = G^* &:= \mathcal{IF}(\delta,\rho) + \epsilon; \qquad R^* := 1 - \mathcal{IL}(\delta,\rho,1-\rho) - \epsilon; \\ S^* &:= 1 + \mathcal{IU}(\delta,\rho,1-\rho) + \epsilon; \qquad T^* := 1 + \mathcal{IU}(\delta,\rho,\rho) + \epsilon. \end{aligned} \tag{5.57}$$

Using (5.57), we deduce from (5.55) that

$$\begin{aligned}
& \mathbb{P}(\overline{NSP_{\underline{\alpha}}}) \\
\leq\ & \mathbb{P}\left\{\cup_{i\in\Theta_n^1}\left(\Phi[\rho_n, F_{\Gamma_i}, G_{\Gamma_i}, R_{\Gamma_i}, S_{\Gamma_i}, T_{\Gamma_i}] \geq \Phi[\rho_n, F^*, G^*, R^*, S^*, T^*]\right)\right\} && (5.58)\\
+\ & \mathbb{P}\left\{\Phi[\rho_n, F^*, G^*, R^*, S^*, T^*] \geq \Phi[\rho, F^*, G^*, R^*, S^*, T^*] + \epsilon\right\} && (5.59)\\
+\ & \mathbb{P}\left\{\Phi[\rho, F^*, G^*, R^*, S^*, T^*] + \epsilon \geq 0\right\}, && (5.60)
\end{aligned}$$

since the event in (5.55) lies in the union of the three events in (5.58), (5.59) and (5.60). Now (5.60) is a deterministic event, and $a^*(\delta, \rho; \zeta)$ has been defined in such a way that, for any $\zeta > 0$, provided ϵ is taken sufficiently small, the event has probability 0. This follows from (5.47), (5.48), (5.50), and by the continuity of Φ. The event (5.59) is also deterministic, and by continuity and since $\rho_n \to \rho$, it follows that there exists some $\tilde{n}$ such that

$$\mathbb{P}\left\{\Phi[\rho_n, F^*, G^*, R^*, S^*, T^*] \geq \Phi[\rho, F^*, G^*, R^*, S^*, T^*] + \epsilon\right\} = 0 \quad \text{for all } \ n \geq \tilde{n}.$$

Taking limits as $n \to \infty$, the terms (5.59) and (5.60) are zero, leaving only (5.58), and we have

$$\begin{aligned}
& \lim_{n\to\infty} \mathbb{P}(\overline{NSP_{\underline{\alpha}}}) \\
\leq\ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}\left(\Phi[\rho_n, F_{\Gamma_i}, G_{\Gamma_i}, R_{\Gamma_i}, S_{\Gamma_i}, T_{\Gamma_i}] \geq \Phi[\rho_n, F^*, G^*, R^*, S^*, T^*]\right)\right\} \\
\leq\ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(F_{\Gamma_i} \geq F^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(G_{\Gamma_i} \geq G^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(R_{\Gamma_i} \leq R^*)\right\} \\
+\ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(S_{\Gamma_i} \geq S^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(T_{\Gamma_i} \geq T^*)\right\}, && (5.61)
\end{aligned}$$

where the last line follows from the monotonicity of Φ with respect to F, G, R, S and T. Since $|\Theta_n^1| \leq \binom{N}{k}$, we may apply Lemmas 4.23 and 4.27 with $|S_n| := \binom{N}{k}$ and $\mathcal{S}(\delta, \rho) = H(\delta\rho)/\delta$ to (5.61), and we deduce $\mathbb{P}(\overline{NSP_{\underline{\alpha}}}) \to 0$ as $n \to \infty$, exponentially in n, as required. □

Lemma 5.19. *Consider Problem 1.3. Suppose Assumptions 2 and 3 hold, and suppose that (5.47) holds. Given any $\epsilon > 0$, there exists ζ sufficiently small such that, in the proportional-growth asymptotic, any $\underline{\alpha}(\delta, \rho; \epsilon)$-stable point on Γ_i such that $i \in \Theta_n^2$ satisfies*

$$\|\bar{x} - x^*\| \leq \xi(\delta, \rho) \cdot \Sigma, \tag{5.62}$$

with probability tending to 1 exponentially in n, where $\xi(\delta, \rho)$ is defined in (5.49) and Σ is defined in (4.9).

Proof: Suppose $\bar{x}$ is a minimum-norm solution on Γ, so that $\bar{x}_\Gamma = A_\Gamma^\dagger b$. Then we may follow the argument of Lemma 5.12 to deduce (5.33), where

$$P_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1); \quad Q_\Gamma \sim \frac{k}{n-k+1} F(k, n-k+1). \tag{5.63}$$

Combining (5.33) with (5.51), we may further deduce

$$\begin{aligned} \|\bar{x} - x^*\|^2 &\leq \Sigma^2 \left[a^*(\delta, \rho; \zeta) \cdot \sqrt{P_\Gamma} + \sqrt{Q_\Gamma} \right]^2 + [a^*(\delta, \rho; \zeta)]^2 \cdot \Sigma^2 + \Sigma^2 \\ &= \Sigma^2 \left\{ \left[a^*(\delta, \rho; \zeta) \cdot \sqrt{P_\Gamma} + \sqrt{Q_\Gamma} \right]^2 + 1 + [a^*(\delta, \rho; \zeta)]^2 \right\}. \end{aligned} \tag{5.64}$$

For the sake of brevity, let us define

$$\Psi[P, Q] := \sqrt{\left(a^*(\delta, \rho; \zeta) \cdot \sqrt{P} + \sqrt{Q} \right)^2 + 1 + a^*(\delta, \rho; \zeta)^2}, \tag{5.65}$$

so that (5.64) may equivalently be written as

$$\|\bar{x} - x^*\| \leq \Sigma \cdot \Psi \left[P_\Gamma, Q_\Gamma \right]. \tag{5.66}$$

Given $\zeta > 0$, let us define

$$P^* = Q^* := \mathcal{IF}(\delta, \rho) + \zeta. \tag{5.67}$$

Now we use (5.66) to perform a union bound over all Γ_i such that $i \in \Theta_n^2$, writing $\bar{x}_i$ for the minimum-norm solution on Γ_i, giving

$$\begin{aligned} & \mathbb{P}\left\{ \exists \text{ some } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and } \|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[P^*, Q^*\right] \right\} \\ = \; & \mathbb{P}\left\{ \bigcup_{i \in \Theta_n^2} \left(\|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[P^*, Q^*\right] \right) \right\} && (5.68) \\ \leq \; & \mathbb{P}\left\{ \bigcup_{i \in \Theta_n^2} \left(\|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[P_{\Gamma_i}, Q_{\Gamma_i}\right] \right) \right\} && (5.69) \\ + \; & \mathbb{P}\left\{ \bigcup_{i \in \Theta_n^2} \left(\Sigma \cdot \Psi\left[P_{\Gamma_i}, Q_{\Gamma_i}\right] \geq \Sigma \cdot \Psi\left[P^*, Q^*\right] \right) \right\}, && (5.70) \end{aligned}$$

since the event in (5.68) lies in the union of the two events in (5.69) and (5.70). It is an immediate consequence of (5.66) that the event in (5.69) has probability 0. Taking limits of

(5.70) as $n \to \infty$, and cancelling Σ, we have

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and } \|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[P^*, Q^*\right]\right\} \\ \leq & \lim_{n\to\infty} \mathbb{P}\left\{ \bigcup_{i\in\Theta_n^2} \left(\Psi\left[P_{\Gamma_i}, Q_{\Gamma_i}\right] \geq \Psi\left[P^*, Q^*\right]\right) \right\} \\ \leq & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^2}(P_{\Gamma_i} \geq P^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^2}(Q_{\Gamma_i} \geq Q^*)\right\}, \end{aligned} \tag{5.71}$$

where we used the monotonicity of Ψ with respect to P and Q in the last line. Since $|\Theta_n^2| \leq \binom{N}{k}$, and using (5.63), we may apply Lemma 4.27 with $|S_n| := \binom{N}{k}$ and $\mathcal{S}(\delta,\rho) = H(\delta\rho)/\delta$ to (5.71), yielding that each of the limits in the right-hand side of (5.71) converges to zero exponentially in n, and so finally

$$\lim_{n\to\infty} \mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and } \|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[P^*, Q^*\right]\right\} = 0,$$

with convergence at a rate exponential in n also by Lemma 4.27. Since by Lemma 3.3, any stable point is necessarily a minimum-norm solution, and recalling the definition of $\Psi(P,Q)$ in (5.35), and the definitions of P^*, Q^* in (5.67), we have

$$\lim_{n\to\infty} \mathbb{P}\left\{ \begin{array}{l} \exists \text{ some } \underline{\alpha}\text{-stable point } \bar{x}_i \text{ on } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and} \\ \|\bar{x}_i - x^*\| > \Sigma\sqrt{\mathcal{IF}(\delta,\rho)\left[1 + a(\delta,\rho) + \zeta\right]^2 + 1 + \left[a(\delta,\rho) + \zeta\right]^2} \end{array} \right\} = 0, \tag{5.72}$$

with convergence exponential in n. Finally, by continuity,

$$\begin{aligned} & \|\bar{x}_i - x^*\| > \Sigma\sqrt{\mathcal{IF}(\delta,\rho)\left[1 + a(\delta,\rho)\right]^2 + 1 + \left[a(\delta,\rho)\right]^2} \\ & \qquad \Longrightarrow \|\bar{x}_i - x^*\| > \Sigma\sqrt{\mathcal{IF}(\delta,\rho)\left[1 + a(\delta,\rho) + \zeta\right]^2 + 1 + \left[a(\delta,\rho) + \zeta\right]^2}, \end{aligned}$$

for some ζ suitably small, and the result now follows from the definition of $\xi(\delta,\rho)$ in (5.49). □

Combining Lemmas 5.18 and 5.19, we have the following recovery result for NIHT.

Theorem 5.20 (**Recovery result for NIHT**). *Consider Problem 1.3. Suppose Assumptions 2 and 3 hold and suppose that (5.47) holds. Then, in the proportional-growth asymptotic, NIHT converges to $\bar{x}$ such that (5.62) holds with probability tending to* 1 *exponentially in n.*

Proof: By Lemma 5.18, there exists $\epsilon > 0$ such that NIHT with shrinkage parameter κ converges to an $\underline{\alpha}(\delta,\rho;\epsilon)$-stable point on some Γ_i such that $i \in \Theta_n^2$, and for this choice of ϵ, we can apply Lemma 5.19 to deduce the result. □

In the case of Problem 1.1, Theorem 5.20 simplifies to an exact recovery result.

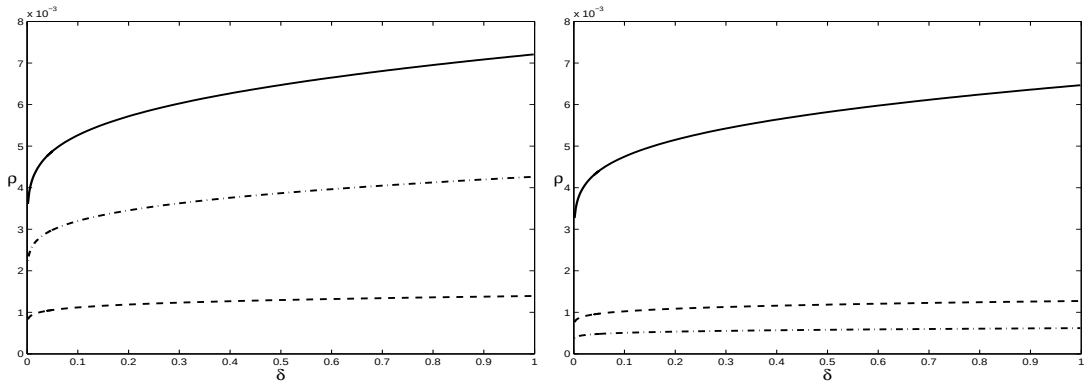
Corollary 5.21 (**Noiseless case**). *Consider Problem 1.1. Suppose Assumption 2 holds and suppose that (5.47) holds. Then, in the proportional-growth asymptotic, NIHT with shrinkage parameter κ converges to x^* with probability tending to* 1 *exponentially in* n.

Proof: The result follows by setting $\Sigma := 0$ in Theorem 5.14. □

5.4 Illustration and discussion of recovery results

5.4.1 Noiseless case

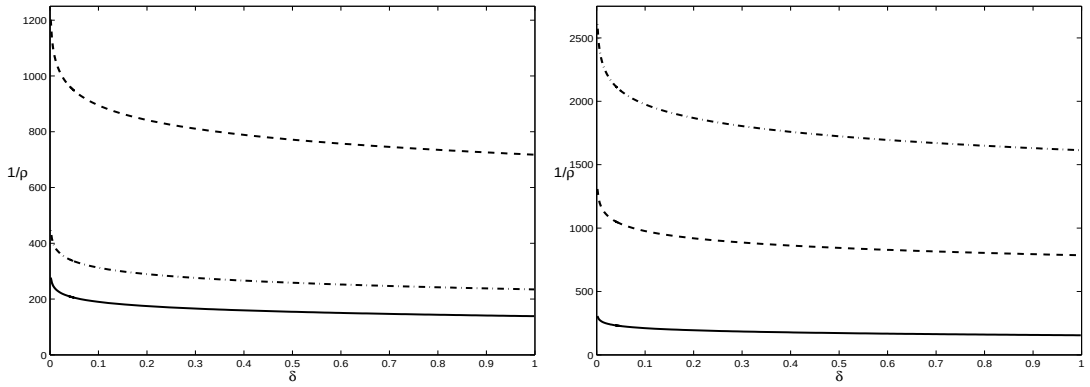
Exact recovery results. Let us first consider the simplified case of Problem 1.1, where we seek to recover an exactly k-sparse signal x^* from noiseless measurements. Corollaries 5.15 and 5.21 establish that, given a sequence of k-sparse signals x^* and independently drawn $n \times N$ Gaussian measurement matrices A, provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{SP}^{IHT}(\delta)$ or $\hat{\rho}_{SP}^{NIHT_\kappa}(\delta)$, defined in (5.4) and (5.5) respectively, the probability of exact recovery of the original signal x^* approaches 1 exponentially fast in n. The functions $\hat{\rho}_{SP}^{IHT}(\delta)$ and $\hat{\rho}_{SP}^{NIHT_\kappa}(\delta)$ define curves in the (δ, ρ)-plane, which are plotted in Figure 5.2. Also plotted in Figure 5.2 to enable a comparison are the current best phase transitions derived from an RIP analysis. We plot both the results obtained in Chapter 2 based upon Foucart's analysis [93] and the results derived in Chapter 3 from our own stable point analysis. As in previous plots in this thesis, a shrinkage parameter of $\kappa := 1.1$ is selected for NIHT in both cases. We observe in both cases an improvement in comparison with the RIP-based phase transitions, by a factor of approximately 1.7 for IHT, and by around an even larger factor of approximately 5 for NIHT. It should also be added that our result for IHT holds for a continuous stepsize range, while the result based upon [93], in keeping with all other similar RIP-based results for IHT surveyed in Section 2.4, holds true only if the stepsize is optimized to a particular value.



(a) (b)

Figure 5.2: Our average-case phase transitions for IHT algorithms (unbroken) compared with the best-known RIP-based phase transitions based on our stable point analysis (dashed) and the analysis in [93] (dash-dot): (a) IHT (b) NIHT.

Figure 5.3 displays the inverse of $\hat{\rho}_{SP}^{IHT}(\delta)$ and $\hat{\rho}_{SP}^{NIHT_{1.1}}(\delta)$. RIP analysis requires a lower bound of $n \geq 234k$ measurements to guarantee recovery using IHT (see Section 2.4), and $n \geq 785k$ using NIHT. By comparison, we reduce these lower bounds to $n \geq 138k$ for IHT and $n \geq 154k$ for NIHT.



(a) (b)

Figure 5.3: Inverse of the phase transitions in Figure 5.2: (a) IHT (b) NIHT.

Interpretation as lower bounds on a weak phase transition. It is important to point out the average-case nature of these results, in contrast to the RIP-based results presented in Chapters 2 and 3, and in Section 5.1 of the present chapter. Worst-case RIP results prove that, with exponentially high probability on the draw of a Gaussian measurement matrix, IHT/NIHT will recover any k-sparse vector from measurements generating by this matrix. The results in Section 5.3 do not provide such strong statements: rather, it is asymptotically guaranteed that, given some k-sparse signal and independent Gaussian measurements, IHT/NIHT will recover this signal from its measurements with exponentially high probability. In other words, we prove results concerning the recovery probability for an independently-chosen k-sparse signal, rather than for all possible k-sparse signals.

There is a strong parallel between the results presented in this chapter and the weak phase transitions for l_1-l_0 equivalence of Donoho and Tanner [77] discussed in Section 1.5. We recall that, below the strong phase transition, it is asymptotically guaranteed that, given a randomly chosen Gaussian measurement matrix, l_1-l_0 equivalence holds for any instance of Problem 1.1 in which that Gaussian matrix is used to generate the measurements. On the other hand, below the weak phase transition, it is asymptotically guaranteed that, given a randomly chosen Gaussian measurement matrix A and a signal x^* with randomly (or independently) chosen support and sign pattern, l_1-l_0 equivalence holds for that particular instance of Problem 1.1 with original signal x^* and measurement matrix A. Our results may be viewed as lower bounds on a particular kind of weak phase transition for recovery using IHT algorithms, in which the signal is assumed to be statistically independent of the measurement matrix, The notions are

comparable but not identical: in the case of l_1-l_0 equivalence, some dependency between the signal and measurement matrix is permitted: it is only required that the support set and sign pattern of the signal are chosen independently of the matrix. However, independence is the only assumption we place upon the signal, and beyond this there is no further restriction upon the signal's coefficients.

We also recall that it is the weak phase transition that is observed empirically for l_1-l_0 equivalence, and the same is also to be expected for IHT algorithms. While we obtain a significant improvement, our lower bound is still pessimistic compared to the weak phase transition observed empirically, though we have succeeded in narrowing the gap between the two. It is no surprise that our results do not give the precise weak phase transition, due to the continued use of worst-case techniques, such as the RIP and large deviations analysis. However, the use of the average-case independence assumption to analyse the stable point condition has allowed us to break free in part from the restrictions of worst-case analysis.

Choice of stepsize for IHT. Corollary 5.15 guarantees exact recovery using IHT provided the stepsize α falls within the interval given in (5.45), provided this interval is well-defined. In fact, an inspection of the proof of these two results reveals that the lower bound in (5.45) arises from the stable point condition, while the upper bound in (5.45) arises from the convergence condition. Figure 5.4.1 illustrates these bounds for the case $\delta = 0.5$. We see that, as ρ is increased, the admissible stepsize range contracts, until a critical ρ-value is reached at which the interval is no longer well-defined.

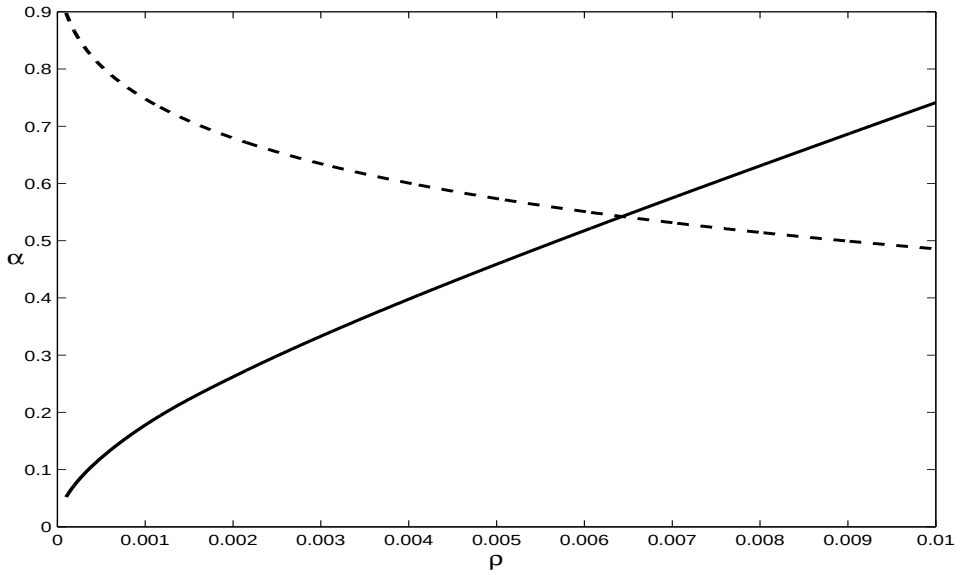


Figure 5.4: Lower bound (unbroken) and upper bound (dashed) on the IHT stepsize for $\delta = 0.5$.

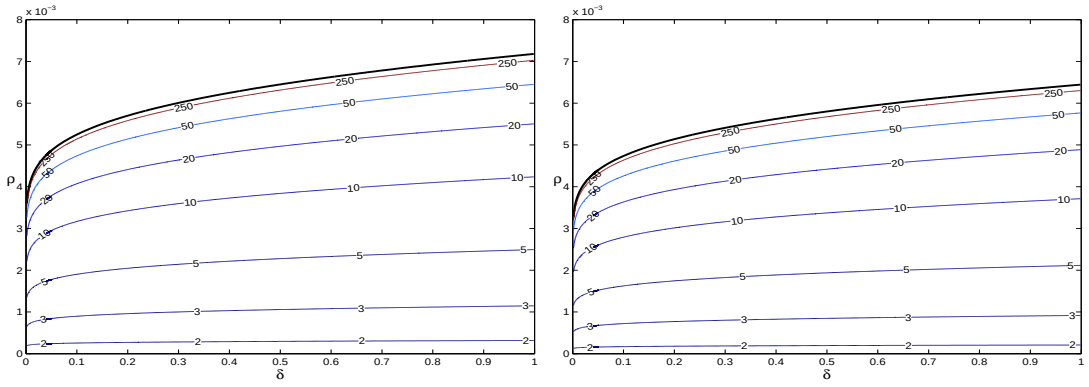
It has been observed empirically [72] that care must be taken to ensure that the IHT stepsize is neither too small or too large. Our analysis gives theoretical insight into this observation:

the stepsize must be small enough to ensure that the algorithm converges, but large enough to ensure that it does not converge to fixed points other than the underlying sparse signal.

5.4.2 General case

Stability factors. In the case of Problem 1.3, where signals are no longer exactly k-sparse and measurements are contaminated by noise, exact recovery of the original signal is impossible. However, Theorem 5.14 guarantees that, under the same condition guaranteeing exact recovery for Problem 1.1, the approximation error of the output of IHT/NIHT is asymptotically bounded by some known stability factor $\xi(\delta, \rho)$ multiplied by the unrecoverable energy Σ. More precisely, consider a sequence of k-compressible signals x^* and independently drawn $n \times N$ Gaussian measurement matrices A. Provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{SP}^{IHT}(\delta)$, then the probability that IHT with stepsize satisfying (5.45) satisfies (5.27) with $\underline{\alpha} := \alpha$ approaches 1 exponentially fast in n. Similarly, provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{SP}^{NIHT_\kappa}(\delta)$, then the probability that NIHT with shrinkage parameter κ satisfies (5.62) with $\underline{\alpha} := \{\kappa[1 + \mathcal{U}(\delta, 2\rho)]\}^{-1}$ approaches 1 exponentially fast in n.

Plots of the stability factor $\xi(\delta, \rho)$ for both IHT and NIHT ($\kappa := 1.1$) are displayed in Figure 5.5. For IHT, there is some freedom in the choice of the stepsize: for these plots we select a stepsize at the upper limit of the permissible interval (5.45) since this minimizes the stability factor. In keeping with the plots in Chapter 2, we observe that the stability factor tends to infinity as the transition point is reached.



(a) (b)

Figure 5.5: Plots of the stability factor $\xi(\delta, \rho)$ for (a) IHT (b) NIHT.

Comparison with previous stability results. Let us now compare these stability results with those presented in Chapter 2 for the previous RIP-based analysis. The stability factors plotted in Figure 2.2 are in the context of Problem 1.2, where the underlying signal is assumed to be k-sparse, and therefore we must first interpret these results in the context of Problem 1.3 before we can make a comparison. In this case, we see from Lemma 2.2 that the stability factor

in Figure 2.2 multiplies a different measure of the unrecoverable energy $\tilde{\Sigma}$, namely

$$\tilde{\Sigma} := \|e\| + \sqrt{1+U_k}\left(\|x^*_{\Lambda^C}\| + \frac{1}{\sqrt{k}}\|x^*_{\Lambda^C}\|_1\right). \tag{5.73}$$

It follows from (5.73) that the corresponding stability factor for Problem 1.3 is necessarily greater than the corresponding factor for Problem 1.2, due to amplification by the RIP term. Furthermore, the l_1-norm term is dependent upon the decay properties of the tail of the underlying signal. In the very worst case in which there is negligible decay, the l_1 term is bounded by $\sqrt{\frac{N}{k}}\|x^*_{\Lambda^C}\|$. Translated into the proportional-dimensional framework, this would give a much worsened bound for small values of ρ. In practice, it may be realistic to expect that a signal exhibits faster decay, in which case the l_1 term could become negligible, for example in the case where the signal coefficients decay according to a power law [7]. Nonetheless, we make the point that quantitative results for Problem 1.3 obtainable from RIP analysis are dependent upon the decay properties of the tail of the signal, which is not the case for our results.

A comparison of Figure 5.5 with Figure 2.2 shows that, for both IHT and NIHT, in the region for which the stability factors derived in this chapter are defined, they are everywhere lower than the corresponding stability factors derived from the previous analysis in Chapter 2. Furthermore, the discussion in the previous paragraph shows that, if a fair comparison of stability factors is made in the context of Problem 1.3, the improvement is even greater than appears at first glance from the plots.

We obtain our results without having to make any implicit assumptions about the decay of the tail of the signal. On the other hand, we do impose additional restrictions upon the noise, namely that the noise is Gaussian distributed and independent of the signal and measurement matrix. This assumption is in keeping with our aim of performing average-case analysis. Our analysis could, however, be altered to deal with the case of non-independent noise by making more use of the RIP, though this would lead to larger stability constants.

Chapter 6

ITP algorithms for tree-based recovery

We now shift our focus to the extension of CS recovery analysis to tree-based signal models.[1] In Section 6.1 of this chapter, we describe in more technical detail the tree-based sparsity model and the Iterative Tree Projection (ITP) algorithm [24, 7], both of which were described in Section 1.9. ITP may be viewed as the extension of IHT to the tree-based setting. We also introduce a variable stepsize variant called Normalised Iterative Tree Projection (NITP), which is the extension of NIHT [27] to the tree-based setting. A crucial question for algorithms such as ITP and NITP is how to perform the tree projection, and in Section 6.2 we propose a dynamic programming (DP) algorithm which is guaranteed to perform an exact tree projection in polynomial time. To the best of our knowledge, it is the only current tree projection algorithm with sub-exponential complexity which is guaranteed to exactly perform the projection. Then, in Section 6.3, we extend the stable point recovery analysis of Chapter 3 to ITP algorithms. The final goal is to quantify these results for Gaussian matrices within a simplified proportional-growth asymptotic framework, which is the focus of Chapter 7.

6.1 Tree-based models and Iterative Tree Projection (ITP)

We recall from Section 1.8 that more refined sparsity models are appropriate for many signals. One such model, which was introduced in Section 1.9, and which applies to the wavelet representation of piecewise smooth signals, is the *tree-based* model. Let us begin by giving some background on wavelets and explaining why the wavelet coefficients of piecewise smooth signals have an inherent tree structure.

Wavelet transforms arrived on the scene in the 1980s as an alternative to the Fourier transform and, as the name suggests, consist of small pieces of a wave. In particular, a wavelet transform consists in convolving a signal with particular instances of a given wavelet function (often called the mother wavelet) at different scales and positions. The theory of wavelets,

[1] Material in the following two chapters is in preparation for submission in [43], which is a joint authorship with C. Cartis whose permission has been obtained for the inclusion of the material.

first developed in the context of continuous signals, also gave rise to the discrete wavelet transform (DWT), which is where our interest lies. We refer the reader to [123] and [59] for more background on the theory of wavelets.

The origins of the DWT can in fact be traced back to the proposal of the *Haar wavelet* in 1909 in Haar's PhD thesis [104], a long time before wavelets became popular in signal processing. The late 1980s saw an explosion of research into DWTs, pioneered by the work of Strömberg [144], Daubechies [58], Meyer [125] and Mallat [122]. There is now a rich bank of DWTs available, for example the family of *Daubechies wavelets* (which can be viewed as a generalization of the Haar wavelet), *Coiflets*, *Legendre wavelets*, and many more.

Despite their differences, all these DWTs have in common a *dyadic* multi-scale tree structure in which the signal is represented at progressively finer scales by repeatedly dividing the support of the signal in two. This induces a natural tree structure on the wavelet coefficients, and a parent-child relationship between wavelet coefficients at different scales. For 1-dimensional wavelets, a binary tree structure results in which each parent node has precisely two children. More general D-dimensional wavelets give rise to a 2^D-ary tree structure: for example, 2-dimensional wavelets have a *quad-tree* structure in which each parent node has precisely four children [7].

By exploiting their nested-support tree structure, many DWTs can be implemented as fast transforms which have complexity $\mathcal{O}(N)$, which improves upon the Fast Fourier Transform (FFT) that requires $\mathcal{O}(N \log N)$ operations [123]. In addition, some DWTs are also orthogonal, which means that the inverse transform can be computed equally fast [123]. We will assume the use of an orthogonal DWT in our analysis in Chapter 7, so that we may model the measurement matrix as Gaussian.

The great advantage of wavelets is their ability to locally detect features of interest. For example, the Haar mother wavelet is a step-function, which leads to large wavelet coefficients in regions in which there is a discontinuity in the signal. Other wavelets are designed so that the first few derivatives vanish (*vanishing moments*), and so detect discontinuities in higher derivatives of the signal. Since wavelets essentially work as local discontinuity detectors, signal discontinuities give rise to a chain of large coefficients along a single branch of the tree. For this reason, if a particular wavelet coefficient is large, its parent wavelet coefficient is also likely to be large, which means that the large wavelet coefficients of piecewise smooth signals can be modelled as forming a connected subset of the whole tree which is itself a rooted tree [7].

Let us frame this alternative simplicity model within the notational framework of Section 1.3. We now choose to expand our signal z^* in a wavelet basis $\{\psi_1, \psi_2, \dots, \psi_N\}$. Letting x^* be the coefficients with respect to this basis, we may write this transformation as the matrix equation

$$z^* = \Psi x^*, \tag{6.1}$$

where Ψ is the $N \times N$ matrix with $\psi_1, \psi_2, \ldots, \psi_N$ as its columns. By analogy with the case of simple sparsity, we consider both exact and approximate tree-based models: either x^* is *k-tree sparse* for some $k < N$, meaning that it is supported on a rooted tree of cardinality k; or x^* is *k-tree compressible*, meaning that it is well-approximated by a k-tree sparse vector. See Section 6.2 in which we make the concept of tree-sparsity more precise in the context of a canonical model for dyadic wavelets.

The tree-based model. We continue to make the same assumption concerning the measurement scheme, namely that we obtain the noisy linear measurements $b \in \mathbb{R}^n$ as

$$b = \Phi z^* + e, \tag{6.2}$$

where $e \in \mathbb{R}^n$ is measurement noise. Defining the measurement matrix to be $A := \Phi\Psi \in \mathbb{R}^{n \times N}$, we may combine (6.2) and (6.1) and write

$$b = Ax^* + e. \tag{6.3}$$

We consider two variants of this problem below.

Problem 6.1 (**Tree-sparse recovery from exact measurements**). *Recover exactly a k-tree sparse $x^* \in \mathbb{R}^N$ from the noiseless measurements $b = Ax^* \in \mathbb{R}^n$, where $2k \leq n \leq N$.*

Problem 6.2 (**Tree-compressible recovery from noisy measurements**). *Recover an approximation to a k-tree compressible $x^* \in \mathbb{R}^N$ from the noisy measurements $b = Ax^* + e \in \mathbb{R}^n$, where $2k \leq n \leq N$.*

Writing $\mathcal{T}_k$ for the set of supports which form a rooted tree of cardinality k, we can frame signal recovery as the optimization problem (1.30). See Section 6.2 for a more precise definition of $\mathcal{T}_k$ in the context of dyadic wavelet trees. The generic family of ITP algorithms, which solves (1.30) by gradient projection, has already been described in Section 1.9, but we now summarize it more formally in Algorithm 6.1 below. In Section 1.9, we also introduced the notation $\mathcal{P}_k(\cdot)$ for the projection onto the set of vectors with support in $\mathcal{T}_k$, which we now define more formally.

$$\mathcal{P}_k(x) := \underset{\{z \in \mathbb{R}^N : \mathrm{supp}(z) \in \mathcal{T}_k\}}{\arg\min} \|z - x\|. \tag{6.4}$$

In Section 6.2, we show that this projection is well-defined in the context of a canonical model for dyadic wavelets. For the moment, we will simply assume that $\mathcal{P}_k(\cdot)$ is well-defined. The next lemma establishes a property of $\mathcal{P}_k(\cdot)$ which is shared with the hard threshold operator $\mathcal{H}_k(\cdot)$, namely that it preserves the value of selected coefficients.

Lemma 6.1. *Let $\mathcal{P}_k(\cdot)$ be defined as in (6.4). Then*

$$\{\mathcal{P}_k(x)\}_i := \begin{cases} x_i & i \in \Gamma \\ 0 & i \notin \Gamma \end{cases} \quad \textit{where} \quad \Gamma := \underset{\Gamma \in \mathcal{T}_k}{\arg\max} \, \|x_\Gamma\| \, . \tag{6.5}$$

Proof: We argue similarly to in the proof of Lemma 1.10. Let $\mathrm{supp}(z) = \Gamma \in \mathcal{T}_k$. Then

$$\|z - x\|^2 = \|(z-x)_\Gamma\|^2 + \|x_{\Gamma^C}\|^2,$$

and $\|(z-x)_\Gamma\|^2$ is minimized by setting $z_i = x_i$ for all $i \in \Gamma$. Meanwhile, $\|x_{\Gamma^C}\|$ is minimized by choosing Γ to be the set in $\mathcal{T}_k$ which maximizes $\|x_\Gamma\|$, and the result now follows. □

Algorithm 6.1 Generic ITP [24, 7]

Inputs: A, b, k.
Initialize $x^0 = 0, \ m = 0$.
While some termination criterion is not satisfied, do:

1. Compute a stepsize α^m.

2. Compute $x^{m+1} := \mathcal{P}_k\left\{x^m + \alpha^m A^T(b - Ax^m)\right\}$,
where $\mathcal{P}_k(\cdot)$ is defined in (6.4).

3. Set $m := m + 1$.

End; output $\hat{x} = x^m$.

The comments concerning practical termination criteria for IHT algorithms given in Section 2.1 also apply for ITP algorithms. For Algorithm 2.1 to be well-defined, it remains to define a stepsize scheme $\{\alpha^m\}$. We will consider the same two options as for IHT algorithms, which give rise to the ITP (*constant* stepsize) and NITP (*normalised* stepsize) variants respectively.

Algorithm 6.2 ITP [24, 7]

Given some $\alpha > 0$, on **step 1** of each iteration $m \geq 0$ of generic ITP, set

$$\alpha^m := \alpha. \tag{6.6}$$

The comments concerning the choice of the shrinkage parameter κ for NIHT given in Section 2.1 also applies to NITP. We will make use of a refinement to the notion of the RIP in the tree-based context, which we next define.

Definition 6.2 (**Tree-based RIP [24, 7]**). *For a given matrix A, define TL_s and TU_s, the lower and upper tree-based RIP constants of order s, to be, respectively,*

$$TL_s := 1 - \min_{\emptyset \neq supp(y) \subseteq \Gamma \in \mathcal{T}_s} \frac{\|Ay\|^2}{\|y\|^2} \quad \textit{and} \quad TU_s := \max_{\emptyset \neq supp(y) \subseteq \Gamma \in \mathcal{T}_s} \frac{\|Ay\|^2}{\|y\|^2} - 1. \tag{6.8}$$

Since tree-based RIP considers only certain support sets of cardinality k, it represents a tightening of the standard notion of the RIP, so that, for a given measurement matrix, $TL_k \leq L_k$

Algorithm 6.3 NITP

Given some $c \in (0,1)$ and $\kappa > 1/(1-c)$, on **step 1** of each iteration $m \geq 0$ of generic ITP, do:

1.1. Exact linesearch.

(a) Set $\Gamma^m := \text{supp}(x^m)$.

(b) Compute

$$\alpha^m := \frac{\|A_{\Gamma^m}^T(b - Ax^m)\|^2}{\|A_{\Gamma^m}A_{\Gamma^m}^T(b - Ax^m)\|^2}. \tag{6.7}$$

(c) Let $\tilde{x}^{m+1} := \mathcal{P}_k\left\{x^m + \alpha^m A^T(b - Ax^m)\right\}$.

1.2. Backtracking. If $\text{supp}(\tilde{x}^{m+1}) = \text{supp}(x^m)$, end; output α^m.
Else, while $\alpha^m \geq (1-c)\frac{\|\tilde{x}^{m+1} - x^m\|^2}{\|A(\tilde{x}^{m+1} - x^m)\|^2}$, do:

(a) $\alpha^m := \alpha^m/(\kappa(1-c))$.

(b) $\tilde{x}^{m+1} := \mathcal{P}_k\left\{x^m + \alpha^m A^T(b - Ax^m)\right\}$.

End; output α^m.

and $TU_k \leq U_k$ for all k.

By analogy with (2.3), we have the following bound on the NITP stepsize.

Lemma 6.3 (**NITP stepsize bounds**). *Let α^m be chosen according to Algorithm 6.1. Then*

$$\frac{1}{\kappa(1 + TU_{2k})} \leq \alpha^m \leq \frac{1}{1 - TL_k}. \tag{6.9}$$

Proof: If (6.7) is accepted, then $\alpha^m \leq 1/(1 - TL_k)$ by (6.8). On the other hand, if (6.7) is rejected, the backtracking phase can only reduce the stepsize further, which proves the upper bound in (6.9). To prove the lower bound, we also distinguish two cases. If (6.7) is accepted, then $\alpha^m \geq 1/(1 + TU_k)$ by (6.8). Since $\kappa > 1$, and since $TU_{2k} \geq TU_k$ by the nonincreasing property of tree-based RIP constants, the lower bound in (6.9) holds in this case. On the other hand, if (6.7) is rejected, the penultimate stepsize calculated in the backtracking phase must also have been rejected. Writing $\tilde{\alpha}^m$ for the penultimate stepsize, since $\tilde{\alpha}^m$ was rejected, we have

$$\tilde{\alpha}^m \geq (1-c)\frac{\|\tilde{x}^{m+1} - x^m\|^2}{\|A(\tilde{x}^{m+1} - x^m)\|^2} \geq \frac{1-c}{1 + TU_{2k}}, \tag{6.10}$$

where the last step follows from (6.8). But $\alpha^m = \tilde{\alpha}^m/[\kappa(1-c)]$, which combines with (6.10) to give the lower bound in (6.9) in this case also. □

It follows that, for both ITP and NITP, there exist lower and upper bounds on the stepsize, $\underline{\alpha} > 0$ and $\overline{\alpha} > 0$ respectively, such that $\underline{\alpha} \leq \alpha^m \leq \overline{\alpha}$ for all $m \geq 0$, by (6.9) in the case of NITP, and trivially in the case of ITP.

6.2 An algorithm for Exact Tree Projection (ETP)

Worst-case recovery guarantees were obtained for ITP with unit stepsize in [7] in terms of tree-based RIP. These results assume that the exact projection $\mathcal{P}_k$ onto the set of all rooted trees of cardinality k is performed in each iteration of ITP. In Section 6.3, we will perform our own recovery analysis of ITP algorithms, in which we will also make the same assumption that the projection is computed exactly. Since theoretical results assume an exact projection, a crucial question is whether it is in fact feasible to perform such a projection. In this section, we first survey the algorithms that have already been proposed in this regard, and argue that none of them is guaranteed to return an exact tree projection for a given sparsity k (Section 6.2.1). In Section 6.2.2, we propose a dynamic programming (DP) algorithm which does give such a guarantee. In Section 6.2.3, we obtain a complexity bound for this algorithm, demonstrating that it has low-order polynomial complexity (in fact $\mathcal{O}(Nk)$ for binary trees). We frame the discussion in terms of standard D-dimensional Cartesian-product dyadic wavelets, which have a particular canonical 2^D-ary tree structure. The DP algorithm described could also be applied, however, to any tree structure, so could equally be used for other tree-based representations such as shearlets [118] and curvelets [33]. We first establish some notation.

A canonical tree structure. We suppose a tree structure in which each coefficient has a maximum of d children, where d is the tree order. Assuming the dyadic tree structure of DWTs, we suppose that $d = 2^D$, where the transform dimension D is some fixed positive integer, so that d is a fixed integer with $d \geq 2$. Constrained by the d-ary tree structure, we will require that $N = d^J$ for some $J \geq 2$, where J is the number of levels in the tree structure. We assign a numbering to the nodes of the tree starting with the root node and proceeding from coarse to fine levels. We consider the following tree structure on $\{i : 1 \leq i \leq N\}$, which would cover any dyadic DWT. Let the root node $i = 1$ be at level 0, and have $d - 1$ children at level 1, namely $\{2, \ldots, d\}$. Let each of the remaining nodes except those in level J (the finest level), that is $\{i : 2 \leq i \leq d^{J-1} = N/d\}$, each have d children, namely $\{d(i-1)+1, \ldots, di\}$ respectively. Let the nodes in the finest level J, that is $\{i : i > d^{J-1} = N/d\}$, each have no children. It follows that every node except the root node has precisely one parent, and we have a d-ary tree structure. Figure 6.2 illustrate this tree structure and numbering scheme for the case of $d = 2$ (binary tree) and $J = 4$.

We may now define a rooted tree of cardinality k to be some subtree consisting of k nodes of the full tree just described, subject to the restrictions that it must include the root node, and if any node other than the root node is included then its parent node must also be included. Now consider a vector $y \in \mathbb{R}^N$, whose coefficient indices have the tree structure just outlined. We are interested in finding the projection $\mathcal{P}_k(y)$ of a vector $y \in \mathbb{R}^N$ onto $\mathcal{T}_k$, the set of rooted trees of cardinality k, where $\mathcal{P}_k(\cdot)$ is defined in (6.4). Provided some prior ordering is used to

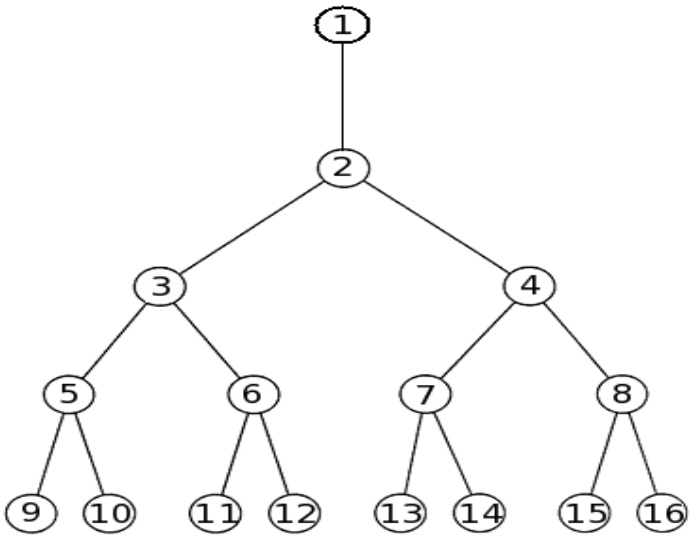


Figure 6.1: An illustration of the canonical dyadic tree structure considered in this section, for the case of $d = 2$ (binary tree) and $J = 4$.

distinguish between vectors which have precisely the same approximation error, having made $\mathcal{T}_k$ concrete, it is now clear that $\mathcal{P}_k(\cdot)$ is well-defined.

An Integer Programming formulation. The problem of projecting a vector $y \in \mathbb{R}^N$ onto the set of rooted d-ary trees may be formulated as the following Integer Program (IP).

$$\max_{\tau \in \mathbb{Z}^N} \sum_{i=1}^{N} y_i^2 \tau_i \quad \text{subject to} \quad \begin{array}{l} \tau \geq 0 \\ \{\tau_i\} \text{ tree-nonincreasing} \\ \sum \tau_i = k \\ \tau_1 = 1. \end{array} \tag{6.11}$$

Denoting the optimal solution of (6.11) by τ^*, the coefficients of the best tree approximation are then $y_i \tau_i^*$. Here tree-nonincreasing means that the coefficients do not increase along the branches, a condition which may be translated into a series of linear constraints. Note that the assumption of integrality together with the constraints in fact forces $\tau_i \in \{0, 1\}$ for each i, so that τ may be viewed as a mask for the nonzero coefficients. To see that (6.11) leads to an exact tree projection, note that the feasible set of (6.11) is precisely the set of vectors which are supported on a rooted tree of cardinality k and whose nonzero coefficients are equal to one. Meanwhile, the objective function of (6.11) is maximized by choosing Γ, the support of τ, so as to maximize $\|y_\Gamma\|^2$, and it now follows from Lemma 6.1 that $y_i \tau_i^* = \{\mathcal{P}_k(y)\}_i$ for all i.

6.2.1 Tree projection algorithms: a brief survey

In [5], Baraniuk presents a survey of three possible algorithms for tree projection: greedy tree approximation, the Condensing Sort and Select Algorithm (CSSA) [8] and a Lagrangian relaxation DP approach due to Donoho [63, 69]. We here present a brief overview, referring the reader to [5] for more details.

Greedy tree approximation. This simple approach involves applying a hard threshold of order $\tilde{k}$ to y for some $\tilde{k} \leq k$, and then forming the smallest rooted tree containing all these $\tilde{k}$ coefficients. One then increases $\tilde{k}$ steadily until the size of the resulting tree is approximately equal to k. While this simple algorithm may yield a good approximation in some circumstances [5], in general there is no reason to expect that it would give anything other than an approximate tree projection.

Condensing Sort and Select Algorithm (CSSA). This algorithm, proposed originally by Baraniuk and Jones in the slightly different context of optimal kernel design [8], solves the following Linear Programming (LP) relaxation of (6.11), in which one dispenses with the assumption of integrality while retaining all other constraints:

$$\max_{\tau \in \mathbb{R}^N} \sum_{i=1}^{N} y_i^2 \tau_i \quad \text{subject to} \quad \begin{array}{l} \tau \geq 0 \\ \{\tau_i\} \text{ tree-nonincreasing} \\ \sum \tau_i = k \\ \tau_1 = 1. \end{array} \tag{6.12}$$

The authors observe that the two solutions are sometimes equal, but not always: a value of $\tau_i = 1$ may be assigned to all but a few coefficients, which are each assigned a value $\bar{\tau}$ for some $0 < \bar{\tau} < 1$. In this case, the result is a tree of size strictly greater than k. Moreover, there appears to be no straightforward method for 'adjusting' the solution *a posteriori* to obtain the optimal k-sparse tree projection. Our own numerical experimentation suggests that the CSSA often identifies an exact tree projection for an approximate value of the sparsity. However, for a given sparsity k, it is not guaranteed to find that particular tree approximation.

Lagrangian relaxation DP approach. Rather than imposing a hard constraint on the required sparsity of the tree projection, Donoho et al. [63, 69] proposed penalizing sparsity using a Lagrangian relaxation. More precisely, the equality constraint $\sum \tau_i = k$ is removed and instead penalized in the objective, giving the relaxation

$$\max_{\tau \in \mathbb{Z}^N} \sum_{i=1}^{N} y_i^2 \tau_i \; - \lambda \cdot \#\{\tau_i = 1\} \quad \text{subject to} \quad \begin{array}{l} \tau \geq 0 \\ \{\tau_i\} \text{ tree-nonincreasing} \\ \tau_1 = 1. \end{array} \tag{6.13}$$

This reformulation can be solved using fine-to-coarse dynamic programming on the tree [69]. The drawback is the non-obvious relationship between the Lagrange multiplier λ and the required sparsity, which means that at best it can only find tree projections for an approximate value of the required sparsity. Because (6.11) is an IP, it is not guaranteed to share the same optimal solution as (6.13). Indeed, our experimentation suggests that the Lagrangian approach finds, like CSSA, an exact tree projection for an approximate value of the sparsity, but not necessarily for a specific required sparsity.

Summary. Greedy tree approximation generally only gives approximate tree projections, while both CSSA and the Lagrangian approach give exact tree projections for an approximation to the sparsity. The underlying reason in the latter two cases is that both approaches rely upon solving a relaxation of the original problem which does not necessarily share the same solution. The solution to the IP (6.11) could be obtained by means of an exhaustive search, which has combinatorial complexity. However, the question of how to calculate exact tree projections in polynomial time remains unanswered. We next propose a DP algorithm of our own which achieves this goal.

6.2.2 An algorithm for exact tree projection

Our algorithm (Algorithm 6.4) falls into the broad category of dynamic programming (DP) algorithms which optimize on directed graphs by utilizing a principle of optimality, namely that optimal solutions at a given node may be determined entirely from optimal solutions at 'preceding' nodes [13]. Our algorithm makes two passes through the tree: firstly, a fine-to-coarse pass finds optimal subtrees at each node for all $\tilde{k} \leq k$. Secondly, a coarse-to-fine pass tracks back to identify the optimal choices at each stage.

In Algorithm 6.4, we use the notation $Q(l,d)$ for the set of all partitions of l into d positive integers, viewed as vectors in $\mathbb{Z}_+^d$, where we must consider different orderings to be distinct. More formally

$$Q(l,d) = \left\{ \mathbf{G} \in \mathbb{Z}_+^d \ : \ \begin{array}{ll} 0 \leq G_r \leq l & \forall r \\ \sum_{r=1}^d G_r = l & \end{array} \right\}.$$

The algorithm starts at the finest level and moves up the tree, finding optimal subtrees rooted at each node, each decision requiring only the information already obtained within that particular subtree. Optimal subtrees are found at each node for all cardinalities which could possibly contribute to a rooted tree of cardinality k, which imposes two restrictions. Firstly, by summing the appropriate geometric series, the maximum cardinality of a subtree rooted at a node in level j is $\frac{d^{J+1-j}-1}{d-1}$. Secondly, any subtree rooted at a node in level j of cardinality greater than $k-j$ would necessarily contribute to a rooted tree of cardinality greater than k. Therefore, for nodes at level j, we need only find optimal subtrees for cardinalities up to $\min\left(\frac{d^{J+1-j}-1}{d-1}, k-j\right)$.

Algorithm 6.4 Exact tree projection (ETP)

Inputs: $\begin{cases} y \in \mathbb{R}^N & (N = d^J; J \geq 2; d \geq 2) \\ k \in \mathbb{N} & (k \leq N) \end{cases}$

Initializations:

for $i = d^{J-1} + 1 : d^J$ **do**
$F(i, 0) = 0$
$F(i, 1) = y_i^2$
end for

Find all optimal subtrees:

for $j = J - 1 : -1 : 1$ **do**
for $i = d^{j-1} + 1 : d^j$ **do**
$F(i, 0) = 0$
$\mathbf{G}(i, 0) = \mathbf{0}$
$F(i, 1) = y_i^2$
$\mathbf{G}(i, 1) = \mathbf{0}$
for $l = 2 : \min(d^{J+1-j} - 1, k - j)$ **do**

$$F(i, l) = y_i^2 + \max_{\mathbf{G} \in Q(l-1,d)} \sum_{r=1}^{d} F\left[d(i-1) + r, G_r\right]$$

$$\mathbf{G}(i, l) = \arg \max_{\mathbf{G} \in Q(l-1,d)} \sum_{r=1}^{d} F\left[d(i-1) + r, G_r\right]$$

end for
end for
end for

$$F(1, k) = y_1^2 + \max_{\mathbf{G} \in Q(k-1,d)} \sum_{r=2}^{d} F(r, G_r)$$

$$\mathbf{G}(1, k) = y_1^2 + \arg \max_{\mathbf{G} \in Q(k-1,d)} \sum_{r=2}^{d} F(r, G_r)$$

Backtrack to identify optimal tree of size k:

$\Omega = [1]$
$\Gamma_1 = k$
for $j = 1 : J - 1$ **do**
for $i = d^{j-1} + 1 : d^j$ **do**
if $i \in \Omega$ **then**
for $r = 1 : d$ **do**
if $G_r(i, \Gamma_i) > 0$ **then**
$\Omega = \Omega \cup \{d(i-1) + r\}$
$\Gamma_{d(i-1)+r} = G_r(i, \Gamma_i)$
end if
end for
end if
end for
end for

Outputs: $\begin{cases} \hat{y}_\Omega = y_\Omega \\ \hat{y}_{\Omega^C} = 0 \end{cases}$

End

$F(i,l)$ denotes the total energy of the optimal subtree rooted at node i, of size l. $\mathbf{G}(i,l) \in \mathbb{R}^d$ is a vector of positive integers, giving the number of nodes contributing to this optimal solution from the subtree rooted at each of the children of node i. Storing this precedence information is a crucial step in many DP algorithms, allowing the optimal path to be determined by backtracking through the graph at the end. We then proceed recursively and eventually determine $F(1,k)$, the energy of the optimal tree of size k. Finally, we use the precedence information to trace the optimal solution back along the branches.

That the algorithm does indeed calculate optimal subtrees at each node follows since we have a principle of optimality, namely that the optimal subtree of size l at a given node i is found by searching through all possible combinations of optimal subtrees of the children of i, such that the sum of the sizes of these subtrees is equal to $l-1$. More formally,

$$F(i,l) = y_i^2 + \max_{\mathbf{G} \in Q(l-1,d)} \sum_{r=1}^{d} F\left[d(i-1)+r, G_r\right], \tag{6.14}$$

which is precisely how optimal subtrees are calculated in the algorithm.

The algorithm actually does more than is asked for: by the time the root node is reached at the end of the first pass, enough information has been obtained to determine optimal trees for all $\tilde{k} \leq k$. An important question to ask is how its complexity compares with other approximate methods, which we address next.

6.2.3 Complexity analysis

The first observation to make is that the second (backtracking) pass through the tree simply entails a fixed number of operations per node of the tree, and therefore has complexity $\mathcal{O}(N)$. We will see that the leading order complexity is in fact determined by the first pass. The complexity of the first pass is essentially determined by the aggregated cardinality of the candidate sets $Q(l,d)$ over which we maximize for each node, since for each we perform a fixed number of operations ($d-1$ additions and a comparison). Fortunately, $|Q(l,d)|$ is nothing other than the number of weak compositions (ordered partitions) of l into d nonnegative integer addends [108], which has a simple closed-form expression.

Lemma 6.4 (**Weak compositions [142, p.18]**). *Let l and d be positive integers. Then*

$$|Q(l,d)| = \binom{l+d-1}{d-1}. \tag{6.15}$$

The next lemma gives a simple bound on $|Q(l,d)|$.

Lemma 6.5 (**Weak composition bound**).

$$|Q(l,d)| \leq (l+1)^{d-1}.$$

Proof: Starting from (6.15), we deduce

$$\begin{aligned} |Q(l,d)| &= \frac{(l+d-1)!}{l!(d-1)!} = \frac{(l+1)(l+2)\dots(l+d-1)}{1\cdot 2\dots(d-1)} \\ &= \prod_{i=1}^{d-1}\left(\frac{l+i}{i}\right) = \prod_{i=1}^{d-1}\left(\frac{l}{i}+1\right) \leq (l+1)^{d-1}. \end{aligned}$$ □

We now proceed to our complexity result.

Theorem 6.6 (**Complexity of ETP**). *The ETP algorithm has complexity* $\mathcal{O}(Nk^{d-1})$.

Proof: We bound the aggregrated cardinality of all candidate sets $Q(l,d)$ in the first pass of the algorithm. In each level j with $1 \leq j \leq J-1$, there are $(d-1)\cdot d^{j-1}$ nodes. Writing $l(j)$ for the number of evaluations of F that are required for a node in level j, each evaluation of F involves the optimization over a candidate set $Q(l(j)-1,d)$, the cardinality of each of which is bounded, by means of Lemma 6.5, by $|Q(l(j)-1,d)| \leq [l(j)]^{d-1}$. It follows that the aggregated cardinality $\mathcal{Q}$ of all $Q(l,d)$ sets for levels 1 to $J-1$ is given by

$$\mathcal{Q} = \sum_{j=1}^{J-1}\left\{(d-1)d^{j-1}\cdot l(j)\,[l(j)]^{d-1}\right\}, \tag{6.16}$$

while, in addition, Algorithm 6.4 guarantees that

$$l(j) \leq \min\left(\frac{d^{J+1-j}-1}{d-1}, k-j\right) \leq \min(d^{J+1-j}, k). \tag{6.17}$$

To determine which of d^{J+1-j} or k gives a tighter bound, let $1 \leq p \leq J-1$ be the unique positive integer such that

$$d^{p-1} \leq k \leq d^p. \tag{6.18}$$

First let us assume $p > 2$. We have

$$k \leq d^{J+1-j} \iff d^p \leq d^{J+1-j} \iff p \leq J+1-j \iff j \leq J+1-p,$$

which we may combine with (6.16) and (6.17) to deduce

$$\begin{aligned} \mathcal{Q} &\leq \sum_{j=1}^{J+1-p}(d-1)d^{j-1}\cdot k\cdot k^{d-1} + \sum_{J+2-p}^{J-1}(d-1)d^{j-1}\cdot d^{J+1-j}\cdot (d^{J+1-j})^{d-1} \qquad (6.19) \\ &= (d-1)\left\{k^d\sum_{j=1}^{J+1-p}d^{j-1} + \sum_{J+2-p}^{J-1}d^{dJ-(d-1)(j-1)}\right\} \\ &= (d-1)\left\{k^d\left(\frac{d^{J+1-p}-1}{d-1}\right) + d^{J+(d-1)(p-1)}\left[\frac{1-\left(\frac{1}{d^{d-1}}\right)^{p-1}}{1-\left(\frac{1}{d^{d-1}}\right)}\right]\right\}. \qquad (6.20) \end{aligned}$$

Since

$$(d-1)\left[\frac{1-\left(\frac{1}{d^{d-1}}\right)^{p-1}}{1-\left(\frac{1}{d^{d-1}}\right)}\right] \leq \left(\frac{d-1}{1-\frac{1}{d}}\right) = d,$$

and since $d^{J+1-p} - 1 \leq d^{J+1-p}$ holds trivially, we can further deduce from (6.20) that

$$\mathcal{Q} \leq k^d \cdot d^{J+1-p} + d^{J+(d-1)(p-1)+1},$$

to which we can make the substitution $N = d^J$ and apply the bounds (6.18) to conclude that

$$\mathcal{Q} \leq k^d \cdot \frac{dN}{k} + dNk^{d-1} = 2dNk^{d-1} = \mathcal{O}(Nk^{d-1}).$$

On the other hand, if $p \leq 2$, then $k \leq d^2 \leq d^{J+1-j}$ for all $j \leq J-1$, and so the second summation in (6.19) is empty. We may then follow the same argument to bound the first summation, obtaining $\mathcal{Q} = \mathcal{O}(Nk^{d-1})$ in this case also. Finally, the observations at the start of this section, that the aggregated cardinality bounded here determines the order of complexity of the forward pass through the tree, and that the backward pass through the tree is $\mathcal{O}(N)$, now lead us to the desired conclusion. □

The result assumes d to be fixed, and we should draw attention to the dependence upon d, and in particular the linear growth of the exponent in d. For a binary tree ($d = 2$), for example, the ETP algorithm is $\mathcal{O}(Nk)$; whereas for quad-trees ($d = 4$) we obtain $\mathcal{O}(Nk^3)$.

We may compare the complexity of ETP with that of the other approximate tree projection methods: greedy tree approximation and CSSA are both $\mathcal{O}(N \log N)$ while the Lagrangian approach is $\mathcal{O}(N)$. Provided $\log N \ll k$, we can see that the price we pay for guaranteeing an exact tree projection is an increased order of complexity, with the change being especially marked for trees of higher order.

In this thesis, we require an exact tree projection for the theoretical results to hold. However, in practice, one of the approximate solutions to the tree projection problem may be considered quite satisfactory, and in fact to be preferred due to it likely being faster. In particular, the CSSA appears to be a good practical choice when an accurate approximation to a tree projection for a given sparsity is required. This having been said, our ETP algorithm is the only one guaranteed to perform the task of tree projection exactly, thereby clarifying an issue of some ambiguity in the literature. In the recovery analysis that follows in the next section, we assume that the tree projection $\mathcal{P}_k$ is computed exactly.

6.3 A new recovery analysis of ITP algorithms

In this section, we define and analyse the notion of an $\underline{\alpha}$-stable point for ITP algorithms (Section 6.3.1), and we prove conditions guaranteeing the convergence of both ITP and NITP

variants to a stable point (Section 6.3.2). We follow a similar approach to Chapter 3, proving our results in the most general setting of noisy measurements and k-compressible signals (Problem 6.2). We next adapt some of the notation introduced for the standard sparsity model at the start of Chapter 3 to the tree-based setting.

Notation. We now redefine x_k^* to be the best k-tree sparse approximation to x^*, so that

$$x_k^* := \underset{\{z \in \mathbb{R}^N : \mathrm{supp}(z) \in \mathcal{T}_k\}}{\arg\min} \|z - x^*\|. \tag{6.21}$$

We define $\Lambda \in \mathcal{T}_k$ to be the support of the best k-tree sparse approximation to x^*, namely

$$\Lambda := \mathrm{supp}(x_k^*). \tag{6.22}$$

We will assume that x^* itself is at least k-tree sparse (it may not be tree sparse at all), so that x_k^* is exactly k-tree sparse and Λ has cardinality $|\Lambda| = k$. The combined noise due both to compression error and noise in the measurements, E, may be written as

$$E := A(x^* - x_k^*) + e = A_{\Lambda^C} x_{\Lambda^C}^* + e. \tag{6.23}$$

Given some index set $\Gamma \in \mathcal{T}_k$, we also define

$$\tilde{e}_\Gamma := e + A_{(\Gamma \cup \Lambda)^C} x_{(\Gamma \cup \Lambda)^C}^*. \tag{6.24}$$

6.3.1 Stable point analysis

As in Section 3.1, we begin our considerations with ITP, the constant stepsize variant. Recalling the algorithm summary in Algorithm 6.1, let us write

$$\phi(x) := \mathcal{P}_k\{x + \alpha A^T(b - Ax)\}, \tag{6.25}$$

so that the ITP iteration can be expressed as $x^{m+1} = \phi(x^m)$. Then a *fixed point* of ITP is defined as any $\bar{x} \in \mathbb{R}^N$ such that $\phi(\bar{x}) = \bar{x}$.

We have the following necessary conditions for some $\bar{x}$ to be a fixed point of ITP.

Lemma 6.7 (ITP fixed point necessary condition). *Suppose $\bar{x} \in \mathbb{R}^N$ is a fixed point of ITP with stepsize $\alpha > 0$. Then there exist some set Γ with $\mathrm{supp}(\bar{x}) \subseteq \Gamma$ and $|\Gamma| = k$ such that the following two conditions hold.*

$$\left\{A^T(b - A\bar{x})\right\}_\Gamma = 0; \tag{6.26}$$

$$\|\bar{x}_{\Gamma \setminus \Omega}\| \geq \underline{\alpha} \|A_{\Omega \setminus \Gamma}^T (b - A\bar{x})\| \quad \forall\, \Omega \in \mathcal{T}_k. \tag{6.27}$$

Proof: We can follow the argument of Lemma 3.1 to deduce (6.26), since all that is assumed in the argument about the projection $\mathcal{H}_k(\cdot)$ is that it preserves the value of selected coefficients, which is also the case for $\mathcal{P}_k(\cdot)$ by (6.5). To prove (6.27), let us suppose that $\bar{x}$ is a fixed point of ITP with stepsize $\alpha > 0$, and that $\text{supp}(\bar{x}) \subseteq \Gamma$ with $|\Gamma| = k$, and let us write

$$\bar{a} := \bar{x} + \alpha A^T(b - A\bar{x}) \tag{6.28}$$

so that $\phi(\bar{x}) = \mathcal{P}_k(\bar{a})$. Then it follows from (6.5) that, for any $\Omega \in \mathcal{T}_k$,

$$\|\bar{a}_\Gamma\|^2 \geq \|\bar{a}_{\text{supp}(\bar{x})}\|^2 \geq \|\bar{a}_\Omega\|^2,$$

which may be simplified to give

$$\|\bar{a}_{\Gamma\setminus\Omega}\|^2 \geq \|\bar{a}_{\Omega\setminus\Gamma}\|^2. \tag{6.29}$$

Now $\bar{a}_{\Omega\setminus\Gamma} = \alpha A^T_{\Omega\setminus\Gamma}(b - A\bar{x})$ using (6.28), while (6.26) implies that $\bar{a}_{\Gamma\setminus\Omega} = \bar{x}_{\Gamma\setminus\Omega}$, and these observations may be combined with (6.29) to deduce (6.27). □

Like the corresponding conditions for IHT, the conditions (6.26) and (6.27) have a simple intuitive interpretation. If a further iteration of ITP is applied at $\bar{x}$, there is no change in the support set, and it follows from (6.5) that we then require

$$\| \left\{\bar{x} + \alpha A^T(b - A\bar{x})\right\}_\Gamma \|^2 \geq \| \left\{\bar{x} + \alpha A^T(b - A\bar{x})\right\}_\Omega \|^2,$$

which simple manipulation shows to be equivalent to (6.27). Meanwhile, the coefficients on the support of $\bar{x}$ remain unchanged, for which we require the gradient on the support of $\bar{x}$ be zero, namely (6.26) must hold.

If the inequality in (6.27) is replaced by a strict inequality, then it can be shown that the conditions in Lemma 6.7 are also sufficient for $\bar{x}$ to be a fixed point of ITP.

Recall from Section 3.1 that the initial stepsize choice for NIHT is not well-defined, forcing the introduction of the concept of an $\underline{\alpha}$-stable point. Since the NITP stepsize initialization (6.7) is the same as for NIHT, we must also introduce the equivalent concept in the tree-based setting.

Definition 6.8 (Stable points of generic ITP). *Given $\underline{\alpha} > 0$ and an index set $\Gamma \in \mathcal{T}_k$, we say $\bar{x} \in \mathbb{R}^N$ is an $\underline{\alpha}$-stable point of generic ITP on Γ if supp$(\bar{x}) \subseteq \Gamma$ and*

$$\left\{A^T(b - A\bar{x})\right\}_\Gamma = 0 \quad \textit{and} \tag{6.30}$$

$$\|\bar{x}_{\Gamma\setminus\Omega}\| \geq \underline{\alpha}\|A^T_{\Omega\setminus\Gamma}(b - A\bar{x})\| \quad \forall\, \Omega \in \mathcal{T}_k. \tag{6.31}$$

For brevity's sake, we will often drop the 'of generic ITP' label, and at times we will also

drop the reference to the support set Γ. We can also view an $\underline{\alpha}$-stable point as a generalization of the notion of a fixed point for ITP. In particular, note that, by Lemma 6.7, any fixed point of ITP with stepsize $\alpha > 0$ is an α-stable point of generic ITP.

An $\underline{\alpha}$-stable point of generic ITP may also be characterized as a minimum-norm solution on some k-subspace.

Lemma 6.9. *Suppose Assumption 1 holds and suppose $\bar{x}$ is an $\underline{\alpha}$-stable point of generic ITP on Γ for some $\underline{\alpha} > 0$. Then $\bar{x}_\Gamma = A_\Gamma^\dagger b$, where the Moore-Penrose pseudoinverse $A_\Gamma^\dagger$ is defined in (2.7).*

Proof: It follows from (6.30) that $A_\Gamma^T(b - A_\Gamma \bar{x}_\Gamma) = 0$ where $\mathrm{supp}(\bar{x}) \subseteq \Gamma$ and $|\Gamma| = k$. Under Assumption 1, the pseudoinverse $A_\Gamma^\dagger$ in (2.7) is well-defined and we may rearrange to give $\bar{x}_\Gamma = A_\Gamma^\dagger b$. □

The next result, which translates Theorem 3.4 to the tree-based setting, gives a necessary condition for a stable point on a given support in terms of only x^*, A and e and their restrictions to certain support sets.

Theorem 6.10 (**Stable point condition**). *Consider Problem 6.2. Suppose Assumption 1 holds and suppose there exists an $\underline{\alpha}$-stable point on some Γ such that $\Gamma \neq \Lambda$. Then*

$$\left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| + \left\|x^*_{\Gamma\setminus\Lambda}\right\| + \left\|A_\Gamma^\dagger \tilde{e}_\Gamma\right\| \geq \underline{\alpha}\left\{\left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| - \left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)\tilde{e}_\Gamma\right\|\right\}, \tag{6.32}$$

where Λ is defined in (6.22) and $\tilde{e}_\Gamma$ is defined in (6.24).

Proof: Supposing that $\bar{x}$ is an $\underline{\alpha}$-stable point on Γ, choosing $\Omega := \Lambda$ in (6.31) yields

$$\|\bar{x}_{\Gamma\setminus\Lambda}\|^2 \geq \underline{\alpha}^2 \|A^T_{\Lambda\setminus\Gamma}(b - A\bar{x})\|^2.$$

We may now follow the argument of Theorem 3.4 to deduce (6.32). □

In the case of Problem 6.1, Theorem 6.10 simplifies to the following corollary.

Corollary 6.11 (**Noiseless case**). *Consider Problem 6.1. Suppose Assumption 1 holds and suppose there exists an $\underline{\alpha}$-stable point on some Γ such that $\Gamma \neq \Lambda$. Then*

$$\left\|A_\Gamma^\dagger A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\| \geq \underline{\alpha}\left\|A^T_{\Lambda\setminus\Gamma}(I - A_\Gamma A_\Gamma^\dagger)A_{\Lambda\setminus\Gamma} x^*_{\Lambda\setminus\Gamma}\right\|, \tag{6.33}$$

where Λ is defined in (6.22).

Proof: Since $x^*_{\Lambda^C} = 0$ and $e = 0$, we have $\tilde{e}_\Gamma = 0$ and $x^*_{\Gamma\setminus\Lambda} = 0$ for all Γ, and making both these substitutions in (6.32) yields the required result. □

6.3.2 Convergence analysis

In this section, we derive convergence conditions for both ITP and NITP. Some useful properties of the iterates of ITP are given in the next lemma, which is the counterpart of Lemma 3.7.

Lemma 6.12. *The iterates of generic ITP satisfy (3.25) and (3.26) for all $m \geq 0$, where $\Psi(\cdot)$ is defined in (1.24) and g^m is defined in (3.24).*

Proof: Using (6.4) and (3.24), we may rewrite the generic ITP iteration (1.29) as

$$x^{m+1} = \underset{\{z \in \mathbb{R}^N : \mathrm{supp}(z) \in \mathcal{T}_k\}}{\arg\min} \|z - \{x^m - \alpha^m g^m\}\|^2,$$

from which we may deduce

$$\|x^{m+1} - (x^m - \alpha^m g^m)\|^2 \leq \|x^m - (x^m - \alpha^m g^m)\|^2 = (\alpha^m)^2 \|g^m\|^2,$$

which expands to give

$$\|x^{m+1} - x^m\|^2 + 2\alpha^m (g^m)^T (x^{m+1} - x^m) + (\alpha^m)^2 \|g^m\|^2 \leq (\alpha^m)^2 \|g^m\|^2,$$

and so (3.25) holds. We may follow the argument of Lemma 3.7 to deduce (3.26). □

A sufficient condition for convergence of generic ITP is given next.

Lemma 6.13 (**Sufficient condition for convergence**). *Consider Problem 6.2. Suppose Assumption 1 holds, and suppose the iterates of generic ITP satisfy (3.27) for some $d > 0$ which does not depend upon m, where $\Psi(\cdot)$ is defined in (1.24). Assume that there exist $\overline{\alpha} \geq \underline{\alpha} > 0$ such that (3.28) holds. Then $x^m \to \bar{x}$ as $m \to \infty$, where $\bar{x}$ is an $\underline{\alpha}$-stable point of generic ITP.*

Proof: We may follow the proof of Lemma 3.8 to deduce that $x^m \to \bar{x}$, where $\bar{x}_\Gamma = A_\Gamma^\dagger b$ and $\bar{x}_{\Gamma^C} = 0$, for some Γ such that $|\Gamma| = k$. The proof still holds since all that is assumed about the hard threshold projection $\mathcal{H}_k(\cdot)$ is that it preserves the value of selected coefficients, a property which is also shared by the tree projection $\mathcal{P}_k(\cdot)$ by (6.5). Since $\Gamma = \Gamma^m$ for some $m \geq 0$, it follows that, in the case of ITP, $\Gamma \in \mathcal{T}_k$. Therefore (6.30) holds for $\bar{x}$.
It remains to establish that $\bar{x}$ satisfies (6.31). Defining

$$\Gamma_1 = \{i \in \Gamma \ : \ \bar{x}_i \neq 0\}, \tag{6.34}$$

it follows that $\Gamma_1 \subseteq \Gamma^m$ for all m sufficiently large. It follows from (6.5) that, for any $\Omega \in \mathcal{T}_k$,

$$\|x^{m+1}_{\Gamma^{m+1}}\|^2 \geq \|\{x^m - \alpha^m g^m\}_\Omega\|^2, \ \text{ for all } \ m \geq 0.$$

and therefore, for all m sufficiently large,

$$\|x^{m+1}_{\Gamma_1}\|^2 + \|x^{m+1}_{\Gamma^{m+1}\setminus\Gamma_1}\|^2 \geq \|x^{m+1}_{\Omega\cap\Gamma_1}\|^2 + \|\{x^m - \alpha^m g^m\}_{\Omega\setminus\Gamma_1}\|^2,$$

which cancels to

$$\|x^{m+1}_{\Gamma_1\setminus\Omega}\|^2 + \|x^{m+1}_{\Gamma^{m+1}\setminus\Gamma_1}\|^2 \geq \|\{x^m + \alpha^m g^m\}_{\Omega\setminus\Gamma_1}\|^2. \tag{6.35}$$

Furthermore, it follows from (6.34) that

$$\|x^{m+1}_{\Gamma^{m+1}\setminus\Gamma_1}\|^2 \to 0. \tag{6.36}$$

By (3.28), there exists a convergent subsequence of stepsizes,

$$\alpha^{m_r} \to \tilde{\alpha} \geq \underline{\alpha} \text{ as } r \to \infty \tag{6.37}$$

Passing to the limit in (6.35) on the subsequence m_r for which (6.37) holds, we deduce that $\|\bar{x}_{\Gamma_1\setminus\Omega}\| \geq \underline{\alpha}\|\{A^T(b - A\bar{x})\}_{\Omega\setminus\Gamma_1}\|$, from which it follows trivially that

$$\|\bar{x}_{\Gamma\setminus\Omega}\| \geq \underline{\alpha}\|\{A^T(b - A\bar{x})\}_{\Omega\setminus\Gamma}\|. \tag{6.38}$$

Since (6.38) holds for any $\Omega \in \mathcal{T}_k$, $\bar{x}$ satisfies (6.31), and the result is proved. □

Theorem 6.14 (**ITP convergence**). *Consider Problem 6.2. Suppose that Assumption 1 holds, and suppose that the stepsize in ITP satisfies*

$$\alpha < \frac{1}{1 + TU_{2k}}. \tag{6.39}$$

Then ITP with stepsize α converges to an α-stable point $\bar{x}$ of generic ITP.

Proof: We mimic the proof of Theorem 3.9. Let $m \geq 0$. Since the support size of the change to the iterates $x^{m+1} - x^m$ is at most $2k$, and since a union of two rooted trees is also a rooted tree, (6.8) with $s = 2k$ provides $\|A(x^{m+1} - x^m)\|^2 \leq (1 + TU_{2k})\|x^{m+1} - x^m\|^2$. Using this bound, and (3.25) with the choice (6.39), in (3.26), we obtain

$$\begin{aligned} \Psi(x^{m+1}) - \Psi(x^m) &\leq -\tfrac{1}{2\alpha}\|x^{m+1} - x^m\|^2 + \tfrac{1}{2}(1 + TU_{2k})\|x^{m+1} - x^m\|^2 \\ &= \frac{\alpha(1 + TU_{2k}) - 1}{2\alpha}\|x^{m+1} - x^m\|^2, \end{aligned}$$

which, due to (6.39), implies that (3.27) holds with $d := 2\alpha/[1 - \alpha(1 + TU_{2k})]$. Due to (6.39), (3.28) trivially holds with $\overline{\alpha} = \underline{\alpha} = \alpha$. Thus Lemma 6.13 applies, and the ITP iterates x^m converge to an α-stable point. □

We next obtain a convergence result for NITP. In this case, there is no explicit requirement

for a tree-based RIP condition to be satisfied; however, tree-based RIP this time appears in the choice of $\underline{\alpha}$.

Theorem 6.15 (**NITP convergence**). *Suppose Assumption 1 holds. Then NITP with shrinkage parameter κ converges to a $[\kappa(1+TU_{2k})]^{-1}$-stable point $\bar{x}$ of generic ITP.*

Proof: We mimic the proof of Theorem 3.10. Firstly, we consider the case when the exact linesearch choice is accepted, so that α^m is given by (6.7). Then (3.24) implies $\Gamma^{m+1} = \Gamma^m$, and (1.29) implies

$$x^{m+1}_{\Gamma^m} = x^m_{\Gamma^m} - \alpha^m g^m_{\Gamma^m}. \tag{6.40}$$

Using (6.40), (6.7) becomes

$$\alpha^m = \frac{\|g^m_{\Gamma^m}\|^2}{\|A_{\Gamma^m} g^m_{\Gamma^m}\|^2} = \frac{\|x^{m+1} - x^m\|^2}{\|A(x^{m+1} - x^m)\|^2}. \tag{6.41}$$

Using that $x^{m+1} - x^m$ is supported on Γ^m, expressing $g^m_{\Gamma^m}$ from (6.40) and substituting into (3.26), we deduce that

$$\begin{aligned} \Psi(x^{m+1}) - \Psi(x^m) &= -\frac{1}{\alpha^m}(x^{m+1}_{\Gamma^m} - x^m_{\Gamma^m})^T(x^{m+1}_{\Gamma} - x^m_{\Gamma}) + \frac{1}{2}\|A(x^{m+1} - x^m)\|^2 \\ &= -\frac{1}{\alpha^m}\|x^{m+1} - x^m\|^2 + \frac{1}{2\alpha^m}\|x^{m+1} - x^m\|^2 = -\frac{1}{2\alpha^m}\|x^{m+1} - x^m\|^2, \end{aligned} \tag{6.42}$$

where to obtain the second equality, we also used (6.41). Alternatively, when α^m is computed by shrinkage, we deduce that

$$\|A(x^{m+1} - x^m)\|^2 \leq \frac{1-c}{2\alpha^m}\|x^{m+1} - x^m\|^2.$$

Substituting this and (3.25) into (3.26), we obtain

$$\Psi(x^{m+1}) - \Psi(x^m) \leq -\frac{1}{2\alpha^m}\|x^{m+1} - x^m\|^2 + \frac{1-c}{2\alpha^m}\|x^{m+1} - x^m\|^2 = -\frac{c}{2\alpha^m}\|x^{m+1} - x^m\|^2. \tag{6.43}$$

Thus (6.42), (6.43) and $c \in (0,1)$ imply that, for all $m \geq 0$,

$$\|x^{m+1} - x^m\|^2 \leq \frac{2\alpha^m}{c}[\Psi(x^m) - \Psi(x^{m+1})] \leq \frac{2(1-c)}{c(1-TL_{2k})}[\Psi(x^m) - \Psi(x^{m+1})],$$

due to (6.9). Hence (3.27) holds with $d := 2(1-c)/[c(1-TL_{2k})]$, and so does (3.28) due to (6.9). Lemma 6.13 applies and, together with (6.9), provides the required conclusion. □

The final goal is to translate the results of the present chapter into quantitative asymptotic results for Gaussian measurement matrices, towards which we now direct our attention in Chapter 7.

Chapter 7

Oversampling thresholds for ITP algorithms

In this chapter, we extend the analysis of Chapter 5 to the tree-based setting, and obtain average-case recovery guarantees for ITP algorithms in the case of Gaussian measurement matrices. In Section 7.1, by making use of tree counting results, we obtain special cases of the large deviations bounds of Chapter 4 for the tree-based model. We show that these results can be expressed within a simplified proportional-growth asymptotic, and we also prove bounds on the upper tree-based RIP constant within this simplified framework. Section 7.2 contains the recovery analysis and main results for the tree-based model, which combines the results from Chapter 6 and Section 7.1. In Section 7.2.1, we define our recovery phase transitions and establish that they are well-defined. We prove our oversampling thresholds for ITP in Section 7.2.2, and for NITP in Section 7.2.3. We illustrate our results in Section 7.3, and present a discussion, firstly for the noiseless case of Problem 6.1, and secondly for the general case of Problem 6.2.

7.1 Large deviations results in the tree-based setting

In order to prove quantitative recovery conditions for ITP algorithms, we need to apply the large deviations results in Chapter 4 to the tree-based model. We also need to obtain a bound on the upper tree-based RIP constant for Gaussian matrices by extending the analysis in [21]. We accomplish both of these two tasks in this section. The assumption of a rooted tree structure means that the number of permissible support sets for iterates of the algorithm is much diminished in comparison with the standard sparsity model, which means that union bound arguments can be tightened.

In what follows, consider d to be some fixed integer with $d \geq 2$. We need to count $|\mathcal{T}_k|$, the number of permissible support sets in the d-ary tree-based framework, which is bounded above by $T(k)$, the total number of ordered, rooted d-ary trees of cardinality k. Fortunately, there is a simple answer to this question, which is given in the following lemma.

Lemma 7.1 (**Tree counting result [115, Note 12]**). *The total number of ordered, rooted d-ary trees of cardinality k is*

$$T(k) = \frac{1}{(d-1)k+1}\binom{dk}{k}. \tag{7.1}$$

These numbers are also known in the literature as the Pfaff-Fuss-Catalan numbers and the k-Raney numbers [115, Note 12]. In the special case of $d = 2$, we obtain

$$T(k) = \frac{1}{k+1}\binom{2k}{k},$$

which is the Catalan number [115, Note 6].

A similar result was proved in [7, Proposition 1] for the case of binary trees ($d = 2$), though the result given above represents a generalization to any $d > 2$, and in fact also gives a tightening of the result in [7] in the case where $\log_2(N) > k$. Note also that we have an upper bound on $|\mathcal{T}_k|$ which is independent of N. This is in contrast to the total number of supports, which is $\binom{N}{k}$. However, $|\mathcal{T}_k|$ may not attain this upper bound, for two reasons. Firstly, if additional structure is imposed upon the tree structure (for example, in the dyadic wavelet tree model of Section 6.2, the root node has only $d-1$ children), then some trees are excluded from consideration. Secondly, the number of levels in the tree structure is limited to $J = \log_d(N)$, which means that if $\log_d(N) < k$, there is a further significant restriction. Indeed, this restriction is likely to be in force, for example in the high-dimensional limit as k and N grow proportionally. It follows that $|\mathcal{T}_k|$ may well exhibit dependence upon both k and N, while it is still possible to give an upper bound which is valid for any N.

The tail bounds given in Section 4.3 depend upon the quantity $\mathcal{S}(\delta, \rho)$, which is defined in (4.87). By setting $|S_n| := T(k)$ in (4.87), we obtain the following expression for $\mathcal{S}(\delta, \rho)$.

Lemma 7.2 (**Tree counting limit**). *Let the sequence of sets $\{S_n\}$ be such that $|S_n| = T(k)$, where $T(k)$ is defined in (7.1). Let $\mathcal{S}(\delta, \rho)$ be defined in terms of S_n as in (4.87). Then*

$$\mathcal{S}(\delta, \rho) = d\rho \cdot H(d^{-1}), \tag{7.2}$$

where $H(\cdot)$ is defined in (2.34).

Proof:

$$\begin{aligned}
\lim_{k\to\infty} \frac{1}{k} \ln T(k) &= \lim_{k\to\infty} \frac{1}{k} \ln \left[\frac{1}{(d-1)k+1} \binom{dk}{k} \right] \\
&= \lim_{k\to\infty} \frac{1}{k} \ln \left[\frac{1}{(d-1)k+1} \right] + \lim_{k\to\infty} \frac{1}{k} \ln \binom{dk}{k} \\
&= 0 + \lim_{k\to\infty} d \cdot \frac{1}{dk} \ln \binom{dk}{k} \\
&= d \cdot H(d^{-1}),
\end{aligned} \tag{7.3}$$

where the last step follows from Stirling's formula. Setting $|S_n| := T(k)$, where $T(k)$ is defined in (7.1), it follows from (4.87) and (7.3) that

$$\mathcal{S}(\delta, \rho) = \lim_{n\to\infty} \frac{1}{n} \ln |S_n| = \lim_{n\to\infty} \frac{k}{n} \cdot \frac{1}{k} \ln T(k) = d\rho \cdot H(d^{-1}), \tag{7.4}$$

as required. □

The most striking aspect of the expression in (7.2) is that $\mathcal{S}(\delta, \rho)$ has no dependence upon δ. Since, in the definitions of the tail bound functions in Definitions 4.22 and 4.26, the only dependence upon δ arises from the $\mathcal{S}(\delta, \rho)$ term, it follows that the tail bound functions $\mathcal{IU}_{\mathcal{S}}(\delta, \rho)$, $\mathcal{IL}_{\mathcal{S}}(\delta, \rho)$ and $\mathcal{IF}_{\mathcal{S}}(\delta, \rho)$ themselves have no dependence upon δ. It is therefore legitimate to omit the δ variable in the following definitions.

Definition 7.3 (**Tree-based χ^2 tail bounds**). *Let $\rho \in (0, 1)$ and $\lambda \in (0, 1]$. Let $\mathcal{TIU}(\rho, \lambda)$ be the unique solution to*

$$\nu - \ln(1 + \nu) = \frac{2d\rho \cdot H(d^{-1})}{\lambda} \quad \textit{for} \quad \nu > 0, \tag{7.5}$$

and let $\mathcal{TIL}(\rho, \lambda)$ be the unique solution to

$$-\nu - \ln(1 - \nu) = \frac{2d\rho \cdot H(d^{-1})}{\lambda} \quad \textit{for} \quad \nu \in (0, 1), \tag{7.6}$$

where $H(\cdot)$ is defined in (2.34).

Definition 7.4 (**Tree-based F tail bound**). *Let $\rho \in (0, 1/2]$. Let $\mathcal{TIF}(\rho)$ be the unique solution in f to*

$$\ln(1 + f) - \rho \ln f = 2d\rho \cdot H(d^{-1}) + H(\rho) \quad \textit{for} \quad f > \frac{\rho}{1 - \rho}, \tag{7.7}$$

where $H(\cdot)$ is defined in (2.34).

That these tail bounds are well-defined follows since they are a special case of the tail bounds originally given in Definitions 4.22 and 4.26. The tail bounds in Definitions 7.3 and 7.4 suggest that results in the tree-based setting can be captured within a simplified proportional-dimensional asymptotic framework in which we retain the ρ variable but dispense with the δ variable. Let us then formally define the simplified proportional-growth asymptotic.

Definition 7.5 (**Simplified proportional-growth asymptotic**). *We say that a sequence of problem sizes (k, n, N), where $0 < k \leq n \leq N$, obeys the simplified proportional-growth asymptotic if, for some $\rho \in (0, 1]$,*

$$\frac{k}{n} \to \rho \quad \textit{as } (k, n, N) \to \infty.$$

Note that the only restriction that the simplified proportional-growth asymptotic places upon N is that we must have $N \to \infty$ such that $N \geq n$. We argued in Section 1.9 that the simplified proportional-growth asymptotic is also suggested by the results for ITP in [7], where the tree-based RIP was used to show that $n = C \cdot k$ measurements guarantees recovery using random matrices, for some constant C. In this chapter, we will determine conditions of the form $\rho < \hat{\rho}$ which asymptotically guarantee recovery using ITP algorithms and Gaussian measurement matrices, thereby quantifying the constant C. As was noted in Section 1.5, the factor ρ may be interpreted as an oversampling factor, revealing how many measurements must be taken as a multiple of the sparsity to guarantee recovery. We therefore refer to our results as *oversampling thresholds*. Framing the results within the simplified proportional-growth asymptotic, we have the following tail bound results in the tree-based setting.

Lemma 7.6 (**Tree-based large deviations result for** χ^2)**.** *Let $l \in \{1, \ldots, n\}$ and let the random variables $X_l^i \sim \frac{1}{l}\chi_l^2$ for all $i \in S_n$, where $|S_n| = T(k)$, and let $\epsilon > 0$. In the simplified proportional growth asymptotic, let $l/n \to \lambda \in (0, 1]$. Then*

$$\mathbb{P}\left\{\cup_{i \in S_n}[X_l^i \geq 1 + \mathcal{TIU}(\rho, \lambda) + \epsilon]\right\} \to 0 \tag{7.8}$$

and

$$\mathbb{P}\left\{\cup_{i \in S_n}[X_l^i \leq 1 - \mathcal{TIL}(\rho, \lambda) - \epsilon]\right\} \to 0, \tag{7.9}$$

exponentially in n, where $\mathcal{TIU}(\rho, \lambda)$ and $\mathcal{TIL}(\rho, \lambda)$ are defined in (7.5) and (7.6) respectively.

Proof: The result follows by combining Lemma 4.23, Lemma 7.2 and Definition 7.3. □

Lemma 7.7 (**Tree-based large deviations results for** F)**.** *Let the random variables $X_n^i \sim \frac{k}{n-k+1}\,\mathcal{F}(k, n-k+1)$ for all $i \in S_n$, where $|S_n| = T(k)$, and let $\epsilon > 0$. In the simplified proportional growth asymptotic,*

$$\mathbb{P}\left\{\cup_{i \in S_n}[X_n^i \geq \mathcal{TIF}(\rho) + \epsilon]\right\} \to 0, \tag{7.10}$$

exponentially in n, where $\mathcal{TIF}(\rho)$ is defined in (7.7).

Proof: The result follows by combining Lemma 4.27, Lemma 7.2 and Definition 7.7. □

We next obtain a bound on the upper tree-based RIP constant for Gaussian matrices, within the simplified proportional-growth asymptotic. We define the following bound, which extends to the tree-based setting the RIP bounds derived in [21].

Definition 7.8 (**Tree-based RIP bound**)**.** *Define*

$$\psi_{max}(\lambda, \rho) = \frac{1}{2}\left[(1+\rho)\ln\lambda + 1 + \rho - \rho\ln\rho - \lambda\right]. \tag{7.11}$$

Define $\lambda^{max}(\rho)$ *as the solution to*

$$\psi_{max}[\lambda^{max}(\rho), \rho] + d\rho \cdot H(d^{-1}) = 0 \quad \textit{for} \quad \lambda^{max}(\rho) > 1 + \rho, \tag{7.12}$$

and define $\mathcal{TU}(\rho) = \lambda^{max}(\rho) - 1$.

In order to prove the validity of the bound in Definition 7.8, we will require the following two lemmas concerning the pdf of the largest eigenvalue of a Wishart matrix.

Lemma 7.9 (**Pdf of the largest eigenvalue of a Wishart matrix [85, Lemma 4.2]**). *Let* $A \sim \mathcal{N}_{n,N}(0, 1/n)$ *and let* $A_\Gamma \in \mathbb{R}^{n\times k}$ *be the restriction of* A *to the* k *columns indexed by* Γ. *Let* $f_{max}(k, n; \lambda)$ *denote the pdf of the largest eigenvalue of the* $k \times k$ *Wishart matrix* $A_\Gamma^T A_\Gamma$. *Then* $f_{max}(k, n; \lambda)$ *satisfies*

$$f_{max}(k, n; \lambda) \leq \left[(2\pi)^{1/2} (n\lambda)^{-3/2} \left(\frac{n\lambda}{2} \right)^{(n+k)/2} \frac{1}{\Gamma(\frac{k}{2})\Gamma(\frac{n}{2})} \right] \cdot e^{-n\lambda/2} := g_{max}(k, n; \lambda). \tag{7.13}$$

Lemma 7.10 (**Exponent of the pdf [21, Lemma 2.5]**). *Given* $\rho = \frac{k}{n} \in (0, 1)$,

$$f_{max}(k, n; \lambda) \leq p_{max}(n, \lambda) \exp\left[n \cdot \psi_{max}(\lambda, \rho) \right], \tag{7.14}$$

where $p_{max}(n, \lambda)$ *is a polynomial in* (n, λ).

We now establish that $\mathcal{TU}(\rho)$ is indeed an asymptotic bound on the upper tree-based RIP constant.

Theorem 7.11 (**Validity of tree-based RIP bound**). *Let* $A \sim \mathcal{N}_{n,N}(0, 1/n)$ *and let* $\epsilon > 0$. *In the simplified proportional-growth asymptotic,*

$$\mathbb{P}\left(TU_k < \mathcal{TU}(\rho) + \epsilon \right) \to 1, \tag{7.15}$$

exponentially in n.

Proof: We follow closely the method of proof in [21, Proposition 2.6]. That there is a unique solution to (7.12) follows entirely as in the proof of [21, Proposition 2.6]. Select $\epsilon > 0$ and let (k, n) be such that $k/n = \rho_n$. Then

$$\begin{aligned} \mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon \right] &= \mathbb{P}\left[TU_k \geq \lambda^{max}(\rho_n) - 1 + \epsilon \right] \\ &= \mathbb{P}\left[1 + TU_k \geq \lambda^{max}(\rho_n) + \epsilon \right] \end{aligned} \tag{7.16}$$

Equivalent to (7.16) is the requirement that the maximum eigenvalue of at least one of the $k \times k$ Wishart matrices $A_\Gamma^T A_\Gamma \sim \mathcal{W}_k(n; 0, , 1/n)$ is greater than or equal to $\lambda^{max}(\rho_n) + \epsilon$, considering all $\Gamma \in \mathcal{T}_k$. Writing $f_{max}(k, n; \lambda)$ for the pdf of a $k \times k$ Wishart matrix distributed

as $\mathcal{W}_k(n; 0, 1/n)$, we may therefore perform a union bound over all $\Gamma \in \mathcal{T}_k$ and write

$$\mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon\right] \leq |\mathcal{T}_k| \int_{\lambda^{max}(\rho_n)+\epsilon}^{\infty} f_{max}(k, n; \lambda) d\lambda. \tag{7.17}$$

Using Lemma 7.9, we may write

$$f_{max}(k, n; \lambda) \leq g_{max}(k, n; \lambda) = \phi(n, \rho_n) \lambda^{-\frac{3}{2}} \lambda^{\frac{n}{2}(1+\rho_n)} e^{-\frac{n}{2}\lambda} \tag{7.18}$$

where

$$\phi(n, \rho_n) = (2\pi)^{1/2} n^{-3/2} \left(\frac{n}{2}\right)^{\frac{n}{2}(1+\rho_n)} \frac{1}{\Gamma(\frac{n}{2}\rho_n)\Gamma(\frac{n}{2})}.$$

Since $\lambda^{max}(\rho_n) > 1+\rho_n$, the quantity $\lambda^{\frac{n}{2}(1+\rho_n)} e^{-\frac{n}{2}\lambda}$ is strictly decreasing in λ on $[\lambda^{max}(\rho_n), \infty)$. Combining with (7.18), we therefore have

$$\begin{aligned}
& \int_{\lambda^{max}(\rho_n)+\epsilon}^{\infty} f_{max}(k, n; \lambda) d\lambda \\
\leq\ & \phi(n, \rho_n) \left[\lambda^{max}(\rho_n) + \epsilon\right]^{\frac{n}{2}(1+\rho_n)} e^{-\frac{n}{2}[\lambda^{max}(\rho_n)+\epsilon]} \int_{\lambda^{max}(\rho_n)+\epsilon}^{\infty} \lambda^{-\frac{3}{2}} d\lambda \\
=\ & \left[\lambda^{max}(\rho_n) + \epsilon\right]^{\frac{3}{2}} g_{max}\left[k, n; \lambda^{max}(\rho_n) + \epsilon\right] \int_{\lambda^{max}(\rho_n)+\epsilon}^{\infty} \lambda^{-\frac{3}{2}} d\lambda \\
=\ & 2\left[\lambda^{max}(\rho_n) + \epsilon\right] g_{max}\left[k, n; \lambda^{max}(\rho_n) + \epsilon\right],
\end{aligned}$$

which may be substituted into (7.17) to obtain

$$\mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon\right] \leq 2|\mathcal{T}_k| \left[\lambda^{max}(\rho_n) + \epsilon\right] g_{max}\left[k, n; \lambda^{max}(\rho_n) + \epsilon\right],$$

and we furthermore apply Lemma 7.10 to give

$$\begin{aligned}
&\mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon\right] \\
&\quad \leq 2|\mathcal{T}_k| \left[\lambda^{max}(\rho_n) + \epsilon\right] p_{max}\left[n, \lambda^{max}(\rho_n) + \epsilon\right] \exp\left\{n \cdot \psi_{max}\left[\lambda^{max}(\rho_n) + \epsilon, \rho\right]\right\},
\end{aligned} \tag{7.19}$$

where $p_{max}(n, \lambda)$ is a polynomial in (n, λ). Taking logarithms and limits of the right-hand side of (7.19), using (7.2), we have

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \ln 2|\mathcal{T}_k| \left[\lambda^{max}(\rho_n) + \epsilon\right] p_{max}\left[n, \lambda^{max}(\rho_n) + \epsilon\right] \exp\left\{n \cdot \psi_{max}\left[\lambda^{max}(\rho_n) + \epsilon, \rho\right]\right\} \\
&\qquad = d\rho \cdot H(d^{-1}) + \psi_{max}\left[\lambda^{max}(\rho) + \epsilon, \rho\right],
\end{aligned}$$

from which it follows that, for any $\eta > 0$,

$$\frac{1}{n} \ln \mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon\right] \leq d\rho \cdot H(d^{-1}) + \psi_{max}\left[\lambda^{max}(\rho) + \epsilon, \rho\right] + \eta, \tag{7.20}$$

for all n sufficiently large. By the definition of $\mathcal{TU}(\rho)$ in Definition 7.8, and since $\psi_{max}(\lambda, \rho)$ is

strictly decreasing in λ, then, for any $\epsilon > 0$, we may choose η sufficiently small to ensure

$$\frac{1}{n} \ln \mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon\right] \leq -c_T \quad \text{for all } n \text{ sufficiently large,}$$

where c_T is some positive constant, from which it follows that

$$\mathbb{P}\left[TU_k \geq \mathcal{TU}(\rho_n) + \epsilon\right] \leq e^{-c_T \cdot n} \quad \text{for all } n \text{ sufficiently large,}$$

and (7.15) now follows. □

7.2 Recovery oversampling thresholds

As in Chapter 5, we will obtain recovery results in the case where the measurement matrix is i.i.d. Gaussian, and independent of the underlying signal x^*, as formalized in Assumption 2. Recalling the notational framework for the tree-based model given in Section 6.1, suppose we obtain Gaussian measurements $b = \Phi x^* + e$ with $\Phi \sim \mathcal{N}_{n,N}(0, 1/n)$. Then, provided the wavelet transform matrix $\Psi \in \mathbb{R}^{N \times N}$ is orthogonal, it follows from Lemma 4.3 that $A := \Phi\Psi$ also follows the same Gaussian distribution. Provided we consider an orthogonal DWT, Assumption 2 is therefore sensible. We will also follow Chapter 5 in assuming the noise vector e to be Gaussian distributed and independent of both A and x^*, as formalized in Assumption 3. Recalling (6.21) and (6.22), we adapt the definition of the *unrecoverable energy* Σ of the problem introduced in Section 4.1.2 to the tree-based setting, as follows.

Definition 7.12 (**Unrecoverable energy**). *Given some index set $\Gamma \in \mathcal{T}_k$, define*

$$\tilde{\sigma}_\Gamma := \sqrt{\sigma^2 + \|x^*_{(\Gamma \cup \Lambda)^C}\|^2}, \tag{7.21}$$

and define

$$\Sigma := \sigma + \|x^*_{\Lambda^C}\|. \tag{7.22}$$

7.2.1 Definitions of oversampling thresholds

We define the following two asymptotic recovery oversampling thresholds for ITP algorithms, in terms of the large deviations bounds developed in the previous section.

Definition 7.13 (**Recovery oversampling thresholds**). *Define $\hat{\rho}^{NITP}_{SP}$ to be the unique solution to*

$$\frac{\sqrt{\mathcal{TIF}(\rho)}}{(1-\rho)\left[1 - \mathcal{TIL}(\rho, 1-\rho)\right]} = \frac{1}{1 + \mathcal{TU}(2\rho)} \quad \textit{for} \quad \rho \in (0, 1/2], \tag{7.23}$$

and define $\hat{\rho}_{SP}^{NITP_\kappa}$ *to be the unique solution to*

$$\frac{\sqrt{\mathcal{TIF}(\rho)}}{(1-\rho)\left[1-\mathcal{TIL}(\rho,1-\rho)\right]} = \frac{1}{\kappa[1+\mathcal{TU}(2\rho)]} \quad \textit{for} \quad \rho \in (0,1/2], \tag{7.24}$$

where $\mathcal{TIF}$ *is defined in (7.7),* $\mathcal{TIL}$ *is defined in (7.6) and* $\mathcal{TU}$ *is defined in Definition 7.11.*

It remains to establish that $\hat{\rho}_{SP}^{ITP}$ and $\hat{\rho}_{SP}^{NITP_\kappa}$ are well-defined, to which we devote the rest of this section. The proof, which follows the same approach as the well-definedness proofs for the phase transitions in Section 5.2, relies upon three lemmas, which we give next. The first lemma shows that each of the tail bound functions in Definition 7.13 is strictly increasing in $\rho \in (0,1/2]$.

Lemma 7.14. *For any* $\rho \in (0,1)$, $\mathcal{TIF}(\rho)$, $\mathcal{TIL}(\rho,1-\rho)$ *and* $\mathcal{TU}(\rho)$ *are strictly increasing in* ρ.

Proof: Writing $f = \mathcal{TIF}(\rho)$ and differentiating (7.7) with respect to ρ gives

$$\frac{\partial f}{\partial \rho}\left(\frac{1}{1+f}-\frac{\rho}{f}\right) = 2d\cdot H(d^{-1}) + \ln\left[\frac{f(1-\rho)}{\rho}\right]. \tag{7.25}$$

Now $f > \rho/(1-\rho)$ by (7.7), which implies that

$$\frac{1}{1+f}-\frac{\rho}{f} > 0 \qquad \text{and} \qquad \frac{f(1-\rho)}{\rho} > 1,$$

making the logarithm in (7.25) strictly positive, and the result now follows for $\mathcal{TIF}(\rho)$. Similarly, writing $l = \mathcal{TIL}(\rho,1-\rho)$ and differentiating (7.6) with $\gamma = 1-\rho$ gives

$$(1-\rho)\left(\frac{1}{1-l}-1\right)\frac{\partial l}{\partial \rho} - [-l-\ln(1-l)] = 2d\cdot H(d^{-1}).$$

A further application of (7.6) followed by some rearrangement gives

$$\frac{\partial l}{\partial \rho}\left(\frac{l(1-\rho)}{1-l}\right) = \frac{2d\cdot H(d^{-1})}{1-\rho}. \tag{7.26}$$

The result for $\mathcal{TIL}(\rho,1-\rho)$ now follows since all terms are strictly positive. Finally, differentiating (7.11) with respect to ρ gives

$$\frac{\partial \psi_{max}[\lambda^{max}(\rho),\rho]}{\partial \rho} = \frac{1}{2}\left[\ln\left(\frac{\lambda^{max}}{\rho}\right) - \frac{\partial \lambda^{max}}{\partial \rho}\left(1-\frac{1+\rho}{\lambda^{max}}\right)\right],$$

and we may then differentiate (7.12) to obtain, after a little rearrangement,

$$\frac{\partial \lambda^{max}}{\partial \rho}\left(1-\frac{1+\rho}{\lambda^{max}}\right) = \ln\left(\frac{\lambda^{max}}{\rho}\right) + 2d\cdot H(d^{-1}).$$

Now, since $\lambda^{max} > 1 + \rho$,

$$1 - \frac{1+\rho}{\lambda^{max}} > 1 - \frac{1+\rho}{1+\rho} = 0 \qquad \text{and} \qquad \frac{\lambda^{max}}{\rho} > \frac{1+\rho}{\rho} \geq 2 > 0,$$

and the result now follows for $\mathcal{TU}(\rho)$. □

The second lemma shows that $\mathcal{TIF}(\rho)$ grows to be much greater than 1 at $\rho = 1/2$.

Lemma 7.15.

$$\mathcal{TIF}(1/2) \geq 31 + 8\sqrt{15}.$$

Proof: Substituting $\rho = 1/2$ into (7.7) implies that $\mathcal{TIF}(1/2)$ solves for f the equation

$$\ln(1+f) - \frac{1}{2}\ln f = \frac{d}{2} \cdot H(d^{-1}) + H\left(\frac{1}{2}\right); \quad f > 1.$$

Now $H(1/2) = \ln 2$, and we may lower bound $d \cdot H(d^{-1})$ by $2\ln 2$ for any $d \geq 2$, which implies that

$$\ln(1+f) - \frac{1}{2}\ln f \geq 3\ln 2,$$

which may be rearranged to give

$$f^2 - 62f + 1 \geq 0,$$

which, together with $f > 1$, yields the required result. □

The third lemma proves that each of the tail bound functions in Definition 7.13 tends to zero as $\rho \to 0$.

Lemma 7.16. *The following limiting results hold:*

$$\lim_{\rho\to 0} \mathcal{TIF}(\rho) = 0; \qquad \lim_{\rho\to 0} \mathcal{TIL}(\rho, 1-\rho) = 0; \qquad \lim_{\rho\to 0} \mathcal{TU}(\rho) = 0.$$

Proof: By (5.9) and (7.2), the right-hand side of (7.7) tends to zero as $\rho \to 0$. Writing $f = \mathcal{TIF}(\rho)$, it follows from (4.100) that

$$\lim_{\rho\to 0} [\ln(1+f) - \rho \ln f] \leq 0. \tag{7.27}$$

However, Lemma 7.15 and (7.7) imply that

$$-\rho\ln(31 + 8\sqrt{15}) \leq -\rho \ln f \leq -\rho \ln\left(\frac{\rho}{1-\rho}\right),$$

from which it follows that $\rho \ln f \to 0$ as $\rho \to 0$, which combines with (7.27) to give $\ln(1+f) \to 0$ as $\rho \to 0$, which yields the result for $\mathcal{TIF}(\rho)$. It also follows from (5.9) and (7.2) that the right-

hand side of (7.6) with $\gamma = 1-\rho$ also tends to zero as $\rho \to 0$, which in turn implies the result for $\mathcal{TIL}(\rho)$. Finally, taking limits of (7.12) as $\rho \to 0$, we see that $\psi_{max}[\lambda^{max}(\rho), \rho] \to 0$ as $\rho \to 0$. It then follows from (7.11) that $\lambda^{max}(\rho) \to 1$ as $\rho \to 0$, from which the result for $\mathcal{TU}(\rho)$ follows.□

That $\hat{\rho}_{SP}^{ITP}$ is well-defined may now be shown as follows. Let us define $\psi^{ITP}(\rho)$ to be

$$\psi^{ITP}(\rho) := \frac{\sqrt{\mathcal{TIF}(\rho)}}{(1-\rho)\left[1-\mathcal{TIL}(\rho, 1-\rho)\right]} - \frac{1}{1+\mathcal{TU}(2\rho)}.$$

By Lemma 7.16, we have

$$\lim_{\rho \to 0} \psi^{ITP}(\rho) = -1 < 0.$$

By Lemma 7.14, $\psi^{ITP}(\rho)$ is strictly increasing in ρ for $\rho \in (0, 1/2]$. Also, Lemma 7.15, $\mathcal{TU} \geq 0$ and $\mathcal{TIL} \geq 0$ together imply that $\psi^{IH}P(1/2) > 0$. It therefore follows that there exists a unique $\rho \in (0, 1/2]$ for which $\psi^{ITP}(\rho) = 0$ and the definition of $\hat{\rho}_{SP}^{ITP}$ is valid. A similar argument applies for $\hat{\rho}_{SP}^{NITP_\kappa}$.

We proceed next to proving our main recovery results for ITP algorithms, beginning with (constant stepsize) ITP in the next section.

7.2.2 Recovery results for ITP

We will follow closely the argument in Section 5.3.1. The only change is that we now switch to using the tree-based tail bounds defined in Section 7.1. Since there is now no dependence upon δ, we also switch to proving results in the simplified proportional-growth asymptotic. We begin by introducing several definitions.

Definition 7.17 (**Stability factor for ITP**). *Consider Problem 6.2. Given $\rho \in (0, 1/2]$ and $\underline{\alpha} > 0$, provided*

$$\underline{\alpha} > \frac{\sqrt{\mathcal{TIF}(\rho)}}{(1-\rho)[1-\mathcal{TIL}(\rho, 1-\rho)]}, \tag{7.28}$$

define

$$a(\rho) := \frac{1 + \sqrt{\mathcal{TIF}(\rho)} + \underline{\alpha}\sqrt{\rho(1-\rho)[1+\mathcal{TIU}(\rho, 1-\rho)][1+\mathcal{TIU}(\rho,\rho)]}}{\underline{\alpha}(1-\rho)[1-\mathcal{TIL}(\rho, 1-\rho)] - \sqrt{\mathcal{TIF}(\rho)}}, \tag{7.29}$$

and

$$\xi(\rho) := \sqrt{\mathcal{TIF}(\rho)\left[1+a(\rho)\right]^2 + 1 + \left[a(\rho)\right]^2}, \tag{7.30}$$

where $\mathcal{TIF}$ is defined in (7.7), and where $\mathcal{TIU}$ and $\mathcal{TIL}$ are defined in (7.5) and (7.6) respectively.

Note that (7.28) ensures that the denominator in (7.29) is strictly positive and that $a(\rho)$ is therefore well-defined. The function $\xi(\rho)$ will represent a stability factor in our results, bounding the approximation error of the output of IHT as a multiple of the unrecoverable energy Σ.

Definition 7.18 (**Support set partition for ITP**). *Consider Problem 6.2 and suppose* $\rho \in (0, 1/2]$ *and* $\alpha > 0$. *Given* $\zeta > 0$, *let us write*

$$a^*(\rho; \zeta) := a(\rho) + \zeta, \tag{7.31}$$

let us write $\{\Gamma_i : i \in \mathcal{T}_k\}$ *for the set of all possible support sets which form a rooted tree of cardinality* k, *and let us disjointly partition* $\mathcal{T}_k := \Theta_n^1 \cup \Theta_n^2$ *such that*

$$\Theta_n^1 := \left\{ i \in \mathcal{T}_k \ : \ \|x^*_{\Lambda \setminus \Gamma_i}\| > \Sigma \cdot a^*(\rho; \zeta) \right\} \quad \textit{and} \quad \Theta_n^2 := \left\{ i \in \mathcal{T}_k \ : \ \|x^*_{\Lambda \setminus \Gamma_i}\| \leq \Sigma \cdot a^*(\rho; \zeta) \right\}, \tag{7.32}$$

where Σ *is defined in (7.22), and where* Λ *is defined in (6.22).*

We proceed by means of three lemmas which mimic Lemmas 5.10, 5.12 and 5.13. Combining all three results, we have convergence to some α-stable point with guaranteed approximation error, provided the conditions in each lemma hold, and combining the conditions leads to the oversampling threshold defined in (7.13). Where the argument is identical except for the use of different tail bound functions, we will merely state what replacements need to be made.

We first show that there are asymptotically no α-stable points on any Γ_i such that $i \in \Theta_n^1$, and we write NSP_α for this event.

Lemma 7.19. *Consider Problem 6.2 and choose* $\zeta > 0$. *Suppose Assumptions 2 and 3 hold, and suppose that (7.28) holds. Then, in the simplified proportional-growth asymptotic, there are no* α*-stable points on any* Γ_i *such that* $i \in \Theta_n^1$, *with probability tending to* 1 *exponentially in* n.

Proof: In the proof of Lemma 5.10, we replace $a(\delta, \rho)$ by $a(\rho)$, which is defined in (7.29), and we replace $a^*(\delta, \rho; \zeta)$ by $a^*(\rho; \zeta)$, which is defined in (7.31). We also now let Θ_n^1 and Θ_n^2 be defined as in (7.32). Then, with these replacements, and replacing the condition (5.11) with the condition (7.28), we may follow the argument of Lemma 5.10 to deduce the equivalent of (5.26), namely that, in the simplified proportional-growth asymptotic,

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}(\overline{NSP_\alpha}) \\ \leq \ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(F_{\Gamma_i} \geq F^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(G_{\Gamma_i} \geq G^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(R_{\Gamma_i} \leq R^*)\right\} \\ + \ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(S_{\Gamma_i} \geq S^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(T_{\Gamma_i} \geq T^*)\right\}, \end{aligned} \tag{7.33}$$

where

$$\begin{gathered} F_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1); \quad G_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1); \\ R_{\Gamma_i} \sim \frac{1}{n-k} \chi^2_{n-k}; \quad S_{\Gamma_i} \sim \frac{1}{n-k} \chi^2_{n-k}; \quad T_{\Gamma_i} \sim \frac{1}{k} \chi^2_k. \end{gathered}$$

and

$$\begin{aligned} F^* = G^* &:= \mathcal{TIF}(\rho) + \epsilon; \qquad R^* := 1 - \mathcal{TIL}(\rho, 1-\rho) - \epsilon; \\ S^* &:= 1 + \mathcal{TIU}(\rho, 1-\rho) + \epsilon; \qquad T^* := 1 + \mathcal{TIU}(\rho, \rho) + \epsilon. \end{aligned}$$

Since $|\Theta_n^1| \leq T(k)$, we may apply Lemmas 7.6 and 7.7 to (7.33), and the result now follows. □

We may deduce from Lemma 7.19 that, in the case of an exactly tree-sparse signal and noiseless measurements (Problem 6.1), a condition can be given guaranteeing that ITP has a single fixed point, namely the underlying tree-sparse signal x^*.

Corollary 7.20 (**Single fixed point condition for ITP**). *Consider Problem 6.1. Suppose Assumption 2 holds, and suppose that (7.28) holds. Then, in the simplified proportional-growth asymptotic, x^* is the only fixed point of ITP with stepsize α, with probability tending to* 1 *exponentially in n.*

Proof: Lemma 7.19 establishes that, if (7.28) holds, there are asymptotically no α-stable points on any Γ_i such that $i \in \Theta_n^1$, while, setting $\Sigma := 0$ in (7.32), we have $i \in \Theta_n^2 \Rightarrow \Gamma_i = \Lambda$. Therefore any α-stable point is supported on Λ, and Lemma 6.9 implies that it must be x^*. However, any fixed point of ITP with stepsize α is necessarily an α-stable point, and therefore x^* is also the only fixed point of ITP with stepsize α. □

Next, we show that any α-stable points on Γ_i such that $i \in \Theta_n^2$ have bounded approximation error.

Lemma 7.21. *Consider Problem 6.2. Suppose Assumptions 2 and 3 hold, and suppose that (7.28) holds. Then there exists ζ sufficiently small such that, in the simplified proportional-growth asymptotic, any α-stable point $\bar{x}$ on Γ_i such that $i \in \Theta_n^2$ satisfies*

$$\|\bar{x} - x^*\| \leq \xi(\rho) \cdot \Sigma, \tag{7.34}$$

with probability tending to 1 *exponentially in n, where $\xi(\rho)$ is defined in (7.30) and Σ is defined in (7.22).*

Proof: In the proof of Lemma 5.12, we replace $a(\delta, \rho)$ by $a(\rho)$, which is defined in (7.29), we replace $\xi(\delta, \rho)$ by $\xi(\rho)$, which is defined in (7.30), and we replace $a^*(\delta, \rho; \zeta)$ by $a^*(\rho; \zeta)$, which is defined in (7.31). We also now let Θ_n^1 and Θ_n^2 be defined as in (7.32). Then, with these replacements, and replacing the condition (5.11) with the condition (7.28), and denoting by $\bar{x}_i$ the minimum-norm solution on Γ_i, we may follow the argument of Lemma 5.12 to deduce the equivalent of (5.41), namely that, in the simplified proportional-growth asymptotic,

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and } \|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[a^*(\rho;\zeta), P^*, Q^*\right]\right\} \\ \leq \; & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^2}(P_{\Gamma_i} \geq P^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^2}(Q_{\Gamma_i} \geq Q^*)\right\}, \end{aligned} \tag{7.35}$$

where

$$P_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1); \quad Q_{\Gamma_i} \sim \frac{k}{n-k+1} F(k, n-k+1),$$

$$\Psi(a,P,Q) := \sqrt{\left(a\cdot\sqrt{P}+\sqrt{Q}\right)^2+1+a^2},$$

and

$$P^* = Q^* := \mathcal{TIF}(\rho)+\zeta.$$

Since $|\Theta_n^1| \leq T(k)$, we may apply Lemmas 7.6 and 7.7 to (7.35), and we may now follow the remainder of the argument of the proof of Lemma 5.12 to deduce the result. □

In the context of ITP, we obtain the following convergence result in the simplified proportional-dimensional asymptotic framework.

Lemma 7.22. *Consider Problem 6.2. Suppose Assumption 2 holds, suppose that the stepsize α of ITP is chosen to satisfy*

$$\alpha < \frac{1}{1+\mathcal{TU}(2\rho)}. \tag{7.36}$$

Then, in the simplified proportional-growth asymptotic, ITP converges to an α-stable point with probability tending to 1 exponentially in n.

Proof: Given (7.36), we may apply Lemma 7.11 with ϵ sufficiently small to deduce $\alpha(1+U_{2k})<1$, with probability tending to 1 exponentially in n. Under Assumption 2, we may then apply Theorem 6.14 and deduce convergence of ITP to an α-stable point. □

We now combine Lemmas 7.19, 7.21 and 7.22 and prove the main recovery result for ITP.

Theorem 7.23 (**Recovery result for ITP**). *Consider Problem 6.2. Suppose Assumptions 2 and 3 hold, suppose that*

$$\rho < \hat{\rho}_{SP}^{ITP}, \tag{7.37}$$

where $\hat{\rho}_{SP}^{ITP}$ is defined in (7.23), and suppose that α satisfies

$$\frac{\sqrt{\mathcal{TIF}(\rho)}}{(1-\rho)\left[1-\mathcal{TIL}(\rho,1-\rho)\right]} < \alpha < \frac{1}{1+\mathcal{TU}(2\rho)}. \tag{7.38}$$

Then, in the simplified proportional-growth asymptotic, ITP converges to $\bar{x}$ such that (7.34) holds with probability tending to 1 exponentially in n.

Proof: First note that (7.37) implies that the interval in (7.38) is well-defined. Provided α is chosen to satisfy (7.38), (6.39) holds, and under Assumption 2, we may apply Lemma 7.22 to deduce convergence of ITP to an α-stable point. On the other hand, Lemma 7.19 establishes that there are asymptotically no α-stable points on any Γ_i such that $i \in \Theta_n^1$, while we may apply Lemma 7.21 to deduce that any α-stable points on any Γ_i such that $i \in \Theta_n^2$ satisfy (7.34).□

In the special case of Problem 6.1, the same oversampling threshold guarantees exact recovery of the underlying signal x^*.

Corollary 7.24 (**Noiseless case**). *Consider Problem 6.1. Suppose Assumption 2 holds, suppose that (7.37) holds, and suppose that α satisfies (7.38). Then, in the simplified proportional-growth asymptotic, ITP converges to x^* with probability tending to 1 exponentially in n.*

Proof: The result follows by setting $\Sigma := 0$ in Theorem 7.23. □

7.2.3 Results for NITP

We will follow closely the argument in Section 5.3.2, with the only changes being the switch to tree-based tail bounds and the simplified proportional-growth asymptotic. We begin with some definitions.

Definition 7.25 (**Stability factor for NITP**). *Consider Problem 6.2. Given $\rho \in (0, 1/2]$, provided*

$$\rho < \hat{\rho}_{SP}^{NITP_\kappa}, \tag{7.39}$$

define

$$a(\rho) := \frac{1 + \sqrt{\mathcal{TIF}(\rho)} + \{\kappa[1 + \mathcal{TU}(2\rho)]\}^{-1}\sqrt{\rho(1-\rho)[1 + \mathcal{TIU}(\rho, 1-\rho)][1 + \mathcal{TIU}(\rho,\rho)]}}{(1-\rho)\{\kappa[1 + \mathcal{TU}(2\rho)]\}^{-1}[1 - \mathcal{TIL}(\rho, 1-\rho)] - \sqrt{\mathcal{TIF}(\rho)}}, \tag{7.40}$$

and

$$\xi(\rho) := \sqrt{\mathcal{TIF}(\rho)\,[1 + a(\rho)]^2 + 1 + [a(\rho)]^2}, \tag{7.41}$$

where $\mathcal{TIF}$ is defined in (7.7), where $\mathcal{TIU}$ and $\mathcal{TIL}$ are defined in (7.5) and (7.6) respectively, and where $\mathcal{TU}$ is defined in Definition 7.11.

Definition 7.26 (**Support set partition for NITP**). *Consider Problem 6.2 and suppose $\rho \in (0, 1/2]$ and $\alpha > 0$. Given $\zeta > 0$, let us write*

$$a^*(\rho; \zeta) := a(\rho) + \zeta, \tag{7.42}$$

let us write $\{\Gamma_i : i \in \mathcal{T}_k\}$ for the set of all possible support sets which form a rooted tree of cardinality k, and let us disjointly partition $\mathcal{T}_k := \Theta_n^1 \cup \Theta_n^2$ such that

$$\Theta_n^1 := \left\{ i \in \mathcal{T}_k \;:\; \|x^*_{\Lambda\setminus\Gamma_i}\| > \Sigma \cdot a^*(\rho;\zeta) \right\} \quad \text{and} \quad \Theta_n^2 := \left\{ i \in \mathcal{T}_k \;:\; \|x^*_{\Lambda\setminus\Gamma_i}\| \le \Sigma \cdot a^*(\rho;\zeta) \right\}, \tag{7.43}$$

where Σ is defined in (7.22), and where Λ is defined in (6.22).

We proceed by means of two lemmas which mimic Lemmas 5.18 and 5.19. Where the argument is identical except for the use of different tail bound functions, we will merely state what replacements need to be made.

We first show that, provided the condition (7.39) holds, NITP is asymptotically guaranteed to converge to an $\underline{\alpha}(\rho;\epsilon)$-stable point on some Γ_i such that $i \in \Theta_n^2$, where

$$\underline{\alpha}(\rho;\epsilon) := \{\kappa[1 + \mathcal{TU}(2\rho) + \epsilon]\}^{-1}, \tag{7.44}$$

for some $\epsilon > 0$. We write $NSP_{\underline{\alpha}}$ for the event that there is no $\underline{\alpha}$-stable point on any Γ_i such that $i \in \Theta_n^1$.

Lemma 7.27. *Consider Problem 6.2 and choose $\zeta > 0$. Suppose Assumptions 2 and 3 hold, and suppose that (7.39) holds. Then there exists ϵ such that, in the simplified proportional-growth asymptotic, NITP with shrinkage parameter κ converges to an $\underline{\alpha}(\rho;\epsilon)$-stable point on some Γ_i such that $i \in \Theta_n^2$, with probability tending to 1 exponentially in n.*

Proof: Under Assumption 2, we have by Theorem 6.15 convergence of NITP to a $[\kappa(1 + TU_{2k})]^{-1}$-stable point. By Definition 6.8, for any $\alpha_1 < \alpha_2$, the set of α_1-stable points includes the set of α_2-stable points, and this observation combines with Theorem 7.11 to imply convergence to an $\underline{\alpha}(\rho;\epsilon)$-stable point, where $\underline{\alpha}(\rho;\epsilon)$ is defined in (7.44), with probability tending to 1 exponentially in n. We now show that, provided ϵ is taken sufficiently small, this stable point must be on Γ_i such that $i \in \Theta_n^2$. In the proof of Lemma 5.18, we replace $\underline{\alpha}(\delta,\rho;\epsilon)$ by $\underline{\alpha}(\rho;\epsilon)$, which is defined in (7.44), we replace $a(\delta,\rho)$ by $a(\rho)$, which is defined in (7.40), and we replace $a^*(\delta,\rho;\zeta)$ by $a^*(\rho;\zeta)$, which is defined in (7.42). We also now let Θ_n^1 and Θ_n^2 be defined as in (7.43). Then, with these replacements, and replacing the condition (5.47) with the condition (7.39), we may follow the argument of Lemma 5.18 to deduce the equivalent of (5.61), namely that, in the simplified proportional-growth asymptotic,

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}(\overline{NSP_{\underline{\alpha}}}) \\ \leq\ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(F_{\Gamma_i} \geq F^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(G_{\Gamma_i} \geq G^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(R_{\Gamma_i} \leq R^*)\right\} \\ +\ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(S_{\Gamma_i} \geq S^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^1}(T_{\Gamma_i} \geq T^*)\right\}, \end{aligned} \tag{7.45}$$

where

$$F_{\Gamma_i} \sim \frac{k}{n-k+1}F(k,n-k+1); \quad G_{\Gamma_i} \sim \frac{k}{n-k+1}F(k,n-k+1);$$
$$R_{\Gamma_i} \sim \frac{1}{n-k}\chi^2_{n-k}; \quad S_{\Gamma_i} \sim \frac{1}{n-k}\chi^2_{n-k}; \quad T_{\Gamma_i} \sim \frac{1}{k}\chi^2_k.$$

and

$$F^* = G^* := \mathcal{TIF}(\rho) + \epsilon; \qquad R^* := 1 - \mathcal{TIL}(\rho, 1-\rho) - \epsilon;$$
$$S^* := 1 + \mathcal{TIU}(\rho, 1-\rho) + \epsilon; \qquad T^* := 1 + \mathcal{TIU}(\rho,\rho) + \epsilon.$$

Since $|\Theta_n^1| \leq T(k)$, we may apply Lemmas 7.6 and 7.7 to (7.45), and the result now follows. □

Next, we show that any $\underline{\alpha}(\rho;\epsilon)$-stable points on Γ_i such that $i \in \Theta_n^2$ have bounded approximation error.

Lemma 7.28. *Consider Problem 6.2. Suppose Assumptions 2 and 3 hold, and suppose that (7.39) holds. Given any $\epsilon > 0$, there exists ζ sufficiently small such that, in the simplified proportional-growth asymptotic, any $\underline{\alpha}(\rho;\epsilon)$-stable point on Γ_i such that $i \in \Theta_n^2$ satisfies*

$$\|\bar{x} - x^*\| \leq \xi(\rho) \cdot \Sigma, \tag{7.46}$$

with probability tending to 1 exponentially in n, where $\xi(\rho)$ is defined in (7.41) and Σ is defined in (7.22).

Proof: In the proof of Lemma 5.19, we replace $\underline{\alpha}(\delta,\rho;\epsilon)$ by $\underline{\alpha}(\rho;\epsilon)$, which is defined in (7.44), we replace $a(\delta,\rho)$ by $a(\rho)$, which is defined in (7.40), we replace $\xi(\delta,\rho)$ by $\xi(\rho)$, which is defined in (7.41), and we replace $a^*(\delta,\rho;\zeta)$ by $a^*(\rho;\zeta)$, which is defined in (7.42). We also now let Θ_n^1 and Θ_n^2 be defined as in (7.43). Then, with these replacements, and replacing the condition (5.47) with the condition (7.39), and denoting by $\bar{x}_i$ the minimum-norm solution on Γ_i, we may follow the argument of Lemma 5.19 to deduce the equivalent of (5.71), namely that, in the simplified proportional-growth asymptotic,

$$\begin{aligned} & \lim_{n\to\infty} \mathbb{P}\left\{\exists \text{ some } \Gamma_i \text{ such that } i \in \Theta_n^2 \text{ and } \|\bar{x}_i - x^*\| > \Sigma \cdot \Psi\left[a^*(\rho;\zeta), P^*, Q^*\right]\right\} \\ \leq\ & \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^2}(P_{\Gamma_i} \geq P^*)\right\} + \lim_{n\to\infty} \mathbb{P}\left\{\cup_{i\in\Theta_n^2}(Q_{\Gamma_i} \geq Q^*)\right\}, \end{aligned} \tag{7.47}$$

where

$$P_{\Gamma_i} \sim \frac{k}{n-k+1}F(k,n-k+1); \quad Q_{\Gamma_i} \sim \frac{k}{n-k+1}F(k,n-k+1),$$

$$\Psi(a,P,Q) := \sqrt{\left(a\cdot\sqrt{P}+\sqrt{Q}\right)^2 + 1 + a^2},$$

and

$$P^* = Q^* := \mathcal{TIF}(\rho) + \zeta.$$

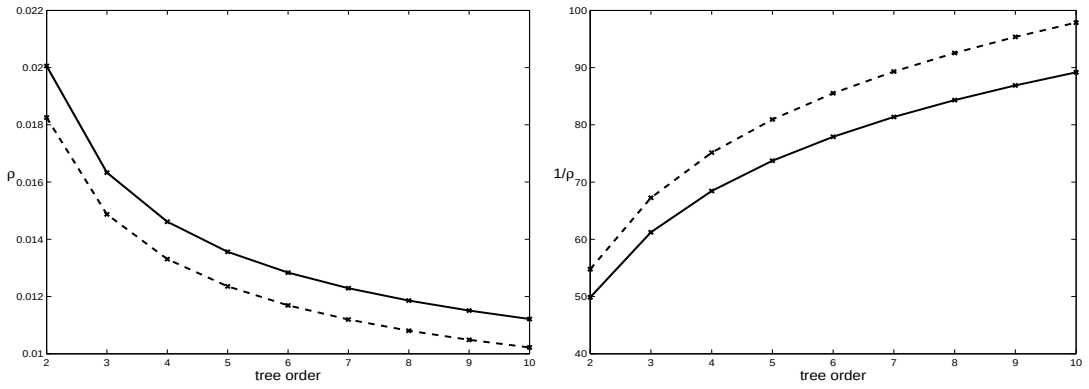
Since $|\Theta_n^1| \leq T(k)$, we may apply Lemmas 7.6 and 7.7 to (7.47), and, we may now follow the remainder of the argument of the proof of Lemma 5.19 to deduce the result. □

Combining Lemmas 7.27 and 7.28, we have the following recovery result for NITP.

Theorem 7.29 (**Recovery result for NITP**). *Consider Problem 6.2. Suppose Assumptions 2 and 3 hold and suppose that (7.39) holds. Then, in the simplified proportional-growth asymptotic, NITP converges to $\bar{x}$ such that (7.46) holds with probability tending to 1 exponentially in n.*

In the case of Problem 6.1, Theorem 7.29 simplifies to an exact recovery result.

Corollary 7.30 (**Noiseless case**). *Consider Problem 6.1. Suppose Assumption 2 holds and suppose that (7.39) holds. Then, in the simplified proportional-growth asymptotic, NITP with shrinkage parameter κ converges to x^* with probability tending to 1 exponentially in n.*



(a) (b)

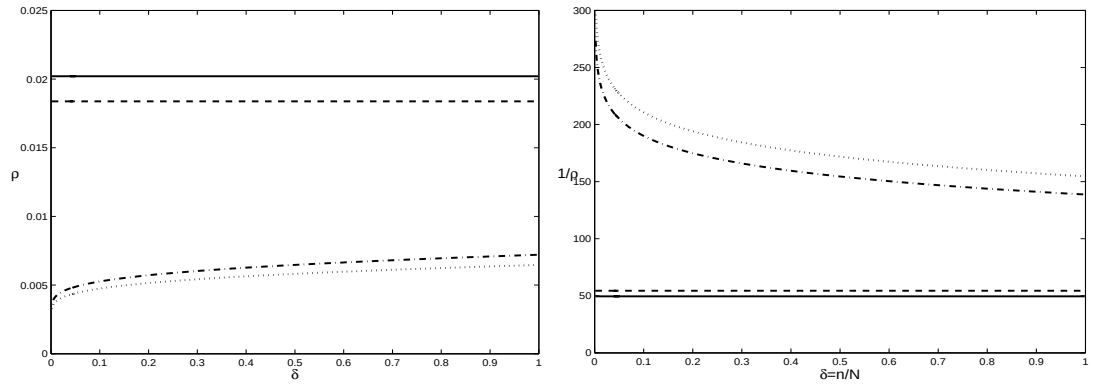
Figure 7.1: (a) Critical ρ-values for different tree orders: ITP – unbroken; NITP – dashed. (b) Corresponding oversampling factors (reciprocals of $\hat{\rho}$).

7.3 Illustration and discussion of recovery results

7.3.1 Noiseless case

Oversampling thresholds for exact recovery. In the simplified case of Problem 6.1, where we seek to recover an exactly k-tree sparse signal x^* from noiseless measurements, Corollaries 7.24 and 7.30 establish that, given a sequence of k-tree sparse signals x^* and independently drawn $n \times N$ Gaussian measurement matrices A, provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{SP}^{ITP}$ or $\hat{\rho}_{SP}^{NITP_\kappa}$, the probability of exact recovery of the original signal x^* approaches 1 exponentially fast in n. These oversampling thresholds also depend on the value of d, the order of the tree structure. For binary trees ($d = 2$), for example, we have $\hat{\rho}_{SP}^{ITP} \approx 0.0201$ for ITP and $\hat{\rho}_{SP}^{NITP_{1.1}} \approx 0.0183$ for NITP. Figure 7.1(a) plots these oversampling thresholds for different tree orders, taking $\kappa = 1.1$ in each case for NITP. Figure 7.1(b) shows the inverse of the oversampling ratio, which indicates the number of measurements required by the analysis as a multiple of the sparsity. For binary trees, we find that $n \geq 49.86k$ measurements guarantees recovery by ITP, while $n \geq 54.78k$ measurements guarantees recovery by NITP. We see a measured deterioration in the results for higher tree orders.

Comparison with results for IHT algorithms. While in the present chapter we have dispensed with the undersampling ratio $\delta = n/N$, we may also frame our results in the (δ, ρ) asymptotic in order to make a comparison with the recovery phase transitions derived for IHT and NIHT presented in Section 5.4. Since there is no dependence upon δ, the phase transitions we obtain are simply horizontal lines in the (δ, ρ)-plane. Phase transitions for binary trees are displayed in Figure 7.2 alongside the phase transitions for IHT algorithms given in Section 5.4, with recovery asymptotically guaranteed beneath the respective curves. We observe that the switch to the tree-based setting leads to significantly improved results, especially for small δ. An improvement is to be expected since we are considering a more refined model in which



(a) (b)

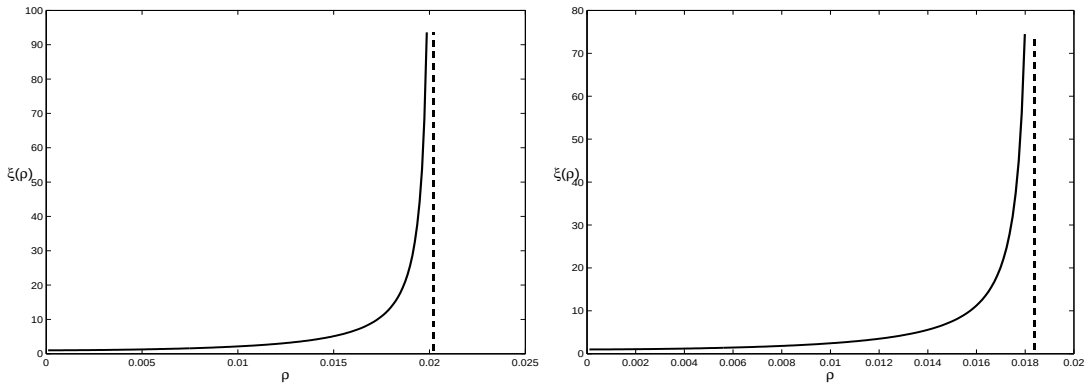
Figure 7.2: (a) Phase transitions in the (δ, ρ) framework for binary trees (ITP – unbroken; NITP – dashed) and non-tree-based (IHT – dash-dot; NIHT – dotted). (b) Corresponding inverses of the phase transition.

a smaller number of support sets are permissible, which allowed us to tighten union bound arguments.

Interpretation as lower bounds on a weak phase transition. Like in Chapter 5, the recovery results presented in this chapter assume independence between the underlying signal and the measurement matrix. The results of this chapter are therefore also average-case results. The phase transitions given in Figure 7.2 are therefore lower bounds on a particular kind of weak phase transition for ITP algorithms, in which the signal is assumed to be statistically independent of the measurement matrix.

Our average-case guarantees are in contrast to worst-case guarantees which apply to all possible signals for a given measurement matrix. Worst-case recovery guarantees for ITP with unit stepsize were obtained in [7] in terms of the tree-based RIP, and the approach of Section 3.3 could be followed to obtain a worst-case recovery condition for NITP. Furthermore, we have seen that the upper tree-based RIP constant can be analysed within the simplified proportional-growth asymptotic framework, and unsurprisingly the same is also true for the lower tree-based RIP constant. It follows that worst-case oversampling thresholds could also be obtained for ITP algorithms in the case of Gaussian measurement matrices. Viewed in the (δ, ρ) asymptotic framework, these results would give lower bounds on the strong phase transition for ITP algorithms and Gaussian matrices. Though we omit details of proofs, we have computed these phase transitions, and they are lower than those displayed in Figure 7.2 based upon the results in this chapter. We saw in Section 5.4 that the switch from RIP to average-case analysis leads to improved phase transitions in the case of IHT algorithms, and so it is not surprising that our results for ITP also outperform those based upon tree-based RIP.

To the best of our knowledge, little is currently known about empirical weak phase transitions for ITP algorithms, though it is expected that they would be at least as high as those for



(a) (b)

Figure 7.3: Plot of the stability factor $\xi(\rho)$ for binary trees: (a) ITP; (b) NITP.

IHT algorithms since the model is more refined. Indeed, due to our continued use of worst-case techniques, such as the tree-based RIP and large deviations analysis, it should be expected that our phase transitions presented here are far from sharp and are pessimistic compared to the average-case behaviour of ITP algorithms in practice. However, as we noted in the context of IHT algorithms in Section 5.4, the use of the average-case independence assumption to analyse the stable point condition has allowed us to break free in part from the restrictions of worst-case analysis.

7.3.2 General case

Stability factors. In the case of Problem 6.2, where signals are only k-tree compressible and measurements are contaminated by noise, exact recovery of the original signal is impossible. However, Theorem 7.23 guarantees that, under the same condition guaranteeing exact recovery for Problem 6.1, the approximation error of the output of ITP/NITP is asymptotically bounded by some known stability factor $\xi(\rho)$ multiplied by the unrecoverable energy Σ. More precisely, consider a sequence of k-tree compressible signals x^* and independently drawn $n \times N$ Gaussian measurement matrices A. Provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{SP}^{ITP}$, then the probability that ITP with stepsize satisfying (7.38) satisfies (7.34) with $\underline{\alpha} := \alpha$ approaches 1 exponentially fast in n. Similarly, provided the ratio $\rho = k/n$ falls below $\hat{\rho}_{SP}^{NITP_\kappa}$, then the probability that NITP with shrinkage parameter κ satisfies (7.46) with $\underline{\alpha} := \{\kappa[1 + \mathcal{TU}(2\rho)]\}^{-1}$ approaches 1 exponentially fast in n.

In the case of binary trees, plots of the stability factor $\xi(\rho)$ for both ITP and NITP ($\kappa := 1.1$) are displayed in Figure 7.3. For ITP, there is some freedom in the choice of the stepsize: for these plots we select a stepsize at the upper limit of the permissible interval (7.38) since this minimizes the stability factor. In keeping with the plots in Section 5.4, we observe that the stability factor tends to infinity as the transition point is reached.

In [7], it was observed that the tree-based RIP is not sufficient to give amplification bounds

on the tail of an inexactly sparse signal (as in Lemma 2.2). Consequently, the authors introduce a further *Restricted Amplification Property* (RAmP), and show that it holds with overwhelming probability for random matrices with $n = \mathcal{O}(k)$. In comparison, our results do not require this additional property. We are able to entirely circumvent the use of the RAmP due to the assumption of independence between signal and measurement matrix, which allows us to bound the amplification of the tail of the signal using the tail bounds derived in this thesis.

Chapter 8

Conclusions and future directions

CS theory was first developed within the framework of l_1-l_0 equivalence, a phenomenon which means that sparse solutions to appropriately designed underdetermined linear systems can be found using l_1-minimization techniques. However, there is growing empirical evidence that other algorithms which do not rely upon l_1-minimization can be equally effective in finding sparse solutions to underdetermined systems (Sections 1.5 and 1.11). In particular, the gradient-based IHT and NIHT algorithms [26, 27] have favourable computational efficiency compared to many other CS algorithms, and therefore represent a competitive alternative to soft-thresholding based gradient methods for l_1-minimization (Section 1.6).

It is important that a CS recovery algorithm is supported by theory which quantitatively determines the degree of undersampling that the algorithm permits. Such results now exist for l_1-minimization, where precise phase transitions have been determined within a proportional-growth asymptotic framework in the case of Gaussian matrices (Section 1.5) [77]. By contrast, both the worst-case and average-case sparse recovery properties of l_0-based algorithms such as IHT and NIHT are much less well understood. In this thesis (Chapter 2), we quantified existing state-of-the-art RIP-based recovery analysis for IHT algorithms within the phase transition framework for Gaussian matrices. Our phase transitions demonstrate that current RIP analysis of IHT algorithms is pessimistic compared to empirically observed performance.

To address this issue, we introduced a new method of recovery analysis for IHT algorithms in which we analysed the algorithms' *stable points*, a generalization of the notion of fixed points (Chapter 3). Our approach consists of two parts: we analyse necessary conditions for the existence of a stable point on a given support (the stable point condition), and meanwhile we derive conditions ensuring convergence to one of its stable points. By analysing the stable point condition in the case of Gaussian matrices under the realistic assumption of independence between the signal and measurement matrix (Chapter 4), we obtained the first average-case recovery guarantees for IHT algorithms in the phase transition framework (Chapter 5). In contrast to RIP analysis, which leads to lower bounds on the strong phase transition (Section 1.5), we obtained lower bounds on a weak phase transition for recovery using IHT algorithms, which is the notion of practical interest. By breaking free in part from the restrictions of worst-case analysis,

we have obtained the highest phase transitions yet guaranteeing exact recovery of sparse signals by means of IHT and NIHT. Our results extend to the realistic model of noisy measurements of compressible signals, guaranteeing an improved robustness to these inaccuracies.

The ultimate remaining goal of the work is to fully close the gap between theoretical guarantees and empirical performance for IHT algorithms. Recalling that our analysis consists of two parts, one of the main reasons for the pessimism of our results is the continued use of the RIP in the convergence analysis. A research direction for the future would be to improve upon the convergence analysis, ideally by making use of average-case assumptions in this case also. More generally, the continued use of worst-case methods of analysis such as union bounds over combinatorially many support sets is at present a hindrance to significant further improvements in phase transitions. For example, such union bounds are required by the present analysis of in order to show that, in the idealized noiseless case, IHT has no spurious fixed points. This raises the question as to whether such a strong requirement is necessary for ensuring signal recovery in practice, which also represents a potential research direction for the future. Though we have obtained quantitative results only for Gaussian matrices here, many other families of random or randomized measurement matrices exhibit similar empirical behaviour and are important to practitioners. Obtaining quantitative guarantees for IHT algorithms applied to such CS measurement schemes is yet another possible future direction.

Interest continues to grow in the CS community in more refined sparsity models, such as the tree-based model for wavelet-based signals. IHT algorithms can be easily adapted to solve the tree-based problem by replacing the hard threshold with a tree projection, leading to ITP algorithms (Section 6.1) [24, 7]. However, none of the tree projection algorithms in the current literature is guaranteed to calculate an exact tree projection for a given sparsity, which is an assumption underpinning all current recovery analysis of ITP algorithms. We proposed a dynamic programming (DP) algorithm which is the first to have such a guarantee, and we also showed that it has low-order polynomial complexity (Section 6.2). By extending our stable point approach to the tree-based setting (Chapter 6), we obtained quantitative average-case results for ITP algorithms in the context of Gaussian matrices (Chapter 7). Since the number of support sets is much reduced, union bound arguments were tightened, leading to improved results. We obtained results in a simplified proportional-growth asymptotic, in which recovery depends only upon the oversampling factor $\rho = k/n$.

To the best of our knowledge, little is currently known empirically about average-case oversampling thresholds for ITP algorithms, and indeed it is not obvious what would be an appropriate signal model on which to base numerical tests. The empirical investigation of the benefits of switching to the tree-based setting remains an area for future investigation. This thesis illustrates the transferability of recovery analysis for algorithms for sparse and tree-based recovery, and it is likely that future advances in theoretical analysis of IHT algorithms will also translate into corresponding advances for ITP algorithms.

Bibliography

[1] M. Abramowitz and I. Stegun. *A handbook of mathematical functions, with formulas, graphs and mathematical tables*. Dover, 9th edition, 1964.

[2] F. Affentrager and R. Schneider. Random projections of regular simplices. *Discrete computational geometry*, 7(3):219–226, 1992.

[3] W. Ambarzumyan. Uber eine frage der eigenwerttheorie. *Zeitschrift fur Physik*, 53(9–10):690–695, 1929.

[4] B. Bah and J. Tanner. Improved bounds on restricted isometry constants for gaussian matrices. *SIAM Journal on Matrix Analysis*, 31(5):2882–2898, 2010.

[5] R. Baraniuk. Optimal tree approximation with wavelets. In *Proceedings of the SPIE Technical Conference on Wavelet Applications in Signal Processing VII*, 1999.

[6] R. Baraniuk. More is less: signal processing and the data deluge. *Science*, 331(6018):717–719, 2011.

[7] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde. Model-based compressive sensing. *IEEE Transactions in Information Theory*, 56(4):1982–2001, 2010.

[8] R. Baraniuk and D. Jones. A signal-dependent time-frequency representation: Fast algorithm for optimal kernel design. *IEEE Transactions on Signal Processing*, 42(12):3530–3535, 1994.

[9] Y. Baryshnikov and R. Vitale. Regular simplices and gaussian samples. *Discrete Computational Geometry*, 11(2):141–147, 1994.

[10] M. Bayati and A. Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Transactions on Information Theory*, 57(2):764–785, 2011.

[11] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[12] S. Becker, J. Bobin, and E. Candès. A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, 4(1):1–39, 2011.

[13] R. Bellman. *Dynamic Programming*. Courier Dover Publications, 2003.

[14] C. Berger, Z. Wang, J. Huang, and S. Zhou. Application of compressive sensing to sparse channel estimation. *IEEE Communications Magazine*, 48(11):164–174, 2010.

[15] J. Bioucas-Dias and M. Figueiredo. A new twist: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image Processing*, 16(12):2992–3004, 2007.

[16] J. Blanchard, C. Cartis, J. Tanner, and A. Thompson. Phase transitions for greedy sparse approximation algorithms; extended technical report. Technical Report ERGO 09-010, School of Mathematics, University of Edinburgh, 2009.

[17] J. Blanchard, C. Cartis, J. Tanner, and A. Thompson. Phase transitions for greedy sparse approximation algorithms. *Applied and Computational Harmonic Analysis*, 30(2):188–203, 2011.

[18] J. Blanchard and J. Tanner. Gpu accelerated greedy algorithms for compressed sensing. `http://ecos.maths.ed.ac.uk/publications.shtml`, 2012.

[19] J. Blanchard and A. Thompson. On support sizes of restricted isometry constants. *Applied and Computational Harmonic Analysis*, 29(3):382–390, 2010.

[20] J. Blanchard and A. Thompson. Pushing the rip phase transition in compressed sensing. In *Proceedings of the European Signal Processing Conference*, 2010.

[21] Jeffrey. D. Blanchard, Coralia Cartis, and Jared Tanner. Compressed sensing: How sharp is the restricted isometry property? *SIAM Review*, 53(1):105–125, 2011.

[22] T. Blumensath and M. Davies. Gradient pursuits. *IEEE Transactions on Signal Processing*, 56(6):2370–2382, 2008.

[23] T. Blumensath and M. Davies. Iterative hard thresholding for compressed sensing. *Applied and Computational Harmonic Analysis*, 27(3):265–274, 2009.

[24] T. Blumensath and M. Davies. Sampling theorems for signals from the union of finite-dimensional linear subspaces. *IEEE Transactions on Information Theory*, 55(4), 2009.

[25] T. Blumensath and M. Davies. Stagewise weak gradient pursuits. *IEEE Transactions on Signal Processing*, 57(11):4333–4346, 2009.

[26] T. Blumensath and M.E. Davies. Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14(5):629–654, 2008.

[27] T. Blumensath and M.E. Davies. Normalized iterative hard thresholding: guaranteed stability and performance. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):298–309, 2010.

[28] J. Bobin, J. Starck, and R. Ottensamer. Compressed sensing in astronomy. *IEEE Journal of Selected Topics in Signal Processing*, 2(5):718–726, 2008.

[29] K. Böröczky and M. Henk. Random projections of regular polytopes. *Archiv der Mathematik*, 73(6):465–473, 1999.

[30] T. Cai, L. Wang, and G. Xu. New bounds for restricted isometry constants. *IEEE Transactions on Information Theory*, 56(9):4388–4394, 2010.

[31] T. Cai, L. Wang, and G. Xu. Shifting inequality and recovery of sparse signals. *IEEE Transactions on Signal Processing*, 58(3):1300–1308, 2010.

[32] E. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes rendus de l'Académie des Sciences*, 346:589–592, 2008.

[33] E. Candès, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling and Simulation*, 5(3), 2006.

[34] E. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the Association for Computing Machinery*, 58(3):1–37, 2011.

[35] E. Candès and J. Romberg. l_1-magic: recovery of sparse signals via convex programming. Technical report, California Institute of Technology, 2005.

[36] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction form highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[37] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[38] E. Candès, M. Rudelson, T. Tao, and R. Vershynin. Error correction via linear programming. In *Proceedings of the 46th IEEE Symposium on Foundations of Computer Science*, pages 295–308, 2005.

[39] E. Candès and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005.

[40] C. Caratheodory. über den varaibilitätsbereich der köffizienten von potenzreihen, die gegebene werte nicht annehmen. *Mathematische Annalen*, 64:95–115, 1907.

[41] C. Caratheodory. über den varaibilitätsbereich der fourierschen konstanten von positiven harmonischen funktionen. *Rendiconti del Circolo Matematico di Palermo*, 32:193–217, 1911.

[42] C. Cartis and A. Thompson. A new and improved recovery analysis for iterative hard thresholding algorithms in compressed sensing. in preparation, 2012.

[43] C. Cartis and A. Thompson. Quantitative recovery guarantees for iterative tree projection algorithms. in preparation, 2012.

[44] A. Cauchy. Méthodes générales pour la rśolution des systèmes d'équations simultanées. *Comptes Rendus de l'Académie to Science, Paris*, 25:536–538, 1847.

[45] R. Chartrand. Exact reconstructions of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters*, 14(10):707–710, 2007.

[46] R. Chartrand. Nonconvex compressed sensing and error correction. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007.

[47] R. Chartrand. Fast algorithms for nonconvex compressive sensing: Mri reconstruction from very few data. In *IEEE International Symposium on Biomedical Imaging*, 2009.

[48] R. Chartrand and V. Staneva. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems*, 24(3):1–14, 2008.

[49] R. Chartrand and W. Yin. Iteratively reweighted algorithms for compressed sensing. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.

[50] J. Chen and X. Huo. Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Transactions on Signal Processing*, 54(12):4634–4643, 2006.

[51] S. Chen, S. Billings, and W. Luo. Orthogonal least squares methods and their application to nonlinear system identification. *International Journal of Control*, 50(5):1873–1896, 1989.

[52] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal of Scientific Computing*, 20(1):33–61, 1999.

[53] J. Claerbout and F. Muir. Robust modelling with erratic data. *Geophysics*, 38(5):826–844, 1973.

[54] P. Combettes and V. Wajs. Signal recovery by proximal forward-backward splitting. *Multisacle Modeling and Simulation*, 4(4):1168–1200, 2005.

[55] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing*, 51(7):2477–2488, 2005.

[56] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Transactions on Information Theory*, 55(5):2230–2249, 2009.

[57] G. Dantzig. *Linear programming and extensions.* Princeton University Press and the RAND Corporation, 1963.

[58] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7), 1988.

[59] I. Daubechies. *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics, 1992.

[60] 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*, 57(11):1413–1457, 2004.

[61] I. Daubechies, R. DeVore, M. Fornasier, and C. Güntürk. Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics*, 63(1):1–38, 2010.

[62] G. Davis, S. Mallat, and Z. Zhang. Adaptive time-frequency decompositions. *Optical engineering*, 33:2183–2191, 1994.

[63] D. Donoho. Cart and best ortho-basis: A connection. *Annals of Statistics*, 25(5):1870–1911, 1997.

[64] D. Donoho. For most large underdetermined systems of linear equations the minimal norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59:797–829, 2004.

[65] D. Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. Technical report, Computational Geometry (online), 2005.

[66] D. Donoho. Neighborly polytopes and sparse solutions of underdetermined linear equations. Technical report, Stanford University, 2005.

[67] D. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[68] D. Donoho. For most large underdetermined systems of equations the minimal l_1-norm solution approximatesthe sparsest near-solution. *Communications on Pure and Applied Mathematics*, 59(7):907–934, 2006.

[69] D. Donoho, N. Dyn, D. Levin, and T. Yu. Smooth multiwavelet duals of alpert bases by moment-interpolating refinement. *Applied and Computational Harmonic Analysis*, 9(2):166–203, 2000.

[70] D. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization. *Proceedings of the National Academy of Science*, 100(5):2197–2202, 2003.

[71] D. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47(7):2845–2862, 2001.

[72] D. Donoho and A. Maleki. Optimally tuned iterative reconstruction algorithms for compressed sensning. *Journal of Selected Topics in Signal Processing*, 4(2):330–341, 2010.

[73] D. Donoho, A. Maleki, and A. Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

[74] D. Donoho and P. Stark. Uncertainty principles and signal recovery. *SIAM Journal on Applied Mathematics*, 49(3):906–931, 1989.

[75] D. Donoho and J. Tanner. Neighborliness of randomly-projected simplices in high dimensions. *Proceedings of the National Academy of Sciences*, 102(27):9452–9457, 2005.

[76] D. Donoho and J. Tanner. Sparse nonnegative solutions of underdetermined linear equations by linear programming. *Proceedings of the National Academy of Sciences*, 102(27):9446–9451, 2005.

[77] D. Donoho and J. Tanner. Counting faces of randomly projected polytopes when the projection radically lowers dimensions. *Journal of the Americal Mathematical Society*, 22(1):1–53, 2009.

[78] D. Donoho and J. Tanner. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Philosophical Transactions of the Royal Society*, 367:4273–4293, 2009.

[79] D. Donoho and J. Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete and Computational Geometry*, 43(3):522–541, 2010.

[80] D. Donoho, Y. Tsaig, I. Drori, and J. Starck. Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit. Technical report, Stanford University, 2006.

[81] M. Duarte and R. Baraniuk. Spectral compressive sensing. `http://dsp.rice.edu/publications/spectral-compressive-sensing`, 2011.

[82] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, 2008.

[83] M. Duarte, M. Wakin, and R. Baraniuk. Fast reconstruction of piecewise smooth signals from random projections. In *Signal Processing with Adaptive Sparse Structured Representations*, 2005.

[84] M. Duarte, M. Wakin, and R. Baraniuk. Wavelet-domain compressive signal reconstruction using a hidden markov tree model. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5137–5140, 2008.

[85] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM Journal on Matrix Analysis and Applications*, 9(4), 1988.

[86] E. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *Annals of Statistics*, 32:407–499, 2004.

[87] M. Efroymson. *Multiple Regression Analysis*, volume 1 of *Mathematical Methods for Digital Computers*, pages 191–203. John Wiley and Sons, 1960.

[88] Y. Eldar, P. Kuppinger, and H. Bölcskei. Block-sparse signals: uncertainty relations and efficient recovery. *IEEE Transactions on Signal Processing*, 58(6):3042–3054, 2010.

[89] Y. Eldar and M. Mishali. Robust recovery of signals from a structured union of subspaces. *IEEE Transactions on Information Theory*, 55(11):5302–5316, 2009.

[90] M. Figueiredo and R. Nowak. An em algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 12(8):906–916, 2003.

[91] M. Figueiredo, R. Nowak, and S. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE J. Selected Topics in Signal Processing*, 1(4), 2007.

[92] S. Foucart. *Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants*, volume XIII of *Springer Proceedings in Mathematics: Approximation Theory*, pages 65–77. San Antonio, 2010.

[93] S. Foucart. Hard thresholding pursuit: an algorithm for compressive sensing. *SIAM Journal on Numerical Analysis*, 49(6):2543–2563, 2011.

[94] S. Foucart and M. Lai. Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$. *Applied and Computational Harmonic Analysis*, 26(3):395–407, 2009.

[95] J. Friedman and W. Stuetzle. Projection pursuit regressions. *Journal of the American Statistical Association*, 76:817–823, 1981.

[96] R. Garg and R. Khandekar. Gradient descent with sparsification: an iterative algorithm for sparse recovery with restricted isometry property. In *Proceedings of the 26th Annual Conference on Machine Learning*, 2009.

[97] M. Garside. The best subset in multiple regression analysis. *Applied Statistics*, 14:196–200, 1965.

[98] A. Gilbert, S. Muthukrishnan, and M. Strauss. Approximation of functions over redundant dictionaries using coherence. In *Proceedings of the Symposium on Discrete Algorithms*, pages 243–252, 2003.

[99] P. Gill, W. Murray, D. Ponceléon, and M. Saunders. Solving reduced kkt systems in barrier methods for linear and quadratic programming. Technical Report SOL 91-7, Stanford University, Standford CA, 1991.

[100] A. Goldstein. Convex programming in hilbert space. *Bulletin of the Americal Mathematical Society*, 70:709–710, 1964.

[101] I. Gorodnitsky and B. Rao. Sparse signal reconstruction from limited data using focuss: a recursive weighted norm minimization algorithm. *IEEE Transactions on Signal Processing*, 45(3):600–616, 1997.

[102] R. Gribonval and M. Nielson. Sparse representations in unions of bases. *IEEE Transactions on Information Theory*, 49(12):3320–3325, 2003.

[103] A. Gupta and D. Nagar. *Matrix Variate Distributions*. Chapman & Hall/CRC, 1999.

[104] A. Haar. Zur theorie der orthogonalen funktionensysteme, (erste mitteilung). *Mathematische Annalen*, 69, 1910.

[105] J. Hadamard. Sur les problemes aux derives partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.

[106] E. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for l_1-minimization: methodology and convergence. *SIAM Journal on Optimization*, 19(3):1107–1130, 2008.

[107] K. Herrity, A. Gilbert, and J. Tropp. Sparse approximation via iterative thresholding. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, 2006.

[108] S. Heubach and T. Mansour. *Combinatorics of Compositions and Words*. CRC Press, 2009.

[109] P. Jain, A. Tewari, and I. Dhillon. Orthogonal matching pursuit with replacement. *Advances in Neural Information Processing Systems*, 24:1215–1223, 2011.

[110] N. Karahanoglu and H. Erdogan. A* orthogonal matching pursuit: best-first search for compressed sensing signal recovery. *Digital Signal Processing*, 22(4):555–568, 2012.

[111] A. Khajehnejad, W. Xu, S. Avestimehr, and B. Hassibi. Analysing weighted l_1-minimization for sparse recovery with nonuniform sparse models. *IEEE Transactions on Signal Processing*, 59(5):1985–2001, 2011.

[112] S-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale l_1-regularized least squares. *IEEE Journal on Selected Topics in Signal Processing*, 1(4):606–617, 2007.

[113] C. La and M. Do. Signal reconstruction using sparse tree representation. In *SPIE Optics and Photonics*, volume Wavelets XI, 2005.

[114] L. Landweber. An iterative formula for fredholm integrals of the first kind. *Americal Journal of Mathematics*, 73:615–624, 1951.

[115] W. Lang. Combinatorial interpretation of generalized stirling numbers. *Journal of Integer Sequences*, 12, 2009.

[116] C. Lawson. *Contributions to the theory of linear least maximum approximation.* PhD thesis, University of California, Los Angeles, 1961.

[117] S. Levy and P. Fullagar. Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution. *Geophysics*, 46(9):1235–1243, 1981.

[118] W. Lim. The discrete shearlet transform: a new directional transform and compactly supported shearlet frames. *IEEE Transactions on Image Processing*, 19, 2010.

[119] T. Lin and F. Herrmann. Compressed wavefield extrapolation. *Geophysics*, 72(5):77–93, 2007.

[120] M. Lustig, D. Donoho, J. Santos, and J. Pauly. Compressed sensing mri. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.

[121] D. Malioutov, M. Çetin, and A. Willsky. Homotopy continuation for sparse signal representation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 5:733–736, 2005.

[122] S. Mallat. Multiresolution approximation and wavelets. *Transactions of the American Mathematical Society*, 315, 1989.

[123] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, San Diego, California, third edition, 2009.

[124] S. Mallat and Z. Zhang. Matching pursuit in a time-frequency dictionary. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993.

[125] Y. Meyer. Ondelettes, fonctions splies et analyses graduées. *Seminario Matematico Università Politecnico di Torina*, 45, 1988.

[126] M. Mishali and Y. Eldar. Blind multi-band signal reconstruction: compressed sensing for analog signals. *IEEE Transactions on Signal Processing*, 57(30):993–1009, 2009.

[127] D. Needell and J. Tropp. Cosamp: iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2008.

[128] D. Needell and R. Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Foundations of Computational Mathematics*, 9(3):317–334, 2009.

[129] Y. Nesterov. A method for unconstrained convex problem with the rate of convergence $\mathcal{O}(1/k^2)$. *Doklady AN USSR*, 269:543–547, 1983.

[130] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 1999.

[131] H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the AIEE*, 47:617–644, 1928.

[132] M. Osborne, B. Presnell, and B. Turlach. A new approach to variable selection in least squares problems. *IMA J. Numerical Analysis*, 20:389–403, 2000.

[133] A. Papoulis and C. Chamzas. Improvement of range resolution by spectral extrapolation. *Ultrasonic Imaging*, 1:121–135, 1979.

[134] V. Patel, G. Easley, D. Healy, and R. Chellappa. Compressed synthetic aperture radar. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):244–254, 2010.

[135] M. Piana and M. Bertero. Projected landweber method and preconditioning. *Inverse problems*, 13:441–463, 1997.

[136] E. Rainville. *Special functions*. Macmillan, 1960.

[137] B. Rao. Analysis and extensions of the focuss algorithm. In *Proceedings of the IEEE Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1218–1223, 1996.

[138] F. Santosa and W. Symes. Inversion of impedance profile from band-limited data. In *Digest, International Geoscience and Remote Sensing Symposium*, 1983.

[139] C. Shannon. Communication in the presence of noise. *Proceedings of the Institute of Radio Engineers*, 37(1):10–21, 1949.

[140] J. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, 1993.

[141] A. Skodras, C. Christopoulos, and T. Ebrahimi. The jpeg2000 still image compression standard. *IEEE Signal Processing Magazine*, 18(5):36–58, 2001.

[142] R. Stanley. *Enumerative Combinatorics*. Cambridge University Press, 2nd edition, 2012.

[143] M. Stojnic, F. Parvaresh, and B. Hassibi. On the reconstruction of block-sparse signals with an optimal number of measurements. *IEEE Transactions on Signal Processing*, 57(8):3075–3085, 2009.

[144] J. Strömberg. A modified franklin system and higher-order spline systems on ρ^n as unconditional bases for hardy spaces. In *Conference on Harmonic Analysis in honour of Antoni Zygmund*, 1983.

[145] B. Sturm and M. Christenson. Cycling matching pursuits with multiscale time-frequency dictionaries. In *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, 2010.

[146] V. Temlyakov. A remark on simultaneous sparse approximation. *East Journal on Approximation*, 100:17–25, 2004.

[147] A. Thompson. Optimization aspects of sparse approximation. Master's thesis, School of Mathematics, University of Edinburgh, 2008.

[148] A. Thompson. End of 1st year report. School of Mathematics, University of Edinburgh, 2009.

[149] A. Thompson. Compressive single-pixel imaging. Technical Report ERGO 11-006, School of Mathematics, University of Edinburgh, 2011.

[150] A. Thompson. Compressive single-pixel imaging. In *2nd IMA Conference on Mathematics in Defence*, 2011.

[151] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 58(1):267–288, 1996.

[152] A. Tikhonov and V. Arsenin. *Solution of ill-posed problems*. Wiley, New York, 1977.

[153] J. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231–2242, 2004.

[154] J. Tropp and A. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, 2007.

[155] J. Tropp, A. Gilbert, and M. Strauss. Algorithms for simultaneous sparse approximation. part i: Greedy pursuit. *Signal Processing*, 86(3):572–588, 2006.

[156] J. Tropp, A. Gilbert, and M. Strauss. Algorithms for simultaneous sparse approximation. part ii: Convex relaxation. *Signal Processing*, 86(3):589–602, 2006.

[157] E. Van den Berg and M. Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computation*, 31(2):890–912, 2008.

[158] B. Varadarajan, S. Khudanpur, and T. Tran. Stepwise optimal subspace pursuit for improved sparse recovery. *IEEE Signal Processing Letters*, 18(1):27–30, 2011.

[159] A. Vershik and P. Sporyshev. Asymptotic behaviour of the number of faces of random polyhedra and the neighbourliness problem. *Selecta Mathematika Sovietica*, 11(2):181–201, 1992.

[160] D. Wipf and S. Nagarajan. Iterative reweighted l_1 and l_2 methods for finding sparse solutions. *Journal of Selected Topics in Signal Processing*, 4(2):1–29, 2010.

[161] S. Wright, R. Nowak, and M. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493, 2009.