# NEW THEORY AND METHODS IN ADAPTIVE AND COMPRESSIVE SAMPLING FOR SPARSE DISCOVERY

by

Jarvis David Haupt

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Electrical Engineering)

at the

**UNIVERSITY OF WISCONSIN – MADISON**

2009

Dedicated to the memory of Bob Fogarty—a truly special engineer, mentor, and friend.

# ACKNOWLEDGEMENTS

I believe that no accomplishment should come without acknowledgements, and no glory or accolade should be accepted without due deference. Similar sentiments have been echoed many times, perhaps most notably by Sir Isaac Newton, who wrote "If I see a little further it is by standing on the shoulders of giants." It is in this spirit that I wish to thank the many who enabled, supported, and assisted me in achieving this professional and personal milestone.

I begin by thanking God for providing me the ability, and the many protectors and defenders of freedom and liberty in the United States of America for the opportunity, to pursue my dreams.

They say that you cannot choose your family, but had I the choice, the outcome would have been the same. I would like to thank my parents, whose love for each other and our family is truly inspiring. In particular, I would like to thank my mother Julie for always making me feel proud of my accomplishments and teaching me true compassion, and my father David for teaching me the inherent value of analytical thought and instilling in me the desire to perpetually learn and evolve, create and improve.

Thank you to my sister Rachael for showing me that true happiness exists in the world and that it can always be found by those who look for it, and to my sister Bridget for always looking out for me and for teaching me to take the realities of life in stride. I would also like to thank my brother Travis for his support and guidance, for teaching me to take a realistic approach to life and its issues, and for being the kind of person that am both proud and humbled to call my best friend.

I would like to thank my nieces Brynn and Kylee Harrington, and my nephew Aiden Haupt, who constantly remind me to take time to enjoy the little pleasures in life.

I owe a debt of tremendous gratitude to my Ph.D. advisor, Rob Nowak, for believing in me, for sharing his wisdom and providing guidance and support, for the opportunity to work and grow together, and for the countless experiences and opportunities he enabled and provided along the way. I would also like to acknowledge and express my gratitude for financial support from the National Science Foundation, the Office of Naval Research, and the Defense Advanced Research Projects Agency.

Thank you to Bill Sethares for his guidance during my first years in graduate school, for his mentoring throughout my masters degree pursuit, and for serving on my Ph.D. committee. I would also like to thank

the other members of my committee—Amos Ron, Akbar Sayeed, and Barry Van Veen—for their time and consideration, and for many enjoyable interactions and discussions during the past few years.

I would like to thank my office mates—Rebecca Willett, Mike Rabbat, Yolanda Tsang, Aline Martin, Rui Castro, Waheed Bajwa, Aarti Singh, Brian Eriksson, Laura Balzano, Jared Doot, Tyler Karrels, Kimiko Parker and Lorne Applebaum—for countless entertaining, enlightening, and productive discussions on research problems (and many unrelated topics).

I would like to thank Steve Meitner for demonstrating to me the power of perseverance and dedication, Matt Rickl for showing me that worrying is not a productive strategy for solving any problem, Ron Klein for showing me that someone can be truly devoted to something larger than himself, Brian Weinfurter for demonstrating to me the value of succinctness, Brent Ziegler for providing a nearly endless source of entertainment, and Nick Marti for being a great friend and for showing me the virtues of a slow temper and a genuine smile.

Thank you to Gil Raz for enjoyable and productive discussions and collaborations on our many shared interests, and for encouraging me to think big(ger) but work smart(er).

I also thank Jerry Fudge for the opportunity to work with him as an intern at L-3 Communications in Greenville, TX. I cannot imagine a better place to take up distance running than in 90 degree weather in the Texas summer. Thank you to my friends and colleagues at Cray, who provided me opportunities to simultaneously hone my computer architecture and foosball skills during my two internships there.

I want to thank Phil Wettstein for showing me that things do not always have to be done according to someone else's template, and that standing for principles is the right option even when it places you squarely in the minority. Thank you also to Pat (Thompson) Johnson, whose teaching inspired in me a "eureka moment," to which I largely attribute my understanding of, and deep appreciation for, mathematics.

I would also like to thank Bob Fogarty, to whom this dissertation is dedicated, for believing in me for my potential, for teaching me to think holistically, and for always encouraging me to think outside the box.

Finally, I want to express my deepest thanks to Christina Gamaldi for being my inspiration, for showing me that true love is synonymous with true support and true devotion, for loving me for who I am, and for encouraging me to be all I can.

# TABLE OF CONTENTS

Page

# LIST OF FIGURES

# ABSTRACT

The study of sparsity has recently garnered significant attention in the signal processing and statistics communities. Generally speaking, sparsity describes the phenomenon where a large data set may be succinctly represented or approximated using only a small number of summary values or coefficients. The implications are clear—the presence of sparsity suggests the potential for efficient methods to extract only the relevant information, conserving acquisition and/or processing resources which can often be scarce or expensive. The overall theme of this work is the identification of efficient and effective ways to exploit sparsity in a variety of settings and applications.

The first part of this work comprises contributions to the emerging theory of compressive sampling (also called compressed sensing or compressive sensing). Compressive sampling describes a framework under which sparse, high-dimensional signals (vectors) can be recovered from a relatively small number of non-adaptive observations. The theory of compressive sampling is extended here, where it is shown that compressive sampling can be an effective tool for sparse recovery in noisy environments, in applications in sensor networking, system identification (channel sensing), and wideband RF surveillance.

The second part of this work examines adaptivity in sampling. Adaptive sampling strategies are those which direct subsequent observations based on the results of previous observations, in an effort to focus on features of interest. It is shown here that adaptivity can result in dramatic improvements in recoverability of sparse signals, providing new insight into the fundamental theoretical limits of sparse recovery in noisy settings. A simple adaptive procedure called *distilled sensing* is proposed and shown to dramatically outperform the best possible non-adaptive strategies, in the sense that distilled sensing enables the recovery (detection and estimation) of signals whose features are otherwise too weak to be recoverable using any methods based on the best non-adaptive sampling strategies.

**NEW THEORY AND METHODS IN**
**ADAPTIVE AND COMPRESSIVE SAMPLING FOR SPARSE DISCOVERY**

Jarvis David Haupt

Under the supervision of Professor Robert D. Nowak
At the University of Wisconsin - Madison

The study of sparsity has recently garnered significant attention in the signal processing and statistics communities. Generally speaking, sparsity describes the phenomenon where a large data set may be succinctly represented or approximated using only a small number of summary values or coefficients. The implications are clear—the presence of sparsity suggests the potential for efficient methods to extract only the relevant information, conserving acquisition and/or processing resources which can often be scarce or expensive. The overall theme of this work is the identification of efficient and effective ways to exploit sparsity in a variety of settings and applications.

The first part of this work comprises contributions to the emerging theory of compressive sampling (also called compressed sensing or compressive sensing). Compressive sampling describes a framework under which sparse, high-dimensional signals (vectors) can be recovered from a relatively small number of non-adaptive observations. The theory of compressive sampling is extended here, where it is shown that compressive sampling can be an effective tool for sparse recovery in noisy environments, in applications in sensor networking, system identification (channel sensing), and wideband RF surveillance.

The second part of this work examines adaptivity in sampling. Adaptive sampling strategies are those which direct subsequent observations based on the results of previous observations, in an effort to focus on features of interest. It is shown here that adaptivity can result in dramatic improvements in recoverability of sparse signals, providing new insight into the fundamental theoretical limits of sparse recovery in noisy settings. A simple adaptive procedure called *distilled sensing* is proposed and shown to dramatically outperform the best possible non-adaptive strategies, in the sense that distilled sensing enables the recovery (detection and estimation) of signals whose features are otherwise too weak to be recoverable using any methods based on the best non-adaptive sampling strategies.

Approved:

Professor Robert D. Nowak
Department of Electrical and Computer Engineering
University of Wisconsin - Madison

# Chapter 1

# Introduction

The study of sparsity in data-rich applications has recently garnered significant attention in the signal processing and statistics communities. Generally speaking, sparsity describes the phenomenon where a large collection of data may be succinctly represented using only a small number of summary values or coefficients. The implications are clear—the presence of sparsity suggests the potential for efficient methods to extract relevant information without wasting often-scarce acquisition or processing resources. Identifying ways to effectively exploit sparsity has become a primary focus in many application domains, including network science, imaging, genomics, and communications.

For example, sensor networking involves the deployment of many sensor nodes (hundreds or thousands, or more), each equipped with modest sensing capabilities and limited battery life, with the goal of performing distributed inference. In monitoring applications, the data sensed collectively by the nodes often correspond to observations of a natural process (diffusion of a chemical or biological agent, acoustic or vibrational measurements, temperature gradients, etc.), and the inherent smoothness of the "target signals" suggests they can be described using a small number of Fourier coefficients. Sparsity in the frequency domain is also present in applications in signals intelligence, where waveforms of interest are often narrow-band communication or radar signals. In emerging technologies such as cognitive radio, the ability to efficiently estimate time-varying communication channels is crucial, and in many environments these channels exhibit sparsity in a time-frequency representation. A wide variety of applications where sparsity is salient exist also in the life sciences—for example, in the identification of differentially-expressed genes in high-throughput microarray experiments. Applications in astronomy, cosmology, bioimaging, and remote sensing are also easy to envision.

In each of these applications, procedures that exploit sparsity can result in significant improvements (real-time processing, reduced exposure, improved imaging capabilities, etc.), and increased efficiency, for example, in the form of reduced operational cost. The general focus of this research involves the theoretical treatment of techniques that exploit sparsity in high-dimensional recovery problems, specifically utilizing observation models that provide flexibility in the data acquisition process, including aggregation sampling

models (as in *compressive sampling*, where observations are linear summaries of the available data), and models that permit sampling with feedback (also called *adaptive sampling*). Advances in these two areas comprise the main contributions of this work.

The first part of this work details contributions to the theory of compressive sampling (CS). The canonical CS observation model allows samples in the form of projections of the unknown signal (a sparse high-dimensional vector) onto a set of specially constructed, non-adaptive "test vectors." Initial breakthrough theoretical results established that a relatively small number of such observations, formed by projecting the signal onto a collection of random test vectors, can be utilized to exactly recover the unknown signal with high probability, provided the signal is sparse in some appropriate representation and the observations are noise-free [1–3]. Our own work provided the first theoretical analysis of, and reconstruction error bounds for, CS when observations are corrupted by additive Gaussian noise [4]. This work appears here in Section 2.1. In a subsequent effort we considered CS in the context of random probing for sparse system identification, developing a novel technique that allows theoretical treatment of CS in settings where the entries of the test vectors exhibit structured statistical dependencies—a marked departure from initial efforts, which primarily treated the case of random test vectors with fully-independent entries [2, 3]. It is shown in [5] that CS can be an effective tool in linear, time-invariant (LTI) system identification problems, making it particularly applicable in channel sensing applications inherent in cognitive radio. That work appears here in Section 2.2. The last part of Chapter 2, Section 2.3, represents our recent effort to identify ensembles of deterministic test vectors, comprised of specially-selected rows of Fourier matrices, which can be used in CS applications. It follows from this work that a structured, non-uniform time domain sampling strategy can be utilized to effectively and efficiently recover time-varying signals that have sparse spectral content over a large bandwidth. This effort took place in the context of the Defense Advanced Research Projects Agency Analog-to-Information (DARPA A2I) program, an initiative whose stated goal is to advance the limits of analog-to-digital technology by developing efficient methods for acquiring signals in sparse broadband RF environments.

The second part of the dissertation details advances in adaptive sampling for recovery of sparse signals in noise, and provides significant new insight into the fundamental theoretical limits of sparse signal recovery. The overall theme of Chapter 3 is that a simple adaptive sampling procedure, in which subsequent observations depend on the outcomes of prior measurements, can dramatically outperform the *best possible* nonadaptive sampling strategies, in a variety of recovery tasks. Motivated by the problem of detecting sparse signals in noise, the adaptive procedure proposed in [6] was shown to enable detection of sparse signals whose nonzero entries have significantly smaller amplitudes than what are required for the signal to be detectable by any nonadaptive method. This work appears here in Section 3.1. A follow-on effort, motivated by the problem of support recovery—identifying the the locations of the nonzero entries of a sparse signal—appeared in [7]. In that work, it was shown that a refinement of the simple adaptive sampling strategy, now termed

*distilled sensing*, enabled even more dramatic improvements in the thresholds of recoverability. This work appears here in Section 3.2. These approaches are generally applicable in settings that allow a tradeoff between resolution and accuracy, having been proposed as viable strategies in a variety of applications including microarray analysis, astronomical imaging, sensor networking, and cognitive radio.

In the final portion, Chapter 4, several directions for future study based on extensions, refinements, and combinations of the works presented in Chapters 2 and 3 are discussed.

# Chapter 2

# Advances in Compressive Sensing

This chapter contains three contributions to the theory of compressive sensing. Section 2.1 is a reprint of our journal article [4], which appeared in the IEEE Transactions in Information Theory and provided the first theoretical analysis of compressive sensing in settings where the observations were corrupted by additive Gaussian noise.

Section 2.2 examines the utility of compressive sensing methods (specifically, random probing) in linear, time-invariant system identification problems. A primary contribution here is the development of a novel technique that enables theoretical analysis in such settings, where the convolutional nature of the observation model introduces structured statistical dependencies among observations. It is shown that CS can be an effective tool in linear, time-invariant (LTI) system identification problems, making it particularly applicable in channel sensing applications inherent to cognitive radio. This work has been submitted to the IEEE Transactions on Information Theory, and is currently under review [5].

Section 2.3 establishes a new class of deterministic compressive sensing measurement matrices comprised of rows of Fourier matrices. One application of these results is in wideband surveillance; it follows from the work presented in Section 2.3 that a structured, non-uniform time domain sampling strategy can be utilized to effectively and efficiently recover time-varying signals that have sparse spectral content over a large bandwidth.

Note that each section is essentially a stand-alone entity, so there may be some overlap in introductory material throughout. In addition, while the notation is consistent within each section, it may vary across sections.

## 2.1 Signal Reconstruction from Noisy Random Projections

Recent results show that a relatively small number of random projections of a signal can contain most of its salient information. It follows that if a signal is compressible in some orthonormal basis, then a very accurate reconstruction can be obtained from random projections. This "compressive sampling" approach is extended here to show that signals can be accurately recovered from random projections contaminated with noise. A practical iterative algorithm for signal reconstruction is proposed, and potential applications to coding, A/D conversion, and remote wireless sensing are discussed.[1]

### 2.1.1 Introduction

Recent theory informs us that, with high probability, a relatively small number of random projections of a signal can contain most of its relevant information. For example, the groundbreaking work in [1] has shown that $k$ random Fourier projections contain enough information to reconstruct piecewise smooth signals at a distortion level nearly equivalent to that attainable from $k$ optimally selected observations. Similar results hold for random Gaussian and Rademacher projections (i.e., projections consisting of independent and identically distributed Gaussian or Rademacher random variables) [2, 3]. The results presented in these works can be roughly summarized as follows. Assume that a signal $f \in \mathbb{R}^n$ is "compressible" in some orthonormal basis in the following sense. Let $f^{(m)}$ denote the best $m$-term approximation of $f$ in terms of this basis and suppose that the average squared error obeys

$$\frac{\|f - f^{(m)}\|^2}{n} = \frac{1}{n}\sum_{i=1}^{n}(f_i - f_i^{(m)})^2 \leq C_A m^{-2\alpha} \tag{2.1}$$

for some $\alpha \geq 0$ and some constant $C_A > 0$. The parameter $\alpha$ governs the degree to which $f$ is compressible with respect to the basis. In a noiseless setting, it can be shown that an approximation of such a signal can be recovered from $k$ random projections with an average squared error that is upper bounded by a constant times $(k/\log n)^{-2\alpha}$, nearly as good as the best $k$-term approximation error. For this reason, these procedures are often referred to as "compressive sampling" since the number of samples required is directly related to the sparsity or compressibility of the signal, rather than its temporal/spatial extent or bandwidth.

This paper takes the investigation of compressive sampling a step further by considering the performance of sampling via random projections in noisy conditions. We show that if the projections are contaminated with zero-mean Gaussian noise, then compressible signals can be reconstructed with an expected average squared error that is upper bounded by a constant times $(k/\log n)^{\frac{-2\alpha}{2\alpha+1}}$. For truly sparse signals (with only a small number of non-zero terms) a stronger result is obtained; the expected average squared reconstruction error is upper bounded by a constant times $(k/\log n)^{-1}$. These bounds demonstrate a remarkable capability

[1] The material in Section 2.1 is ©2006 IEEE. Reprinted, with permission, from *IEEE Transactions on Information Theory*, "Signal Reconstruction from Noisy Random Projections," J. Haupt and R. Nowak.

of compressive sampling – accurate reconstructions can be obtained even when the signal dimension $n$ greatly exceeds the number of samples $k$ *and* the samples themselves are contaminated with significant levels of noise.

This effect is highlighted by the following "needle in a haystack" problem. Suppose the signal $f^*$ is a vector of length $n$ with one nonzero entry of amplitude $\sqrt{n}$. If we sample the vector at $k$ random locations (akin to conventional sampling schemes), then the probability of completely missing the non-zero entry is $(1 - 1/n)^k$, which is very close to 1 when $k$ is significantly smaller than $n$. This implies that the expected average squared error may be almost 1, or larger if noise is present. On the other hand, by sampling with randomized projections our results guarantee that the expected average squared error will be no larger than a constant times $(k/\log n)^{-1}$, which can be close to 0 even when $k \ll n$, provided $k > \log n$.

A closely related problem is the reconstruction of signals with sparse Fourier spectra from a relatively small number of non-uniform time samples (e.g., random samples in time) [8–11]. Most of this work concerns noiseless situations, but [9] addresses the problem of reconstruction from noise-corrupted samples. Another area of work related to our results concerns the reconstruction of signals with finite degrees of freedom using a small number of non-traditional samples [12, 13]. A special instance of this setup is the case of signals that are sparse in time (the dual of the spectrally sparse case). Reconstruction from noise-corrupted samples is the focus of [13]. In a sense, the sampling and reconstruction problems addressed in the papers above are special cases of the class of problems considered here, where we allow signals that are sparse in some arbitrary domain. Again, this more universal perspective is precisely the focus of [2, 3], which consider signal reconstruction from noiseless random projections. An interesting line of similar work concerns the related problem of signal reconstruction from random projections corrupted by an unknown but bounded perturbation [14, 15]. In this paper we consider unbounded, Gaussian noise contamination in the sampling process. Finally, while this paper was under review a related investigation was reported in [16] pertaining to the statistical estimation of sparse signals from underdetermined and noisy observations. That work proposes a linear program to obtain an estimator with quantitative bounds for sparse signal reconstruction similar to ours, but is based on a uniform uncertainty principle rather than randomized designs as here.

The paper is organized as follows. In Section 2.1.2 we state the basic problem and main theoretical results of the paper. In Section 2.1.3 we derive bounds on the accuracy of signal reconstructions from noisy random projections. In Section 2.1.4 we specialize the bounds to cases in which the underlying signal is compressible in terms of a certain orthonormal basis. In Section 2.1.5 we propose a simple iterative algorithm for signal reconstruction. In Section 2.1.6 we discuss applications to encoding, A/D conversion, and wireless sensing, and we make concluding remarks in Section 2.1.7. Detailed derivations are relegated to the Appendix.

### 2.1.2 Main Results

Consider a vector $f^* = [f_1^* \ f_2^* \dots f_n^*]^T \in \mathbb{R}^n$ and assume that $\sum_{i=1}^n (f_i^*)^2 \equiv \|f^*\|^2 \leq nB^2$ for a known constant $B > 0$. The assumption simply implies that the average per element energy is bounded by a constant. This is a fairly weak restriction since it permits a very large class of signals, including signals with peak magnitudes as large as $O(\sqrt{n})$. Now suppose that we are able to make $k$ measurements of $f^*$ in the form of noisy, random projections. Specifically, let $\Phi = \{\phi_{i,j}\}$ be an $n \times k$ array of bounded, i.i.d. zero-mean random variables of variance $E[\phi_{i,j}^2] = 1/n$. Samples take the form

$$y_j = \sum_{i=1}^{n} \phi_{i,j} f_i^* + w_j, \quad j = 1, \dots, k \tag{2.2}$$

where $w = \{w_j\}$ are i.i.d. zero-mean random variables, independent of $\{\phi_{i,j}\}$, with variance $\sigma^2$. The goal is to recover an estimate of $f^*$ using $\{y_j\}$ and $\{\phi_{i,j}\}$.

Define the risk of a candidate reconstruction $f$ to be

$$R(f) \ = \ \frac{\|f^* - f\|^2}{n} + \sigma^2 \tag{2.3}$$

where the norm is the Euclidean distance. Next assume that both $\{\phi_{i,j}\}$ and $\{y_j\}$ are available. Then we can compute the empirical risk

$$\widehat{R}(f) = \frac{1}{k} \sum_{j=1}^{k} \left( y_j - \sum_{i=1}^{n} \phi_{i,j} f_i \right)^2 . \tag{2.4}$$

It is easy to verify that $E[\widehat{R}(f)] = R(f)$ using the facts that $\{\phi_{i,j}\}$ and $\{w_j\}$ are independent random variables and $E[\phi_{i,j}^2] = 1/n$. Thus, $\widehat{R}(f)$ is an unbiased estimator of $R(f)$. We will use the empirical risk to obtain an estimator $\widehat{f}$ of $f^*$, and bound the resulting error $E[\|\widehat{f} - f^*\|^2]$. The estimator is based on a complexity-regularized empirical risk minimization, and we use the Craig-Bernstein concentration inequality to control the estimation error of the reconstruction process. That inequality entails the verification of certain moment conditions, which depend on the nature of $\Phi$ and $w$. In this paper we focus on (normalized) Rademacher projections, in which case each $\phi_{i,j}$ is $\pm 1/\sqrt{n}$ with equal probability, and assume that $w$ is a sequence of zero-mean Gaussian noises. Generalizations to other random projections and noise models may be possible following our approach; this would only require one to verify the moment conditions required by the Craig-Bernstein inequality.

Suppose that we have a countable collection $\mathcal{F}$ of candidate reconstruction functions and a non-negative number $c(f)$ assigned to each $f \in \mathcal{F}$ such that $\sum_{f \in \mathcal{F}} 2^{-c(f)} \leq 1$. Furthermore, assume that each $f \in \mathcal{F}$ satisfies $\|f\|^2 \leq nB^2$. Select a reconstruction according to the complexity-regularized empirical risk minimization

$$\widehat{f}_k \ = \ \arg\min_{f \in \mathcal{F}} \left\{ \widehat{R}(f) + \frac{c(f) \log 2}{k\epsilon} \right\} \tag{2.5}$$

where $\epsilon > 0$ is a constant that depends on $B$ and $\sigma$. Then we have the following oracle inequality.

**Theorem 2.1** *Let $\epsilon = 1/(50(B+\sigma)^2)$, then*

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 \min_{f\in\mathcal{F}}\left\{\frac{\|f-f^*\|^2}{n} + \frac{c(f)\log 2 + 4}{k\epsilon}\right\}, \tag{2.6}$$

*where the constant $C_1$ is given by*

$$C_1 = \frac{(27-4e)S^2 + (50-4\sqrt{2})S + 26}{(23-4e)S^2 + (50-4\sqrt{2})S + 24} \tag{2.7}$$

*with $S = B/\sigma$, the signal-to-noise ratio.*

An important point regarding the constants above is that they depend only on $\sigma^2$ and $B^2$, the noise power and the average signal power, respectively. More specifically, note that $1.08 \leq C_1 \leq 1.33$ and that $\epsilon$ is inversely proportional to $\sigma^2$, so the bound degrades gracefully as the noise level $\sigma$ increases, in the manner one would expect. The dependence of $\epsilon$ on the average signal power is also not surprising, since there is a small chance of "missing" significant components of the signal in our random samples and this source of error scales with $B^2$.

If $f^*$ is compressible with respect to some known orthonormal basis, then we can obtain explicit bounds on the reconstruction error in terms of the number of random projections $k$ and the degree to which $f^*$ is compressible. Let $f^{(m)}$ denote the best $m$-term approximation of $f^*$ in the basis. That is, if $f^*$ has a representation $f^* = \sum_{i=1}^n \theta_i\psi_i$ in the basis $\{\psi_i\}$, then $f^{(m)} = \sum_{i=1}^m \theta_{(i)}\psi_{(i)}$, where coefficients and basis functions are ordered such that $|\theta_{(1)}| \geq |\theta_{(2)}| \geq \cdots \geq |\theta_{(n)}|$. Assume that the average squared error $\|f^* - f^{(m)}\|^2/n \equiv \frac{1}{n}\sum_{i=1}^n (f_i^* - f_i^{(m)})^2$ satisfies

$$\frac{\|f^* - f^{(m)}\|^2}{n} \leq C_A m^{-2\alpha} \tag{2.8}$$

for some $\alpha \geq 0$ and some constant $C_A > 0$. Power-law decays like this arise quite commonly in applications. For example, smooth and piecewise smooth signals as well as signals of bounded variation exhibit this sort of behavior [2, 3]. It is also unnecessary to restrict our attention to orthonormal basis expansions. Much more general approximation strategies can be accommodated [2], but to keep the presentation as simple as possible we will not delve further into such extensions.

Take $\mathcal{F}_c(B,\alpha,C_A) = \{f : \|f\|^2 < nB^2,\ \|f - f^{(m)}\|^2 \leq nC_A m^{-2\alpha}\}$ to be the class of functions to which $f^*$ belongs, and let $\mathcal{F}$ to be a suitably quantized collection of functions represented in terms of the basis $\{\psi_i\}$ (the construction of $\mathcal{F}$ is discussed in Section 2.1.4). We have the following error bound.

**Theorem 2.2** *If*

$$c(f) = 2\log(n) \times (\#\ \textit{non-zero coefficients of } f \textit{ in the basis } \{\psi_i\}) \tag{2.9}$$

*then there exists a constant* $C_2 = C_2(B, \sigma, C_A) > 0$ *such that*

$$\sup_{f^* \in \mathcal{F}_c(B,\alpha,C_A)} E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 C_2 \left(\frac{k}{\log n}\right)^{-2\alpha/(2\alpha+1)}, \tag{2.10}$$

*where* $C_1$ *is as given in Theorem 2.1.*

Note that the exponent $-2\alpha/(2\alpha+1)$ is the usual exponent governing the rate of convergence in nonparametric function estimation.

A stronger result is obtained if the signal is sparse (i.e. belonging to the class $\mathcal{F}_s(B,m) = \{f : \|f\|^2 < nB^2,\ \|f\|_0 \leq m\}$) as stated in the following Corollary.

**Corollary 2.3** *Suppose that* $f^* \in \mathcal{F}_s(B,m)$*. Then there exists a constant* $C_2' = C_2'(B,\sigma) > 0$ *such that*

$$\sup_{f^* \in \mathcal{F}_s(B,m)} E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 C_2' \left(\frac{k}{m \log n}\right)^{-1}, \tag{2.11}$$

*where* $C_1$ *is as given in Theorem 2.1.*

Similar results hold if the signal is additionally contaminated with noise prior to the random projection process, as described in the following Corollary.

**Corollary 2.4** *Suppose observations take the form*

$$y_j = \sum_{i=1}^{n} \phi_{i,j} \left(f_i^* + \eta_{i,j}\right) + w_j, \quad j = 1, \ldots, k \tag{2.12}$$

*where* $\{\eta_{i,j}\}$ *are i.i.d. zero-mean Gaussian random variables with variance* $\sigma_s^2$ *that are independent of* $\{\phi_{i,j}\}$ *and* $\{w_j\}$*. Then Theorems 2.1 and 2.2 and Corollary 2.3 hold with slightly different constants* $C_1$*,* $C_2$*,* $C_2'$*, and* $\epsilon$*.*

It is important to point out that all the results above hold for arbitrary signal lengths $n$, and the constants do not depend on $n$. The fact that the rate depends only logarithmically on $n$ is significant and illustrates the scalability of this approach. One can interpret these bounds as good indicators of the exceptional performance of random projection sampling in large-$n$ regimes. The dependence on $k$ is shown to be polynomial. In analogy with nonparametric estimation theory (e.g., estimating smooth functions from random point samples), the polynomial rate in $k$ is precisely what one expects in general, and thus we believe the upper bounds are tight (up to constant and logarithmic factors).

To drive this point home, let us again consider the "needle in a haystack" problem, this time in a bit more detail. Suppose the signal $f^*$ is a vector of length $n$ with one nonzero entry of amplitude $\sqrt{n}$ such that $\|f^*\|^2/n = 1$. First, consider random spatial point sampling where observations are noise-free (i.e., each

sample is of the form $y_j = f^*(t_j)$, where $t_j$ is selected uniformly at random from the set $\{1, \ldots, n\}$). The squared reconstruction error is 0 if the spike is located and 1 otherwise, and the probability of not finding the spike in $k$ trials is $(1-1/n)^k$, giving an average squared error of $(1-1/n)^k \cdot 1 + (k/n) \cdot 0 = (1-1/n)^k$. If $n$ is large, we can approximate this by $(1-1/n)^k \approx e^{-k/n}$, which is very close to 1 when $k$ is significantly smaller than $n$. On the other hand, randomized Rademacher projections (corrupted with noise) yield an average squared reconstruction error bound of $C_2' (k/\log n)^{-1}$, as given above in Corollary 2.3. This bound may be close to 0 even when $k \ll n$, provided $k > \log n$. This shows that even given the advantage of being noiseless, the reconstruction error from spatial point sampling may be far greater than that resulting from random projections.

### 2.1.3 Oracle Inequality

In this section we prove Theorem 2.1. For ease of notation, we adopt the shorthand notation $\phi_j = [\phi_{1,j}\ \phi_{2,j} \ldots \phi_{n,j}]^T$ for the vector corresponding to the $j^{th}$ projection. The empirical risk of a vector $f$ can now be written as

$$\widehat{R}(f) = \frac{1}{k} \sum_{j=1}^{k} (y_j - \phi_j^T f)^2. \tag{2.13}$$

We will bound $r(f, f^*) \equiv R(f) - R(f^*)$, the "excess risk" between a candidate reconstruction $f$ and the actual function $f^*$, using the complexity-regularization method introduced in [17]. Note that $r(f, f^*) = \|f - f^*\|^2/n$.

Define the empirical excess risk $\widehat{r}(f, f^*) \equiv \widehat{R}(f) - \widehat{R}(f^*)$. Then

$$\begin{aligned} \widehat{r}(f, f^*) &= -\frac{1}{k} \sum_{j=1}^{k} \left[ (y_j - \phi_j^T f^*)^2 - (y_j - \phi_j^T f)^2 \right] \\ &= -\frac{1}{k} \sum_{j=1}^{k} U_j \end{aligned} \tag{2.14}$$

where $U_j = \left[ (y_j - \phi_j^T f^*)^2 - (y_j - \phi_j^T f)^2 \right]$ are i.i.d. for $j = 1 \ldots k$. Notice that $r(f, f^*) - \widehat{r}(f, f^*) = \frac{1}{k} \sum_{j=1}^{k} (U_j - E[U_j])$. We will make use of the Craig-Bernstein inequality [18], which states that the probability of the event

$$\frac{1}{k} \sum_{j=1}^{k} (U_j - E[U_j]) \geq \frac{t}{k\epsilon} + \frac{\epsilon k\ var\left( \frac{1}{k} \sum_{j=1}^{k} U_j \right)}{2(1-\zeta)} \tag{2.15}$$

is less than or equal to $e^{-t}$ for $0 < \epsilon h \leq \zeta < 1$ and $t > 0$, provided the variables $U_j$ satisfy the moment condition

$$E\left[ |U_j - E[U_j]|^k \right] \ \leq \ \frac{k!\, var(U_j)\, h^{k-2}}{2} \tag{2.16}$$

for some $h > 0$ and all integers $k \geq 2$. If we consider vectors $f^*$ and estimates $f$ where $\|f^*\|^2 \leq nB^2$ and $\|f\|^2 \leq nB^2$, Rademacher projections, and Gaussian noises with variance $\sigma^2$, then the moment condition is

satisfied with $h = 16B^2e + 8\sqrt{2}B\sigma$, as shown in the Appendix. Alternative forms of random projections and noises can also be handled using the approach outlined next, provided the moment conditions are satisfied.

To use the Craig-Bernstein inequality we also need a bound on the variance of $U_j$ itself. Defining $g = f - f^*$, we have

$$var(U_j) = E[(\phi^T g)^4] - \left(\frac{\|g\|^2}{n}\right)^2 + \frac{4\sigma^2\|g\|^2}{n}. \tag{2.17}$$

As shown in the Appendix, for integers $k \geq 1$

$$\left(\frac{\|g\|^2}{n}\right)^k \leq E[(\phi^T g)^{2k}] \leq (2k)!! \left(\frac{\|g\|^2}{n}\right)^k, \tag{2.18}$$

where $(2k)!! \equiv (1)(3)\dots(2k-1)$. Thus we can bound the variance of $U_j$ by

$$var(U_j) \leq \left(2\frac{\|g\|^2}{n} + 4\sigma^2\right)\frac{\|g\|^2}{n}. \tag{2.19}$$

Since $g$ satisfies $\|g\|^2 \leq 4nB^2$ and $r(f, f^*) = \|f - f^*\|^2/n = \|g\|^2/n$, the bound becomes

$$var(U_j) \leq \left(8B^2 + 4\sigma^2\right) r(f, f^*). \tag{2.20}$$

So, we can replace the term in the Craig-Bernstein inequality that depends on the variance by

$$\begin{aligned} k\ var\left(\frac{1}{k}\sum_{j=1}^{k} U_j\right) &= \frac{1}{k}\sum_{j=1}^{k} var(U_j) \\ &\leq \left(8B^2 + 4\sigma^2\right) r(f, f^*). \end{aligned} \tag{2.21}$$

For a given function $f$, we have that the probability of the event

$$r(f, f^*) - \widehat{r}(f, f^*) > \frac{t}{k\epsilon} + \frac{\left(8B^2 + 4\sigma^2\right)\epsilon\ r(f, f^*)}{2(1-\zeta)} \tag{2.22}$$

is less than or equal to $e^{-t}$, or by letting $\delta = e^{-t}$, the probability of

$$r(f, f^*) - \widehat{r}(f, f^*) > \frac{\log\left(\frac{1}{\delta}\right)}{k\epsilon} + \frac{\left(8B^2 + 4\sigma^2\right)\epsilon\ r(f, f^*)}{2(1-\zeta)} \tag{2.23}$$

is less than or equal to $\delta$. Now assign to each $f \in \mathcal{F}$ a non-negative penalty term $c(f)$ such that the penalties satisfy the Kraft Inequality [19]

$$\sum_{f \in \mathcal{F}} 2^{-c(f)} \leq 1 \tag{2.24}$$

and let $\delta(f) = 2^{-c(f)}\delta$. Then by applying the union bound we have for all $f \in \mathcal{F}$ and for all $\delta > 0$

$$r(f, f^*) - \widehat{r}(f, f^*) \leq \frac{c(f)\log 2 + \log\left(\frac{1}{\delta}\right)}{k\epsilon} + \frac{\left(8B^2 + 4\sigma^2\right)\epsilon\ r(f, f^*)}{2(1-\zeta)} \tag{2.25}$$

with probability at least $1 - \delta$. Set $\zeta = \epsilon h$, and define

$$a \equiv \frac{\left(8B^2 + 4\sigma^2\right)\epsilon}{2(1-\zeta)}. \tag{2.26}$$

Choose

$$\epsilon < \frac{1}{(4+16e)B^2 + 8\sqrt{2}B\sigma + 2\sigma^2}, \tag{2.27}$$

and notice that $a < 1$ and $\zeta < 1$ by choice of $\epsilon$. Then

$$(1-a)r(f,f^*) \leq \widehat{r}(f,f^*) + \frac{c(f)\log 2 + \log\left(\frac{1}{\delta}\right)}{k\epsilon} \tag{2.28}$$

holds with probability at least $1-\delta$ for all $f \in \mathcal{F}$ and any $\delta > 0$.

For the given training samples, we can minimize the upper bound by choosing

$$\widehat{f}_k = \arg\min_{f\in\mathcal{F}} \left\{ \widehat{r}(f,f^*) + \frac{c(f)\log 2}{k\epsilon} \right\} \tag{2.29}$$

which is equivalent to

$$\widehat{f}_k = \arg\min_{f\in\mathcal{F}} \left\{ \widehat{R}(f) + \frac{c(f)\log 2}{k\epsilon} \right\} \tag{2.30}$$

since we can ignore $\widehat{R}(f^*)$ when performing the optimization. If we define

$$f_k^* \equiv \arg\min_{f\in\mathcal{F}} \left\{ R(f) + \frac{c(f)\log 2}{k\epsilon} \right\} \tag{2.31}$$

then with probability at least $1-\delta$

$$\begin{aligned} &(1-a)r(\widehat{f}_k,f^*) \\ &\leq \quad \widehat{r}(\widehat{f}_k,f^*) + \frac{c(\widehat{f}_k)\log 2 + \log\left(\frac{1}{\delta}\right)}{k\epsilon} \\ &\leq \quad \widehat{r}(f_k^*,f^*) + \frac{c(f_k^*)\log 2 + \log\left(\frac{1}{\delta}\right)}{k\epsilon} \end{aligned} \tag{2.32}$$

since $\widehat{f}_k$ minimizes the complexity-regularized empirical risk criterion. Using the Craig-Bernstein inequality again to bound $\widehat{r}(f_k^*,f^*) - r(f_k^*,f^*)$ (with the same variance bound as before) we get that with probability at least $1-\delta$

$$\widehat{r}(f_k^*,f^*) - r(f_k^*,f^*) \leq a\, r(f_k^*,f^*) + \frac{\log\left(\frac{1}{\delta}\right)}{k\epsilon}. \tag{2.33}$$

We want both (2.32) and (2.33) to hold simultaneously, so we use the union bound to obtain

$$r(\widehat{f}_k,f^*) \leq \left(\frac{1+a}{1-a}\right) r(f_k^*,f^*) + \frac{1}{1-a}\left(\frac{c(f_k^*)\log 2 + 2\log\left(\frac{1}{\delta}\right)}{k\epsilon}\right) \tag{2.34}$$

holding with probability at least $1-2\delta$.

We will convert this probability deviation bound into a bound on the expected value using the fact that, for positive random variables $X$, $E[X] = \int_0^\infty P(X > t)dt$. Let $\delta = e^{-k\epsilon t(1-a)/2}$ to obtain

$$P\left( r(\widehat{f}_k,f^*) - \left(\frac{1+a}{1-a}\right) r(f_k^*,f^*) - \frac{c(f_k^*)\log 2}{k\epsilon(1-a)} \geq t \right) \leq 2e^{-k\epsilon t(1-a)/2}. \tag{2.35}$$

Integrating this relation gives

$$E\left[r(\widehat{f}_k, f^*)\right] \leq \left(\frac{1+a}{1-a}\right) r(f_k^*, f^*) + \frac{c(f_k^*)\log 2 + 4}{k\epsilon(1-a)}. \tag{2.36}$$

Now, since $a$ is positive,

$$\begin{aligned} E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] = E\left[r(\widehat{f}_k, f^*)\right] &\leq \left(\frac{1+a}{1-a}\right) r(f_k^*, f^*) + \frac{c(f_k^*)\log 2 + 4}{k\epsilon(1-a)} \\ &\leq \left(\frac{1+a}{1-a}\right) r(f_k^*, f^*) + (1+a)\frac{c(f_k^*)\log 2 + 4}{k\epsilon(1-a)} \\ &= \left(\frac{1+a}{1-a}\right)\left\{R(f_k^*) - R(f^*) + \frac{c(f_k^*)\log 2 + 4}{k\epsilon}\right\} \\ &\leq \left(\frac{1+a}{1-a}\right)\min_{f\in\mathcal{F}}\left\{R(f) - R(f^*) + \frac{c(f)\log 2 + 4}{k\epsilon}\right\} \\ &= C_1 \min_{f\in\mathcal{F}}\left\{\frac{\|f - f^*\|^2}{n} + \frac{c(f)\log 2 + 4}{k\epsilon}\right\}, \end{aligned} \tag{2.37}$$

where $C_1 = (1+a)/(1-a)$.

Typical values of $C_1$ can be determined by approximating the constant

$$\epsilon < \frac{1}{(4+16e)B^2 + 8\sqrt{2}B\sigma + 2\sigma^2}. \tag{2.38}$$

Upper bounding the denominator guarantees that the condition is satisfied, so let $\epsilon = 1/\left(50(B+\sigma)^2\right)$. Now

$$\begin{aligned} a &= \frac{\left(8B^2 + 4\sigma^2\right)\epsilon}{2(1-\zeta)} \\ &= \frac{2B^2 + \sigma^2}{(25-4e)B^2 + (50-4\sqrt{2})B\sigma + 25\sigma^2}. \end{aligned} \tag{2.39}$$

If we denote the signal to noise ratio by $S^2 = B^2/\sigma^2$ then

$$a = \frac{2S^2 + 1}{(25-4e)S^2 + (50-4\sqrt{2})S + 25} \tag{2.40}$$

for which the extremes are $a_{min} = 1/25$ and $a_{max} = 2/(25-4e)$, giving constants $C_1$ in the range of $[13/12, (27-4e)/(23-4e)] \approx [1.08, 1.33]$.

### 2.1.4 Error Bounds for Compressible Signals

In this section we prove Theorem 2.2 and Corollary 2.3. Suppose that $f^*$ is compressible in a certain orthonormal basis $\{\psi_i\}_{i=1}^n$. Specifically, let $f^{(m)}$ denote the best $m$-term approximation of $f^*$ in terms of $\{\psi_i\}$, and assume that the error of the approximation obeys

$$\frac{\|f^* - f^{(m)}\|^2}{n} \leq C_A\, m^{-2\alpha} \tag{2.41}$$

for some $\alpha \geq 0$ and a constant $C_A > 0$. Let $\mathcal{F}_c(B, \alpha, C_A) = \{f : \|f\|^2 < nB^2,\ \|f - f^{(m)}\|^2 \leq nC_A m^{-2\alpha}\}$ so $f^* \in \mathcal{F}_c(B, \alpha, C_A)$.

Let us use the basis $\{\psi_i\}$ for the reconstruction process. Any vector $f$ can be expressed in terms of the basis $\{\psi_i\}$ as $f = \sum_{i=1}^n \theta_i \psi_i$, where $\theta = \{\theta_i\}$ are the coefficients of $f$ in this basis. Let $T$ denote the transform that maps coefficients to functions, so that $f = T\theta$. Define $\Theta = \{\theta : \|T\theta\|^2 \leq nB^2, \theta_i$ uniformly quantized to $n^p$ levels$\}$ to be the set of candidate solutions in the basis $\{\psi_i\}$, so that $\mathcal{F} = \{f : f = T\theta,\ \theta \in \Theta\}$. The penalty term $c(f)$ written in terms of the basis $\{\psi_i\}$ is $c(f) = c(\theta) = (1+p)\log(n)\sum_{i=1}^n I_{\theta_i \neq 0} = (1+p)\log(n)\|\theta\|_0$. It is easily verified that $\sum_{f\in\mathcal{F}} 2^{-c(f)} \leq 1$ by noting that each $\theta \in \Theta$ can be uniquely encoded via a prefix code consisting of $(1+p)\log n$ bits per non-zero coefficient ($\log n$ bits for the location and $p \log n$ bits for the quantized value) in which case the codelengths $c(f)$ must satisfy the Kraft inequality [19].

The oracle inequality

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 \min_{f\in\mathcal{F}}\left\{\frac{\|f - f^*\|^2}{n} + \frac{c(f)\log 2 + 4}{k\epsilon}\right\} \tag{2.42}$$

can also be written as

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 \min_{\theta\in\Theta}\left\{\frac{\|\theta - \theta^*\|^2}{n} + \frac{c(\theta)\log 2 + 4}{k\epsilon}\right\} \tag{2.43}$$

where $f^* = T\theta^*$. For each integer $m \geq 1$, let $\theta^{(m)}$ denote the coefficients corresponding to the best $m$-term approximation of $f^*$ and let $\theta_q^{(m)}$ denote the nearest element in $\Theta$. The maximum possible dynamic range for the coefficient magnitudes, $\pm\sqrt{n}B$, is quantized to $n^p$ levels, giving $\|\theta_q^{(m)} - \theta^{(m)}\|^2 \leq 4B^2/n^{2p-2} = C_Q/n^{2p-2}$. Now insert $\theta_q^{(m)}$ in place of $\theta$ in the oracle bound. The first term can be expanded as

$$\begin{aligned}
\|\theta_q^{(m)} - \theta^*\|^2 &= \|\theta_q^{(m)} - \theta^{(m)} + \theta^{(m)} - \theta^*\|^2 \\
&\leq \|\theta_q^{(m)} - \theta^{(m)}\|^2 + 2\|\theta_q^{(m)} - \theta^{(m)}\| \cdot \|\theta^{(m)} - \theta^*\| + \|\theta^{(m)} - \theta^*\|^2 \\
&\leq \frac{C_Q}{n^{2p-2}} + 2m^{-\alpha}\sqrt{\frac{nC_AC_Q}{n^{2p-2}}} + C_A n m^{-2\alpha}.
\end{aligned} \tag{2.44}$$

Now notice that $c(\theta_q^{(m)}) = (1+p)m\log n$, so

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 \min_m\left\{\frac{C_Q}{n^{2p-1}} + \frac{2m^{-\alpha}\sqrt{C_AC_Q}}{n^{p-1/2}} + C_A m^{-2\alpha} + \frac{(1+p)m\log n\log 2}{k\epsilon} + \frac{4}{k\epsilon}\right\}. \tag{2.45}$$

The quantization error terms decay exponentially in $p$, so they can be made arbitrarily small while incurring only a modest (linear) increase in the complexity term. Balancing the third and fourth terms gives

$$m = \left(\frac{(1+p)\log 2}{\epsilon C_A}\right)^{\frac{-1}{2\alpha+1}}\left(\frac{k}{\log n}\right)^{\frac{1}{2\alpha+1}} \tag{2.46}$$

so

$$C_A m^{-2\alpha} = C_A \left( \frac{(1+p)\log 2}{\epsilon C_A} \right)^{\frac{2\alpha}{2\alpha+1}} \left( \frac{k}{\log n} \right)^{\frac{-2\alpha}{2\alpha+1}}, \tag{2.47}$$

and since

$$\frac{1}{k} < \left( \frac{\log n}{k} \right)^{\frac{2\alpha}{2\alpha+1}} \tag{2.48}$$

when $k > 1$ and $n > e$, then

$$E\left[ \frac{\|\widehat{f}_k - f^*\|^2}{n} \right] \leq C_1 C_2 \left( \frac{k}{\log n} \right)^{\frac{-2\alpha}{2\alpha+1}} \tag{2.49}$$

holds for every $f^* \in \mathcal{F}_c(B, \alpha, C_A)$ as claimed in the Theorem, where

$$C_2 = \left\{ 2C_A \left( \frac{(1+p)\log 2}{\epsilon C_A} \right)^{\frac{2\alpha}{2\alpha+1}} + \frac{4}{\epsilon} \right\}. \tag{2.50}$$

Suppose now that $f^* \in \mathcal{F}_s(B, m)$ where $\mathcal{F}_s(B, m) = \{f : \|f\|^2 < nB^2,\ \|f\|_{l_0} \leq m\}$. In this case,

$$\|\theta_q^{(m)} - \theta^*\|^2 \leq \frac{C_Q}{n^{2p-2}} \tag{2.51}$$

since $\|\theta^{(m)} - \theta^*\| = 0$. Now the penalty term dominates in the oracle bound and

$$E\left[ \frac{\|\widehat{f}_k - f^*\|^2}{n} \right] \leq C_1 C_2' \left( \frac{k}{m \log n} \right)^{-1}, \tag{2.52}$$

holds for every $f^* \in \mathcal{F}_s(B, m)$ where

$$C_2' = \left\{ \frac{(1+p)\log 2 + 4}{\epsilon} \right\}. \tag{2.53}$$

### 2.1.5 Optimization Scheme

Although our optimization is non-convex, it does permit a simple, iterative optimization strategy that produces a sequence of reconstructions for which the corresponding sequence of complexity-regularized empirical risk values is non-increasing. This algorithm, which is described below, has demonstrated itself to be quite effective in similar denoising and reconstruction problems [20–22]. A possible alternative strategy might entail "convexifying" the problem by replacing the $l_0$ penalty with an $l_1$ penalty. Recent results show that often the solution to this convex problem coincides with or approximates the solution to the original non-convex problem [23].

Let us assume that we wish to reconstruct our signal in terms of the basis $\{\psi_i\}$. Using the definitions introduced in the previous section, the reconstruction

$$\widehat{f}_k = \arg\min_{f \in \mathcal{F}} \left\{ \widehat{R}(f) + \frac{c(f)\log 2}{k\epsilon} \right\} \tag{2.54}$$

is equivalent to $\widehat{f}_k = T\widehat{\theta}_k$ where

$$\widehat{\theta}_k = \arg\min_{\theta\in\Theta}\left\{\widehat{R}(T\theta) + \frac{c(\theta)\log 2}{k\epsilon}\right\} \tag{2.55}$$

Thus, the optimization problem can then be written as

$$\widehat{\theta}_k = \arg\min_{\theta\in\Theta}\left\{\|y - PT\theta\|^2 + \frac{2\log(2)\log(n)}{\epsilon}\|\theta\|_0\right\} \tag{2.56}$$

where $P = \Phi^T$, the transpose of the $n \times k$ projection matrix $\Phi$, $y$ is a column vector of the $k$ observations, and $\|\theta\|_0 = \sum_{i=1}^{n} I_{\{\theta_i \neq 0\}}$.

To solve this, we use an iterative bound-optimization procedure, as proposed in [20–22]. This procedure entails a two-step iterative process that begins with an initialization $\theta^{(0)}$ and computes:

$$1. \qquad \varphi^{(t)} \;=\; \theta^{(t)} + \frac{1}{\lambda}(PT)^T(y - PT\theta^{(t)}) \tag{2.57}$$

$$2. \qquad \widehat{\theta}_i^{(t+1)} \;=\; \begin{cases} \varphi_i^{(t)} & \text{if } |\varphi_i^{(t)}| \geq \sqrt{\frac{2\log(2)\log(n)}{\lambda\epsilon}} \\ 0 & \text{otherwise} \end{cases} \tag{2.58}$$

where $\lambda$ is the largest eigenvalue of $P'P$. This procedure is desirable since the second step, in which the complexity term plays its role, involves a simple coordinate-wise thresholding operation. It is easy to verify that the iterations produce a monotonically non-increasing sequence of complexity-regularized empirical risk values [22]. Thus, this procedure provides a simple iteration that tends to minimize the original objective function, and appears to give good results in practice [20]. The iterations can be terminated when the entries uniformly satisfy $|\widehat{\theta}_i^{(t+1)} - \widehat{\theta}_i^{(t)}| \leq \delta$, for a small positive tolerance $\delta$.

The computational complexity of the above procedure is quite appealing. Each iteration requires only $O(nk)$ operations, assuming that the transform $T$ can be computed in $O(n)$ operations. For example, the discrete wavelet or Fourier transforms can be computed in $O(n)$ and $O(n\log n)$ operations, respectively. Multiplication by $P$ is the most intensive operation, requiring $O(nk)$ operations. The thresholding step is carried out independently in each coordinate, and this step requires $O(n)$ operations as well. Of course, the number of iterations required is problem-dependent and difficult to predict, but in our experience in this application and others [20, 22] algorithms of this sort tend to converge in a reasonably small number of iterations, even in very high dimensional cases.

One point worthy of mention relates to the factor $1/\epsilon = 50(B+\sigma)^2$ in the penalty. As is often the case with conservative bounds of this type, the theoretical penalty is larger than what is needed in practice to achieve good results. Also, a potential hurdle to calibrating the algorithm is that it depends on knowledge of $B$ and $\sigma$, neither of which may be known a priori. Strictly speaking, these values do not need to be known *independently* but rather we need only estimate $(B+\sigma)^2$. To that end, notice that each observation is a

random variable with variance equal to $\|f\|^2/n + \sigma^2$. Let $B = \sqrt{\|f\|^2/n}$, which is the minimum $B$ satisfying the stated bound $\|f\|^2 \leq nB^2$. Then the variance of each observation is $B^2 + \sigma^2$. Further, it is easy to verify that $2(B^2 + \sigma^2) \geq (B + \sigma)^2$. So, a scheme could be developed whereby the sample variance is used as a surrogate for the unknown quantities in the form in which they appear in the parameter $\epsilon$. This would entail using another concentration inequality to control the error between the sample variance and its mean value, and propagating this additional error through the derivation of the oracle inequality. While this is relatively straightforward, we omit a complete derivation here.

To illustrate the performance of the algorithm above, in Figure 2.1 we consider three standard test signals, each of length $n = 4096$. Rademacher projection samples (contaminated with additive white Gaussian noise) are taken for the Blocks, Bumps, and Doppler test signals. The algorithm described above is employed for reconstruction, with one slight modification. Since the theoretical penalty can be a bit too conservative in practice, the threshold used in this example is about 20% of the theoretical value (i.e., a threshold of $\sqrt{2\log(2)\log(n)/(\lambda\epsilon)}/4.6$ was used). The SNR, defined as $SNR = 10\log_{10}(B^2/\sigma^2)$ where $B^2 = \|f\|^2/n$, is 21dB for each test signal. To convey a sense of the noise level, column (a) of Figure 2.1 shows the original signals contaminated with the same level of noise (i.e., the signal resulting from conventional point samples contaminated with noise of the same power). Column (b) shows the reconstructions obtained from 600 projections; reconstructions from 1200 projections are shown in column (c). The Blocks signal (top row) was reconstructed using the Haar wavelet basis (Daubechies-2), well-suited to the piecewise constant nature of the signal. The Bumps and Doppler signals (middle and bottom row, respectively) were reconstructed using the Daubechies-6 wavelet basis. Of course, the selection of the "best" reconstruction basis is a separate matter that is beyond the scope of this paper.

### 2.1.6 Applications

One immediate application of the results and methods above is to signal coding and A/D conversion. In the noiseless setting, several authors have suggested the use of random projection sampling for such purposes [1–3]. Our results indicate how such schemes might perform in the presence of noise. Suppose that we have an array of $n$ sensors, each making a noisy measurement. The noise could be due to the sensors themselves or environmental factors. The goal of encoding and A/D conversion is to represent the $n$ sensor readings in a compressed form, suitable for digital storage or transmission. Our results suggest that $k$ random Rademacher projections of the $n$ sensor readings can be used for this purpose, and the error bounds suggest guidelines for how many projections might be required for a certain level of precision.

Our theory and method can also be applied to wireless sensing as follows. Consider the problem of sensing a distributed field (e.g., temperature, light, chemical) using a collection of $n$ wireless sensors distributed uniformly over a region of interest. Such systems are often referred to as *sensor networks*. The goal is to

obtain an accurate, high-resolution reconstruction of the field at a remote destination. One approach to this problem is to require each sensor to digitally transmit its measurement to the destination, where field reconstruction is then performed. Alternatively, the sensors might collaboratively process their measurements to reconstruct the field themselves and then transmit the result to the destination (i.e., the nodes collaborate to compress their data prior to transmission). Both approaches pose significant demands on communication resources and infrastructure, and it has recently been suggested that non-collaborative analog communication schemes may offer a more resource-efficient alternative [24–26].

Assume that the sensor data is to be transmitted to the destination over an additive white Gaussian noise channel. Suppose the destination broadcasts (perhaps digitally) a random seed to the sensors. Each node modifies this seed in a unique way known to only itself and the destination (e.g., this seed could be multiplied by the node's address or geographic position) and uses the seed to generate a pseudorandom Rademacher sequence. The sequences can also be reconstructed at the destination. The nodes then transmit the random projections to the destination *phase-coherently* (i.e., beamforming). This is accomplished by requiring each node to simply multiply its reading by an element of its random sequence in each projection/communication step and transmit the result to the destination via amplitude modulation. If the transmissions from all $n$ sensors can be synchronized so that they all arrive in phase at the destination, then the averaging inherent in the multiple access channel computes the desired inner product. After receiving $k$ projections, the destination can employ the reconstruction algorithm above using a basis of choice (e.g., wavelet). The communications procedure is completely non-adaptive and potentially very simple to implement. The collective functioning of the wireless sensors in this process is more akin to an ensemble of phase-coherent emitters than it is to conventional networking operations. Therefore, we prefer the term *sensor ensemble* instead of sensor network in this context.

A remarkable aspect of the sensor ensemble approach is that the power required to achieve a target distortion level can be very minimal. Let $\sigma_s^2$ and $\sigma_c^2$ denote the noise variance due to sensing and communication, respectively. Thus, each projection received at the destination is corrupted by a noise of total power $\sigma_s^2 + \sigma_c^2$. The sensing noise variance is assumed to be a constant and the additional variance due to the communication channel is assumed to scale like the inverse of the total received power

$$\sigma_c^2 \propto \frac{1}{nP} \tag{2.59}$$

where $P$ is the transmit power per sensor. In order to achieve the distortion decay rates given by our upper bounds, it is sufficient that the variance due to the communication channel behaves like a constant. Therefore, we require only that $P \propto n^{-1}$. This results in a rather surprising conclusion. Ideal reconstruction is possible at the destination with per sensor transmit power $P$ tending to zero as the density of sensors increases. If conventional spatial point samples were taken instead (e.g., if a single sensor is selected at

random in each step and transmits its measurement to the destination), then the power required per sensor would be a constant, since only one sensor would be involved in such a transmission. Thus, it appears that random projection sampling may be more desirable in wireless sensing applications. Thorough treatments of compressive sampling via random projections in sensor networking applications can be found in our recent work [27, 28].

### 2.1.7 Conclusions and Future Work

We have shown that compressible signals can be accurately recovered from random projections contaminated with noise. The squared error bounds for compressible signals are $O((k/\log n)^{\frac{-2\alpha}{2\alpha+1}})$, which is within a logarithmic factor of the usual nonparametric estimation rate, and $O((k/\log n)^{-1})$ for sparse signals. We also proposed a practical iterative algorithm for signal reconstruction. One of the most promising potential applications of our theory and method is to wireless sensing, wherein one realizes a large transmission power gain by random projection sampling as opposed to conventional spatial point sampling.

The role of the noise variance in the rates we presented is worthy of further attention. As the noise variance tends to zero, one intuitively expects to attain the fast rates that are known to be achievable in the noiseless setting. Our theory is based in the noisy regime and it does not directly imply the previously established bounds in the noiseless setting. Simply put, our analysis assumes a noise variance strictly greater than zero.

Let us comment briefly on the tightness of the upper bounds given by our theory. In analogy with classical nonparametric estimation theory (e.g., estimating smooth functions from random point samples), the polynomial rate in $k$ is precisely what one expects in general, and thus we believe the upper bounds are tight (up to constant and logarithmic factors). Moreover, in the special case of sparse signals with $m$ non-zero terms, we obtain an error bound of $m/k$ (ignoring constant and logarithmic factors). Standard parametric statistical analysis suggests that one should not expect a rate of better than $m/k$ (degrees-of-freedom/ sample-size) in such cases, which again supports our intuition regarding the tightness of the bounds (in terms of the convergence rate). However, to our knowledge explicit minimax lower bounds have not been established in the context of this problem, and the determination of such bounds is one of our future research directions.

Although we considered only the case of Gaussian noise in the observation model (2.2), the same results could be achieved for any zero-mean, symmetrically distributed noise that is independent of the projection vector elements and satisfies

$$E[w_j^{2k}] \leq (2k)!!var(w_j)h^{k-2} \tag{2.60}$$

for some constant $h$ not depending on $k$, a result that follows immediately using the lemmas presented in the Appendix. Another extension would be the consideration of other random projections instead of the

Rademacher projections considered here. Most of our basic approach would go through in such cases; one would only need to verify the moment conditions of the Craig-Bernstein inequality for particular cases.

### 2.1.8 Appendix

#### 2.1.8.1 The Craig-Bernstein Moment Condition

The central requirement of the Craig-Bernstein inequality is the satisfaction of the moment condition

$$E\left[|X - E[X]|^p\right] \leq \frac{p!\, var(X) h^{p-2}}{2} \tag{2.61}$$

for integers $p \geq 2$ with some positive constant $h$ that does not depend on $p$, a condition that is trivially verified for $p = 2$ with any finite value of $h$. For larger $p$ this condition can be very difficult to verify for several reasons, not the least of which is the absolute value present in the moments. Complexity regularization methods based on the Craig-Bernstein inequality often assume that the observations are bounded [17], a condition that makes verification of the moment condition relatively simple, but which does not hold in our application due to the Gaussian noise model we work with. Moreover, even if we adopted a bounded noise model, the boundedness of the observations alone is not sufficient to obtain the rates of convergence we establish. Indeed, even in the absence of noise, the observations in our set-up may only be pointwise bounded by the inequality $|y_i| \leq \sqrt{n}B$, yielding a constant $h$ that would grow proportionally to $n$. With that motivation, we develop a framework under which the moments and a bounding constant $h$ (that does not depend on $n$) can be determined more directly.

First, observe that the moment condition is usually easier to verify for the even powers because the absolute value need not be dealt with directly. This is sufficient to guarantee the moment condition is satisfied for all integers $p \geq 2$, as proved in the following lemma.

**Lemma 2.5** *Suppose the Craig-Bernstein moment condition holds for all even integers greater than or equal to* 2*, that is there exists an* $h > 0$ *such that*

$$E\left[|X - E[X]|^{2k}\right] \leq \frac{(2k)! var(X) h^{2k-2}}{2} \tag{2.62}$$

*for* $k \geq 2$ *since the* $k = 1$ *case is satisfied trivially for any* $h$*. Then the condition holds also for the odd absolute moments,*

$$E\left[|X - E[X]|^{2k-1}\right] \leq \frac{(2k-1)! var(X) \tilde{h}^{2k-3}}{2} \tag{2.63}$$

*for* $k \geq 2$ *with* $\tilde{h} = 2h$*. Thus*

$$E\left[|X - E[X]|^p\right] \leq \frac{p! var(X)(2h)^{p-2}}{2}, \quad p \geq 2. \tag{2.64}$$

**Proof:** For ease of notation, let $Z = X - E[X]$. Hölder's Inequality states, for any random variables $A$ and $B$,

$$E[|AB|] \leq E[|A|^p]^{1/p} E[|B|^q]^{1/q} \tag{2.65}$$

where $1 < p, q < \infty$ and $1/p + 1/q = 1$. Take $A = Z$, $B = Z^{2k-2}$, and $p = q = 2$ to get

$$E[|Z|^{2k-1}] \leq \sqrt{E[Z^2]E[Z^{4k-4}]} \tag{2.66}$$

where the absolute values inside the square root have been dropped because the exponents are even. Now

$$E[Z^{4k-4}] \leq \frac{(4k-4)!E[Z^2]h^{4k-6}}{2} \tag{2.67}$$

by assumption, so

$$\begin{aligned} E[|Z|^{2k-1}] &\leq \sqrt{\frac{(4k-4)!(E[Z^2])^2 h^{4k-6}}{2}} \\ &\leq \sqrt{\frac{(4k-4)!}{2}} E[Z^2] h^{2k-3}. \end{aligned} \tag{2.68}$$

We want to satisfy the following inequality by choice of $\tilde{h}$

$$E[|Z|^{2k-1}] \leq \frac{(2k-1)!E[Z^2]\tilde{h}^{2k-3}}{2}, \tag{2.69}$$

which means $\tilde{h}$ must satisfy

$$\frac{(2k-1)!}{2}\tilde{h}^{2k-3} \geq \sqrt{\frac{(4k-4)!}{2}} h^{2k-3}. \tag{2.70}$$

If we choose

$$\tilde{h} \geq \max_{k \geq 2} \left\{ \left( \frac{\sqrt{2(4k-4)!}}{(2k-1)!} \right)^{\frac{1}{2k-3}} \right\} h \tag{2.71}$$

then the moment condition will be satisfied for the odd exponents $2k-1$. An upper bound for the term in brackets is 2, as shown here.

For $k \geq 2$, the bound $(2k)! \leq 2^{2k}(k!)^2$ holds and can be verified by induction on $k$. This implies

$$\left( \frac{\sqrt{2(4k-4)!}}{(2k-1)!} \right)^{\frac{1}{2k-3}} \leq 2 \left( \frac{2\sqrt{2}}{2k-1} \right)^{\frac{1}{2k-3}}. \tag{2.72}$$

Now, the term in parentheses on the right hand side of (2.72) is always less than 1 for $k \geq 2$. The final step is to show that

$$\lim_{k \to \infty} \left( \frac{2\sqrt{2}}{2k-1} \right)^{\frac{1}{2k-3}} = 1, \tag{2.73}$$

which is verified by noting that

$$\begin{aligned} \lim_{k \to \infty} \log \left( \frac{2\sqrt{2}}{2k-1} \right)^{\frac{1}{2k-3}} &= \lim_{k \to \infty} \frac{1}{2k-3} \left( \log(2\sqrt{2}) - \log(2k-1) \right) \\ &= 0. \end{aligned} \tag{2.74}$$

Thus, the moment condition is satisfied for odd moments with $\tilde{h} = 2h$. Also, if the moment condition is satisfied for a given $h$, it is also satisfied for any $\tilde{h} \geq h$ so

$$E\left[|Z|^p\right] = E\left[|X - E[X]|^p\right] \leq \frac{p! var(X)(2h)^{p-2}}{2} \tag{2.75}$$

holds for all integers $p \geq 2$ as claimed. ∎

We will also need results for how sums and products of random variables behave with respect to the moment condition. For that, we have the following two lemmas.

**Lemma 2.6** *Let $Z = A + B$ be the sum of two zero-mean random variables with variances $var(A) = E[A^2]$ and $var(B) = E[B^2]$, not both zero, and such that $E[AB] \geq 0$. Suppose both $A$ and $B$ satisfy the moment condition for a given integer $p \geq 3$ with positive constants $h_A$ and $h_B$, respectively. That is*

$$E\left[|A|^p\right] \leq \frac{p! var(A) h_A^{p-2}}{2}, \; E\left[|B|^p\right] \leq \frac{p! var(B) h_B^{p-2}}{2}. \tag{2.76}$$

*Then*

$$E\left[|Z|^p\right] \leq \frac{p! var(Z) h_S^{p-2}}{2} \tag{2.77}$$

*where $h_S = 2^{1/(p-2)}(h_A + h_B)$.*

**Proof:** First, define $V_A = var(A)/(var(A) + var(B))$, $V_B = var(B)/(var(A) + var(B))$, $H_A = h_A/(h_A + h_B)$, and $H_B = h_B/(h_A + h_B)$. Use Minkowski's Inequality to write

$$\begin{aligned}
E[|A+B|^p] &\leq \left[E[|A|^p]^{1/p} + E[|B|^p]^{1/p}\right]^p \\
&\leq \frac{p!}{2}\left[\left(var(A)h_A^{p-2}\right)^{\frac{1}{p}} + \left(var(B)h_B^{p-2}\right)^{\frac{1}{p}}\right]^p \\
&= \frac{p!(var(A)+var(B))}{2} \times \left[\left(V_A h_A^{p-2}\right)^{\frac{1}{p}} + \left(V_B h_B^{p-2}\right)^{\frac{1}{p}}\right]^p \\
&= \frac{p!(var(A)+var(B))\,(h_A+h_B)^{p-2}}{2} \times \left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left(V_B H_B^{p-2}\right)^{\frac{1}{p}}\right]^p \\
&\leq \frac{p! var(A+B)\,(h_A+h_B)^{p-2}}{2} \times \left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left(V_B H_B^{p-2}\right)^{\frac{1}{p}}\right]^p,
\end{aligned} \tag{2.78}$$

where the last step follows from the assumption that $E[AB] \geq 0$, implying $var(A) + var(B) \leq var(A+B)$. Showing that

$$\left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left(V_B H_B^{p-2}\right)^{\frac{1}{p}}\right]^p \leq C^{p-2}, \tag{2.79}$$

or

$$\left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left(V_B H_B^{p-2}\right)^{\frac{1}{p}}\right]^{\frac{p}{p-2}} \leq C, \tag{2.80}$$

where $C = 2^{1/(p-2)}$ will complete the proof. Since $V_B = 1 - V_A$ and $H_B = 1 - H_A$, the result follows by maximizing

$$\left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left((1-V_A)(1-H_A)^{p-2}\right)^{\frac{1}{p}}\right]^{\frac{p}{p-2}} \tag{2.81}$$

by choice of $H_A$, $V_A$, and $p$. The same values of $H_A$ and $V_A$ will maximize

$$\left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left((1-V_A)(1-H_A)^{p-2}\right)^{\frac{1}{p}}\right], \tag{2.82}$$

and simple calculus shows that, for $p \neq 1$, the maximum occurs when $H_A = V_A = 1/2$. Thus

$$\left[\left(V_A H_A^{p-2}\right)^{\frac{1}{p}} + \left(V_B H_B^{p-2}\right)^{\frac{1}{p}}\right]^{\frac{p}{p-2}} \leq 2^{\frac{1}{p-2}}. \tag{2.83}$$

Choosing $h_S = C(h_A + h_B) = 2^{1/(p-2)}(h_A + h_B)$ gives

$$E[|Z|^p] \leq \frac{p! var(Z) h_S^{p-2}}{2}, \tag{2.84}$$

concluding the proof. ∎

**Lemma 2.7** *Let $Z = AB$ be the product of $A$ and $B$, two independent zero-mean random variables satisfying general moment conditions for a given integer $p \geq 3$. That is,*

$$E\left[|A|^p\right] \leq C_p^A var(A) h_A^{p-2}, \;\; E\left[|B|^p\right] \leq C_p^B var(B) h_B^{p-2} \tag{2.85}$$

*for some positive constants $h_A$ and $h_B$ and positive numbers $C_p^A$ and $C_p^B$ possibly depending on $p$. Then*

$$E\left[|Z|^p\right] \leq C_p^A C_p^B var(Z) h_P^{p-2} \tag{2.86}$$

*where $h_P = h_A h_B$.*

**Proof:** Because $A$ and $B$ are independent, we can write

$$E[|AB|^p] = E[|A|^p] E[|B|^p]. \tag{2.87}$$

Substituting in the given bounds and noting that $var(A)var(B) = var(AB)$ by independence, we get

$$E[|Z|^p] \leq C_p^A C_p^B var(Z) h_P^{p-2} \tag{2.88}$$

where $h_P = h_A h_B$. ∎

#### 2.1.8.2 Determination of the Bounding Constant $h$ for Noisy Randomized Projection Encoding

Equipped with the previous lemmas, we are now ready to determine the bounding constant $h$ for the Randomized Projection Encoding setup with binary basis elements $\{\phi_{i,j}\}$ taking values $\pm 1/\sqrt{n}$ with equal probability and additive Gaussian noise. For that we examine the moments of random variables of the form

$$U_j - E[U_j] = \frac{\|g\|^2}{n} - (\phi_j^T g)^2 - 2(\phi_j^T g)w_j, \tag{2.89}$$

where $g = f - f^*$. For ease of notation, let $\tilde{Z} = U_j - E[U_j]$. Since we are dealing with the moments in absolute value, it suffices to consider

$$-\tilde{Z} = (\phi_j^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi_j^T g)w_j. \tag{2.90}$$

To further simplify notation we drop the subscripts, let $Z = -\tilde{Z}$, and consider random variables of the form

$$Z = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)\tilde{w} \tag{2.91}$$

where $\tilde{w}$ is a zero-mean Gaussian random variable independent of $\phi$ having variance $\sigma^2$. Let $Z_1 = (\phi^T g)^2 - \frac{\|g\|^2}{n}$, $Z_2 = (\phi^T g)$, and $Z_3 = 2\tilde{w}$, then $Z = Z_1 + Z_2 Z_3$.

Our procedure for determining the bounding constant for $Z$ will be the following. First, we will find the bounding constants for the even moments of each of the components of $Z$ ($Z_1$, $Z_2$, and $Z_3$). Then, we will apply earlier lemmas to determine the bounding constants for the even moments of $Z_2 Z_3$, and finally $Z = Z_1 + Z_2 Z_3$. The last step will be to extend this final bounding constant so that it is valid for all moments.

Instead of explicitly stating the bounding constants in terms of $B$ and noise variance $\sigma^2$ here, we will derive relationships between the even moments of the $Z_i$, $i = 1, 2, 3$, and their respective variances. Later, we will use these relationships to obtain the explicit bounding constants that hold for all candidate functions $f$.

First, since $Z_3$ is Gaussian, we have

$$E[Z_3^{2k}] = (2k)!!(var(Z_3))^k \tag{2.92}$$

where $(2k)!! \equiv (1)(3)(5)\dots(2k-1)$ for integers $k \geq 1$ and $0!! \equiv 1$. Now we define the multinomial coefficient as

$$\begin{aligned} \binom{N}{K_1 \dots K_{n-1}} &\equiv \binom{N}{K_1}\binom{K_1}{K_2}\cdots\binom{K_{n-2}}{K_{n-1}} \\ &= \frac{N!}{K_1! K_2! \dots K_{n-1}!(N - K_{n-1})!} \end{aligned} \tag{2.93}$$

so that for $n = 2$ this is just the binomial coefficient $\binom{N}{K_1}$. Using this, we can write the even powers of $Z_2$ as

$$Z_2^{2k} = (\phi^T g)^{2k} = \sum_{i_1=0}^{2k} \sum_{i_2=0}^{i_1} \cdots \sum_{i_{n-1}=0}^{i_{n-2}} \left[ \binom{2k}{i_1 \dots i_{n-1}} [\phi_1 g_1]^{2k-i_1} [\phi_2 g_2]^{i_1-i_2} \dots [\phi_n g_n]^{i_{n-1}} \right]. \tag{2.94}$$

All $\phi_i$ raised to an odd power are zero in expectation while $E[\phi_i^m] = (1/\sqrt{n})^m$ if $m$ is even. Using this fact and taking the expectation, we get

$$E[(\phi^T g)^{2k}] = \sum_{i_1=0}^{k} \sum_{i_2=0}^{i_1} \cdots \sum_{i_{n-1}=0}^{i_{n-2}} \left[ \binom{2k}{2i_1 \dots 2i_{n-1}} n^{-k} g_1^{2k-2i_1} g_2^{2i_1-2i_2} \dots g_n^{2i_{n-1}} \right]. \tag{2.95}$$

We would like to "collapse" this sum back into a closed form expression, but for that we need another lemma.

**Lemma 2.8** *The multinomial coefficient satisfies the following equality*

$$\binom{2k}{2i_1 \dots 2i_{n-1}} = \binom{k}{i_1 \dots i_{n-1}} \frac{(2k)!!}{(2i_1)!!(2i_2)!! \dots (2i_{n-1})!!(2k-2i_{n-1})!!}. \tag{2.96}$$

**Proof:** For even integers $2j$ where $j$ is an integer satisfying $j \geq 0$, we can write $(2j)!! = (2j)!/(j!\ 2^j)$. Substituting this into the above equation for all terms $(2j)!!$ we see that the right hand side becomes

$$\binom{k}{i_1 \dots i_{n-1}} \binom{2k}{2i_1 \dots 2i_{n-1}} / \binom{k}{i_1 \dots i_{n-1}} = \binom{2k}{2i_1 \dots 2i_{n-1}}. \tag{2.97}$$

∎

Using Lemma 2.8 we can rewrite the sum in (2.95) as

$$E[(\phi^T g)^{2k}] = \sum_{i_1=0}^{k} \sum_{i_2=0}^{i_1} \cdots \sum_{i_{n-1}=0}^{i_{n-2}} \left[ \binom{k}{i_1 \dots i_{n-1}} \frac{(2k)!! n^{-k}}{(2i_1)!! \dots (2k-2i_{n-1})!!} g_1^{2k-2i_1} g_2^{2i_1-2i_2} \dots g_n^{2i_{n-1}} \right]. \tag{2.98}$$

Now observe that the following bound holds

$$1 \leq \frac{(2k)!!}{(2i_1)!! \dots (2k-2i_{n-1})!!} \leq (2k)!!, \tag{2.99}$$

so we obtain bounds on the even moments of $\phi^T g$,

$$\left( \frac{\|g\|^2}{n} \right)^k \leq E[(\phi^T g)^{2k}] \leq (2k)!! \left( \frac{\|g\|^2}{n} \right)^k. \tag{2.100}$$

Since $E[(\phi^T g)^2] = \left( \|g\|^2/n \right) = var(\phi^T g)$, we have

$$E[Z_2^{2k}] \leq (2k)!! \left( var(Z_2) \right)^k. \tag{2.101}$$

Finally, we need the bounding constant for $Z_1 = (\phi^T g)^2 - \|g\|^2/n$. We can write

$$Z_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} g_i g_j (\phi_i \phi_j - E[\phi_i \phi_j]). \tag{2.102}$$

Notice that if $i = j$ that term in the sum is zero, and we have symmetry such that the $i = k, j = l$ term is equal to the $i = l, j = k$ term so we can write

$$Z_1 = \sum_{i<j}^{n} 2\phi_i \phi_j g_i g_j. \tag{2.103}$$

The random variable $Z_1$ is a Rademacher chaos of order 2. Bounds for the absolute moments of such random variables were derived in [29], from which the following result is obtained. We refer the reader to that work for the proof.

**Lemma 2.9 (Bonami's Inequality)** *Let $F$ be a Rademacher chaos of order $d$. That is, $F$ is of the form*

$$F = \sum_{i_1<\cdots<i_d}^{n} a_{i_1 \ldots i_d} \phi_{i_1} \cdots \phi_{i_d} \tag{2.104}$$

*where $d \geq 1$, $a_{i_1 \ldots i_d}$ are real or complex numbers, and $\{\phi_{i_j}\}$ are i.i.d. random variables taking the values $\pm 1$ each with probability $1/2$. Then*

$$E\left[|F|^p\right] \leq (p-1)^{pd/2} \left(E\left[F^2\right]\right)^{p/2} \tag{2.105}$$

*holds for $p \geq 2$.*

We apply this result by letting $a_{ij} = 2g_i g_j/n$ and $p = 2k$. Then

$$\begin{aligned} E\left[Z_1^{2k}\right] &\leq (2k-1)^{2k} \left(E\left[Z_1^2\right]\right)^k \\ &= (2k-1)^{2k} \left(var(Z_1)\right)^k. \end{aligned} \tag{2.106}$$

We will now collect the results derived above and determine bounding constants that hold for all candidate functions $f$. Our random variable of interest is $Z = Z_1 + Z_2 Z_3$ where $Z_1 = (\phi^T g)^2 - \|g\|^2/n$, $Z_2 = (\phi^T g)$, and $Z_3 = 2\tilde{w}$. The variance of $Z_1$ can be obtained from [29] or by direct calculation and is given in closed form by

$$\begin{aligned} var(Z_1) = E[Z_1^2] &= E[(\phi^T g)^4] - \left(\frac{\|g\|^2}{n}\right)^2 \\ &= \frac{4}{n^2} \sum_{i<j} (g_i g_j)^2 \equiv \beta_g^2. \end{aligned} \tag{2.107}$$

To obtain the bounding constant for this term we simplify notation by letting $2k = p$, and upper bound the right-hand side of (2.106) by choice of $h$. That is,

$$(p-1)^p \beta_g^p \leq \frac{p!}{2} \beta_g^2 h^{p-2}, \tag{2.108}$$

for which $h$ must satisfy

$$h \geq \beta_g \max_{p \geq 3} \left\{ \left( \frac{2(p-1)^p}{p!} \right)^{\frac{1}{p-2}} \right\}. \tag{2.109}$$

It is straightforward to verify, using Stirling's Formula, that

$$\max_{p\geq 3}\left\{\left(\frac{2(p-1)^p}{p!}\right)^{\frac{1}{p-2}}\right\} = e \tag{2.110}$$

and so $h = \beta_g e$, where $e$ is the base of the natural logarithm.

To find a constant $h_1$ such that the moment condition is satisfied for $Z_1$ for any $g$, we solve

$$h_1 = \max_g \{\beta_g\}\, e = \max_g \left\{\sqrt{\frac{4}{n^2}\sum_{i<j}(g_i g_j)^2}\right\} e \tag{2.111}$$

subject to the energy constraint $\|g\|^2 \leq 4nB^2$. A straightforward application of Lagrange Multipliers shows that for a fixed nonzero energy $\|g\|^2 = C$, the variance of $Z_1$ is maximized when $|g_1| = |g_2| = \cdots = |g_n|$. Thus we let $g_i^2 = 4B^2$ to obtain

$$\begin{aligned} h_1 &= \max_g \left\{\sqrt{\frac{4}{n^2}\sum_{i<j}(g_i g_j)^2}\right\} e \\ &= \max\left\{\sqrt{\frac{4}{n^2}\frac{n(n-1)}{2}16B^4}\right\} e \\ &= 4\sqrt{2}B^2 e, \end{aligned} \tag{2.112}$$

and so

$$E[Z_1^{2k}] \leq \frac{(2k)!}{2} var(Z_1) h_1^{2k-2}. \tag{2.113}$$

The next term, $Z_2$, satisfies

$$\begin{aligned} E[Z_2^{2k}] &\leq (2k)!!var(Z_2)\max_g\left\{\sqrt{\left(\frac{\|g\|^2}{n}\right)}\right\}^{2k-2} \\ &\leq (2k)!!var(Z_2)(2B)^{2k-2} = (2k)!!var(Z_2)h_2^{2k-2} \end{aligned} \tag{2.114}$$

where $h_2 = 2B$, and $Z_3$ satisfies

$$E[Z_3^{2k}] \leq (2k)!!var(Z_3)\,(2\sigma)^{2k-2} = (2k)!!var(Z_3)h_3^{2k-2} \tag{2.115}$$

where $h_3 = 2\sigma$.

Notice first that $Z_2$ and $Z_3$ are independent, zero-mean, and both satisfy moment conditions as shown above. Applying Lemma 2.7 gives

$$\begin{aligned} E[(Z_2 Z_3)^{2k}] &\leq ((2k)!!)^2 var(Z_2 Z_3)(4B\sigma)^{2k-2} \\ &\leq \frac{(2k)!}{2} var(Z_2 Z_3)(4B\sigma)^{2k-2} \end{aligned} \tag{2.116}$$

for all integers $k \geq 1$, where the last step follows by inspection:

$$\begin{aligned} \frac{(2k)!}{2} &= (1)(1)(3)(4)\dots(2k-1)(2k) \\ &\geq ((2k)!!)^2 \\ &= (1)(1)(3)(3)\dots(2k-1)(2k-1). \end{aligned} \tag{2.117}$$

Now since $E[Z_1(Z_2Z_3)] = 0$ we can apply Lemma 2.6 to get

$$E[(Z_1 + Z_2Z_3)^{2k}] \leq \frac{(2k)!}{2} var(Z_1 + Z_2Z_3)[\sqrt{2}(4\sqrt{2}B^2e + 4B\sigma)]^{2k-2} \tag{2.118}$$

for integers $k \geq 1$. Finally, to extend this result to all moments we use Lemma 2.5 to obtain

$$E[|Z|^p] = E[|Z_1 + Z_2Z_3|^p] \leq \frac{p!}{2} var(Z_1 + Z_2Z_3)[2\sqrt{2}(4\sqrt{2}B^2e + 4B\sigma)]^{2k-2} \tag{2.119}$$

for all integers $p \geq 2$, so the constant $h = 8\sqrt{2}B(\sqrt{2}Be + \sigma) = 16B^2e + 8\sqrt{2}B\sigma$.

#### 2.1.8.3 Working with Projected Gaussian Noise

In this section, we prove Corollary 2. Suppose the signal is contaminated with additive zero-mean Gaussian noise, prior to projective sampling. In this case, the observations are

$$y_j = \sum_{i=1}^{n} \phi_{i,j} \left(f_i^* + \eta_{i,j}\right) + w_j, \quad j = 1, \dots, k \tag{2.120}$$

where the $\{\eta_{i,j}\}$ are i.i.d. zero-mean Gaussian random variables with variance $\sigma_s^2$ and assumed to be independent $\{\phi_{i,j}\}$ and $\{w_j\}$. With respect to the moment condition, the random variable of interest becomes

$$Z = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)(\phi^T \eta) + 2(\phi^T g)w, \tag{2.121}$$

where the subscripts have been dropped for ease of notation. This is equivalent in distribution to

$$Z_{eq} = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)\tilde{w}, \tag{2.122}$$

where $\tilde{w}$ is a zero-mean Gaussian random variable, independent of $\{\phi_i\}$, with variance $\sigma^2 + \sigma_s^2$, as shown below.

**Lemma 2.10** *The random variables $(\phi^T g)$ and $(\phi^T \eta)$ are independent and thus $(\phi^T \eta)$ is independent of $\{\phi_i\}$.*

**Proof:** To prove independence, we will show that the joint characteristic function of $(\phi^T g)$ and $(\phi^T \eta)$ factorizes. The characteristic function is $M_{(\phi^T g),(\phi^T \eta)} = E[e^{(j\nu_1(\phi^T g) + j\nu_2(\phi^T \eta))}]$. Note that $E[e^{j\nu_2(\phi^T \eta)} \mid \phi] = e^{-\nu_2^2\sigma_s^2/2}$, which does not depend on $\phi$. The result follows from this observation. Furthermore, note that

$e^{-\nu_2^2\sigma_s^2/2}$ is the characteristic function of a zero-mean Gaussian random variable with variance $\sigma_s^2$, implying that $\phi^T\eta$ is Gaussian distributed. ∎

This proof immediately gives the following corollary.

**Corollary 2.11** *The random variables $(\phi^T g)(\phi^T\eta)$ and $(\phi^T g)\eta_i$ are equivalent in distribution.*

By the above results, we see that our random variable of interest

$$Z = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)(\phi^T\eta) + 2(\phi^T g)w \tag{2.123}$$

is equivalent in distribution to

$$Z_{eq} = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)(\eta_i + w). \tag{2.124}$$

Let $\tilde{w} = (\eta_i + w)$ and notice that $\tilde{w}$ is Gaussian with mean zero and variance $\sigma^2 + \sigma_s^2$ and independent of $\{\phi_i\}$. Then

$$Z = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)(\phi^T\eta) + 2(\phi^T g)w \tag{2.125}$$

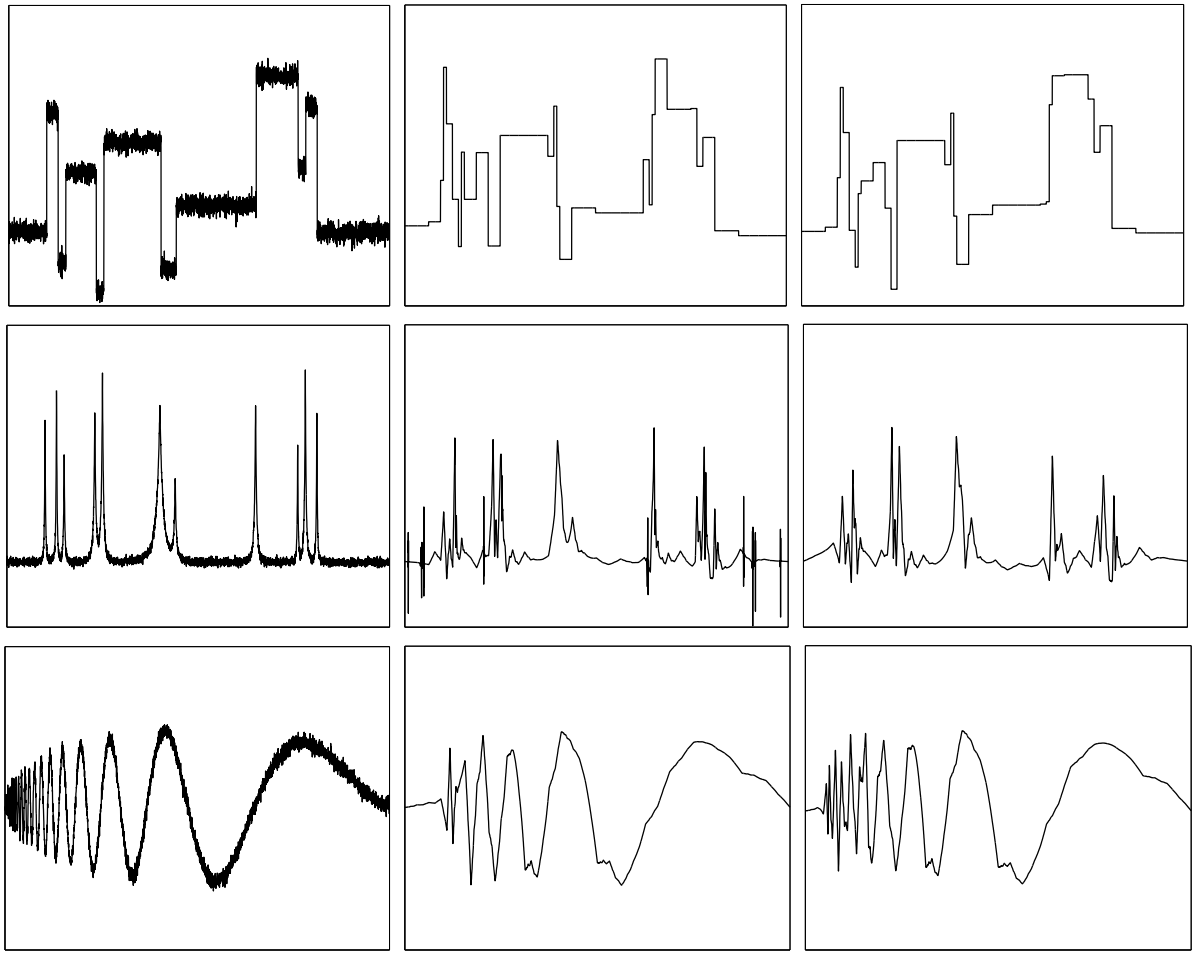
is equivalent in distribution to

$$Z_{eq} = (\phi^T g)^2 - \frac{\|g\|^2}{n} + 2(\phi^T g)\tilde{w}, \tag{2.126}$$

and thus Theorems 1 and 2 and Corollary 2.3 apply in this situation, as well.

### 2.1.9 Acknowledgement

The authors would like to thank Rui Castro for his assistance in all stages of the work presented here. The authors would also like to thank Emmanuel Candès and the anonymous reviewers for their careful readings of the paper and helpful comments.



(a) (b) (c)

Figure 2.1 Reconstruction examples using Blocks, Bumps, and Doppler test signals of length 4096. Column ($a$) shows the original signals with an equivalent level of additive per-pixel noise. Columns ($b$) and ($c$) show reconstructions obtained from 600 and 1200 projections, respectively.

## 2.2 Toeplitz Compressed Sensing Matrices with Applications to Sparse Channel Estimation

Compressed sensing (CS) has recently emerged as a powerful signal acquisition paradigm. In essence, CS enables the recovery of high-dimensional but sparse (or nearly sparse) vectors from relatively few linear observations in the form of projections of the signal onto a collection of test vectors. Existing results show, for example, that if the entries of the test vectors are independent realizations of random variables with certain distributions, such as zero-mean Gaussian, then with high probability the resulting observations sufficiently encode the information in the unknown signal and recovery can be accomplished by solving a tractable convex optimization. This work provides a significant extension of current CS theory. A novel technique is proposed that allows theoretical treatment of CS in settings where the entries of the test vectors exhibit structured statistical dependencies, from which it follows that CS can be effectively utilized in linear, time-invariant (LTI) system identification problems. An immediate application is in the area of sparse channel estimation, where the main results of this work can be applied to the recovery of sparse (or nearly sparse) wireless multipath channels. Specifically, it is shown in the paper that time-domain probing of a wireless channel with a (pseudo-)random binary sequence, along with the utilization of CS reconstruction techniques, provides significant improvements in estimation accuracy when compared with traditional least-squares based linear channel estimation strategies. Abstract extensions of the main results, utilizing the theory of equitable graph coloring to tolerate more general statistical dependencies across the test vectors, are also discussed.

### 2.2.1 Introduction

Compressed sensing (CS) describes a new signal acquisition paradigm in which sparse, high-dimensional vectors $\boldsymbol{\beta} \in \mathbb{R}^n$ can be accurately recovered from a small number of linear observations of the form $\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{y} \in \mathbb{R}^k$. The conditions under which CS succeeds depend on the structure of the measurement matrix $\boldsymbol{X} \in \mathbb{R}^{k \times n}$. Specifically, recent theoretical results establish that the observations $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}$ can be used to efficiently recover any "sparse enough" signal provided that the matrix $\boldsymbol{X}$ satisfies the so-called *restricted isometry property* (RIP) [14, 16, 30]. Essentially, the RIP condition guarantees that all column submatrices of $\boldsymbol{X}$ (the $k \times m$ submatrices formed by all subsets of $m$ columns, where $m$ is proportional to the level of sparsity in $\boldsymbol{\beta}$) are well-conditioned. The use of such measurement matrices to recover sparse vectors leads to remarkable results. First, the number of observations required for accurate recovery, $k$, is on the order of the number of nonzero entries in the signal, which can be far fewer than the ambient signal dimension $n$. In addition, signal recovery can be accomplished by solving a tractable convex optimization. And both of these results remain true even when $\boldsymbol{y}$ is corrupted by some (deterministic or stochastic) additive noise $\boldsymbol{\eta} \in \mathbb{R}^k$,

where "recovery" in these settings means that the signal estimate is close to the true signal in terms of a suitable error metric (such as the mean squared error).

This paper explores whether the sparsity-exploiting power of CS can be used for the identification of discrete, linear, time-invariant (LTI) systems. Specifically, we examine the effectiveness of random probing of LTI systems having sparse impulse responses, coupled with the utilization of CS reconstruction methods. The practical importance of this problem is evidenced by many wireless communication applications in which the underlying multipath channel can be modeled as an LTI system with a sparse impulse response [31]. Compared to the conventional channel estimation methods that do not explicitly account for the underlying multipath sparsity, reliable estimation of the channel impulse response in these settings can lead to significant reductions in transmission energy and improvements in spectral efficiency.

Existing results in CS show that if the entries of the observation (measurement) matrix are independent realizations of random variables with certain distributions, such as zero-mean Gaussian, then with high probability the resulting matrix satisfies the RIP [32]. Toeplitz matrices are matrices having constant diagonals and they arise naturally in the context of estimation of wireless multipath channels. The major contribution of this work is a significant extension of CS theory to observation matrices containing statistical dependencies across rows, from which it follows that random Toeplitz matrices satisfy the RIP with high probability. As a result, recent advances from the theory of CS can be leveraged to devise quantitative error bounds for convex/linear programming based sparse channel estimation schemes. Our proofs rely on a novel technique that facilitates analysis of the (structured) statistical dependencies arising from the Toeplitz structure and, in general, our techniques can be applied in certain settings to obtain deviation bounds for both the $\ell_2$-norms of random vectors having dependent entries, and inner products between certain dependent random vectors. The proofs and techniques outlined here generalize and build upon our own previous works, which were the first to provide theoretical performance guarantees for CS using random Toeplitz-structured matrices [31, 33].

The remainder of this paper is organized as follows. In the rest of this section, we provide the requisite background on existing results in CS. In Section 2.2.2, we describe how estimation of sparse wireless multipath channels fits into the CS framework. In particular, we show that time-domain probing of a wireless channel with a (pseudo-)random binary sequence, along with the utilization of CS reconstruction techniques, provides significant improvements in estimation accuracy when compared with traditional least-squares based linear channel estimation strategies. The major contributions of the paper appear in Section 2.2.3, where we establish the RIP for random Toeplitz matrices comprised of either Gaussian or bounded random variables. Finally, in Section 2.2.4, we present extensions of the main results of the paper to accommodate more general statistical dependencies, and we discuss connections with previous works.

#### 2.2.1.1 Compressed Sensing and the Restricted Isometry Property

Consider the problem of recovering an unknown vector $\boldsymbol{\beta} \in \mathbb{R}^n$ from a collection of linear observations,

$$y_j = \sum_{i=1}^{n} x(j)_i \ \beta_i = \boldsymbol{x}(j)'\boldsymbol{\beta}, \quad j = 1, \ldots, k \tag{2.127}$$

where the vectors $\boldsymbol{x}(j) \in \mathbb{R}^n$ are specified "test vectors." The observation process can be written more compactly in the matrix formulation $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}$ where $\boldsymbol{y} \in \mathbb{R}^k$ and $\boldsymbol{X}$ is the $k \times n$ observation matrix whose rows are the test vectors. This very general signal model encompasses a wide variety of applications, including magnetic resonance imaging, digital imaging, and radio frequency surveillance. When the number of observations, $k$, equals or exceeds the dimension of the unknown signal, $n$, (the so-called overdetermined setting) then results from classical linear algebra show that any unknown signal can be recovered exactly using a suitable set of test vectors. The complete set of basis vectors from any orthonormal transform suffices, for example.

Compressed sensing (CS) primarily addresses the question of what is possible when the number of observations is *fewer* than the dimension of the unknown signal, the so-called underdetermined setting. The seminal works in CS established that signals can theoretically be recovered exactly from such incomplete observations, provided the signals are sparse. Further, CS is a viable practical technique because recovery is tractable—it can be accomplished by convex optimization [1–3].

Proofs of these remarkable initial results all rely on the same properties of the observation matrix, namely that the submatrix of the observation matrix corresponding to the true signal subspace should behave almost like an orthonormal matrix. One concise way to quantify this is by the restricted isometry property (RIP), first introduced in [14]. The RIP, defined below, can be leveraged to establish a series of fundamental results in CS.

**Definition 2.12 (Restricted Isometry Property)** *The observation matrix $\boldsymbol{X}$ is said to satisfy the restricted isometry property of order $S$ with parameter $\delta_S \in (0, 1)$, if*

$$(1 - \delta_S)\|\boldsymbol{z}\|_{\ell_2}^2 \leq \|\boldsymbol{X}\boldsymbol{z}\|_{\ell_2}^2 \leq (1 + \delta_S)\|\boldsymbol{z}\|_{\ell_2}^2 \tag{2.128}$$

*holds for all $S$-sparse vectors $\boldsymbol{z} \in \mathbb{R}^n$.*[2]

**Remark 2.13** *We will sometimes make use of shorthand notation, where instead of saying that a matrix $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$, we will say $\boldsymbol{X}$ satisfies $RIP(S, \delta_S)$.*

While there is currently no known way to test in polynomial time whether a given matrix satisfies RIP, certain probabilistic constructions of matrices can be shown to satisfy RIP with high probability [1–3, 14]. The following result is representative [32].

[2] A vector $\boldsymbol{z} \in \mathbb{R}^n$ is said to be $S$-sparse if $\|\boldsymbol{z}\|_{\ell_0} \leq S$, where $\|\boldsymbol{z}\|_{\ell_0}$ counts the number of nonzero entries in $\boldsymbol{z}$.

**Lemma 2.14** *Let $\boldsymbol{X}$ be a $k \times n$ matrix whose entries are independent and identically distributed (i.i.d.), drawn from one of the following zero-mean distributions, each having variance $1/k$:*

- $x_{i,j} \sim \mathcal{N}(0, 1/k)$,
- $x_{i,j} \sim \begin{cases} 1/\sqrt{k} & \text{with prob. } 1/2 \\ -1/\sqrt{k} & \text{w.p. } 1/2 \end{cases}$,
- $x_{i,j} \sim \begin{cases} \sqrt{3/k} & \text{w.p. } 1/6 \\ 0 & \text{w.p. } 2/3 \\ -\sqrt{3/k} & \text{w.p. } 1/6 \end{cases}$.

*For any $\delta_S \in (0,1)$ and any $c_1 < \delta_S^2(3-\delta_S)/48$, set*

$$c_2 = \frac{192 \log(12/\delta_S)}{3\delta_S^2 - \delta_S^3 - 48c_1}. \tag{2.129}$$

*Then whenever $k \geq c_2 S \log n$, $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability at least $1 - \exp(-c_1 k)$.*

The initial contributions to the theory of CS established, essentially, that any sparse signal can be recovered exactly from a collection of linear observations if the corresponding observation matrix satisfies RIP. The following result is a generalization that also describes the recovery of signals that are not exactly sparse.

**Lemma 2.15 (Noiseless Recovery [30])** *Let $\boldsymbol{X}$ be an observation matrix satisfying RIP of order $2S$ with parameter $\delta_{2S} < \sqrt{2}-1$, and let $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}$ be a vector of observations of any unknown signal $\boldsymbol{\beta}$. The estimate $\widehat{\boldsymbol{\beta}}$ obtained as the solution of*

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{z} \in \mathbb{R}^n}{\arg\min} \|\boldsymbol{z}\|_{\ell_1} \text{ subject to } \boldsymbol{y} = \boldsymbol{X}\boldsymbol{z} \tag{2.130}$$

*satisfies*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq c_0 \frac{\|\boldsymbol{\beta}_S - \boldsymbol{\beta}\|_{\ell_1}^2}{S}, \tag{2.131}$$

*where $\boldsymbol{\beta}_S$ is the vector formed by setting all but the $S$ largest entries (in magnitude) of $\boldsymbol{\beta}$ to zero, and*

$$c_0 = 4\left(\frac{1 - \delta_{2S} + \sqrt{2}\delta_{2S}}{1 - \delta_{2S} - \sqrt{2}\delta_{2S}}\right)^2. \tag{2.132}$$

Note that in the case where the signal $\boldsymbol{\beta}$ has no more than $S$ nonzero entries, this result guarantees that signal recovery is exact. Further, as shown in Lemma 2.14, there exist matrices for which this recovery is possible (with high probability) using only $k = O(S \log n)$ observations. Finally, the optimization program (2.130), which goes by the name of *basis pursuit*, is computationally tractable because it can be recast as a linear program. In this sense the optimality of noiseless CS is evident.

Suppose now that the observations are corrupted by some additive noise. That is, the observations are given by $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta}$, where $\boldsymbol{\eta} \in \mathbb{R}^k$ is either deterministic, or it is a vector whose entries are i.i.d. realizations of some zero-mean random variable. In either case, it turns out that CS can be used to obtain a stable recovery of the unknown signal.

The first work to establish theoretical results in the stochastic noise setting was [4], which used a reconstruction procedure that required a combinatorial search. The result presented here gives similar reconstruction error bounds, but is based on the RIP condition and utilizes a tractable convex optimization that goes by the name of *Dantzig selector*. The original specification of the result in [16] assumed a specific signal class, but the proof actually provides a more general oracle result.

**Lemma 2.16 (Dantzig Selector—Recovery in Stochastic Noise [16])** *Let $\boldsymbol{X}$ be an observation matrix satisfying RIP of order $2S$ such that $\delta_{2S} < \sqrt{2} - 1$ for some integer $S \geq 1$, and let $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta}$ be a vector of noisy observations of $\boldsymbol{\beta} \in \mathbb{R}^n$, where the entries of $\boldsymbol{\eta}$ are i.i.d. zero-mean Gaussian variables with variance $\sigma^2$. Choose $\lambda_n = \sqrt{2(1+a)(1+\delta_1)\log n}$ for any $a \geq 0$. The estimator*

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{z}\in\mathbb{R}^n}{\arg\min} \|\boldsymbol{z}\|_{\ell_1} \text{ subject to } \|\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{z})\|_{\ell_\infty} \leq \sigma\lambda_n \tag{2.133}$$

*satisfies*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq c_0' \min_{1\leq m\leq S} \left( \sigma\lambda_n\sqrt{m} + \frac{\|\boldsymbol{\beta}_m - \boldsymbol{\beta}\|_{\ell_1}}{\sqrt{m}} \right)^2 \tag{2.134}$$

*with probability at least $1 - \left(\sqrt{\pi(1+a)\log n}\cdot n^a\right)^{-1}$. The constant $c_0' = 16/\left(1 - \delta_{2S} - \sqrt{2}\delta_{2S}\right)^2$, and as in the previous lemma, $\boldsymbol{\beta}_m$ is the vector formed by setting all but the $m$ largest entries (in magnitude) of $\boldsymbol{\beta}$ to zero.*

**Remark 2.17** *The sufficient condition stated in the original result in [16] was $\delta_{2S} + \theta_{S,2S} < 1$, where $\theta_{S,2S}$ is called the $S, 2S$-*restricted orthogonality parameter *of the matrix $\boldsymbol{X}$. In general, the $S, S'$-restricted orthogonality parameter is such that*

$$\left|(\boldsymbol{X}\boldsymbol{\alpha})'(\boldsymbol{X}\boldsymbol{\beta})\right| \leq \theta_{S,S'}\|\boldsymbol{\alpha}\|_{\ell_2}\|\boldsymbol{\beta}\|_{\ell_2} \tag{2.135}$$

*holds for all vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ having no more than $S$ and $S'$ nonzero entries, respectively, and such that the nonzero entries of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ occur at disjoint subsets of indices. However, it can be shown that $\theta_{S,2S} \leq \sqrt{2}\delta_{2S}$, from which the RIP condition stated in the lemma follows.*

Notice that the reconstruction error in (2.134) is essentially comprised of two factors. One factor is due to the "estimation error" that arises from determining $m$ unknown quantities from noisy data, while the other is due to the "approximation error" or bias arising from estimating the unknown vector using only $m$

components. For a given signal class, the best rate of error decay is obtained by balancing the two terms. That is, the best value of $m$ is the value $m^*$ such that

$$\frac{\|\boldsymbol{\beta}_{m^*} - \boldsymbol{\beta}\|_{\ell_1}}{m^*} = \sigma\lambda_n. \tag{2.136}$$

Thus, to make the optimal rates achievable, the observation matrix should be chosen to recover signals with sparsity $S$ that is at least as large as the "effective sparsity" $m^*$.

Finally, when the perturbation (noise) vector is deterministic, a relaxation of the optimization program in (2.130) guarantees stable signal recovery. The result given below is from [30], which refines the initial result of [15].

**Lemma 2.18 (Recovery in Deterministic Noise [30])** *Let $\boldsymbol{X}$ be an observation matrix satisfying RIP of order $2S$ such that $\delta_{2S} < \sqrt{2} - 1$ for some integer $S \geq 1$, and let $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta}$ be a vector of noisy observations of $\boldsymbol{\beta} \in \mathbb{R}^n$, where $\|\boldsymbol{\eta}\|_{\ell_2} \leq \epsilon$. The solution of*

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{z}\in\mathbb{R}^n}{\arg\min} \|\boldsymbol{z}\|_{\ell_1} \text{ subject to } \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{z}\|_{\ell_2} \leq \epsilon \tag{2.137}$$

*satisfies*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq c_0'' \left(\epsilon + \frac{\|\boldsymbol{\beta}_S - \boldsymbol{\beta}\|_{\ell_1}}{\sqrt{S}}\right)^2, \tag{2.138}$$

*where $\boldsymbol{\beta}_S$ is as defined earlier and*

$$c_0'' = \frac{16\,(1+\delta_{2S})}{\left(1 - \delta_{2S} - \sqrt{2}\delta_{2S}\right)^2}. \tag{2.139}$$

Note that this result differs significantly from the stochastic noise result. Indeed, applying this result directly to the stochastic noise setting (in which case $\|\boldsymbol{\eta}\|_{\ell_2} \sim \sqrt{k}\sigma$) would yield an approximation error that grows like the *number of observations* times the noise power. On the other hand, the Dantzig selector results in a much better reconstruction error bound, with the estimation error scaling like the *sparsity level* times the noise power. In other words, the estimation error bound of the Dantzig selector is adaptive to the sparsity level, while the error bound of the relaxed $\ell_1$ minimization in (2.137) is not. The difference in the two reconstruction error bounds could be significant, especially when the number of observations is far greater than the sparsity (or effective sparsity).

### 2.2.2 Random Toeplitz Matrices and Sparse Channel Estimation

We now describe a different problem domain which yields an observation model similar to the canonical CS setting. Consider point-to-point communication between two single-antenna transceivers over a wideband wireless multipath channel. Such single-antenna communication channels can be characterized as discrete, linear, time-invariant systems—see, e.g., [31] for further details. Optimal demodulation and decoding in

wireless communication systems often requires accurate knowledge of the channel impulse response. Typically, this is accomplished by probing the channel with a known training sequence and linearly processing the channel output. Many real-world channels of practical interest, such as underwater acoustic channels [34], digital television channels [35] and residential ultrawideband channels [36], however, tend to have sparse or approximately sparse impulse responses. On the other hand, conventional linear channel estimation schemes, such as the least-squares method, fail to capitalize on the anticipated sparsity of the aforementioned channels. In contrast, it is established in this section that a channel estimate obtained as a solution to the Dantzig selector significantly outperforms a least-squares based channel estimate in terms of the mean squared error (MSE) when it comes to learning sparse (or approximately sparse) channels.

To begin with, let $\{x_i\}_{i=1}^{p}$, $p \in \mathbb{N}$, denote the training sequence, and consider using this sequence as the input to a wireless channel characterized by a finite (discrete) impulse response $\boldsymbol{\beta} \in \mathbb{R}^n$. The resulting observations $\boldsymbol{y} \in \mathbb{R}^{n+p-1}$ are described by the discrete-time convolution between the training signal $\boldsymbol{x}$ and the impulse response $\boldsymbol{\beta}$, and corruption by an additive noise vector $\boldsymbol{\eta}$: $\boldsymbol{y} = \boldsymbol{x} * \boldsymbol{\beta} + \boldsymbol{\eta}$. More specifically, if we use the notational convention that $x_i = 0$ for $i \notin \{1, 2, \ldots, p\}$, then each observation can be written as a sum,

$$y_j = \sum_{i=1}^{p} \beta_i \, x_{j+1-i} + \eta_j, \tag{2.140}$$

where the observations are corrupted by independent additive white Gaussian noise $\{\eta_j\}$ of variance $\sigma^2$ and contain a channel impulse response component only for $j = 1, \ldots, n+p-1$. The resulting input-output relation can be expressed as a matrix-vector product

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n+p-2} \\ y_{n+p-1} \end{bmatrix} = \begin{bmatrix} x_1 & & 0 \\ x_2 & \ddots & \\ \vdots & \ddots & x_1 \\ x_p & & x_2 \\ & \ddots & \vdots \\ 0 & & x_p \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} \tag{2.141}$$

and the goal is to obtain an estimate of the channel impulse response $\boldsymbol{\beta}$ from knowledge of the observations $\boldsymbol{y}$ and training signal $\boldsymbol{x}$. Note that, by symmetry, either the training signal or the impulse response could be rewritten as a convolution matrix in the above formulation. Writing it in this way casts the channel estimation problem into the canonical CS framework.

The purpose of channel estimation in communication systems is to aid them in achieving their primary objective, which is reliable communication of data (information) from one point to another. Further, because of the dynamic nature of the wireless medium, the impulse response of a channel is bound to change over

time [37]. As such, the input data sequence at the transmitter is periodically interspersed with the training sequence so as to maintain an up-to-date estimate of the channel impulse response at the receiver. We treat two facets of the sparse channel estimation problem in this section. The first one, corresponding to the lack of a "guard interval" of length $n-1$ between the data and training sequence, most closely resembles the canonical CS observation model, where the number of observations is far fewer than the length of the unknown signal. Specifically, consider a setting where the length of the training sequence $p = n + k - 1$ for some $k \geq 1$ and the training sequence is immediately preceded and succeeded by the data sequence. In this case, the first and last $n-1$ observations in (2.141) also contain contributions from the *unknown* data, rendering them useless for estimation purposes (the 0's in the convolution matrix in (2.141) would be replaced by the data sequence). Therefore, the channel estimation problem in this case reduces to reconstructing the unknown impulse response $\boldsymbol{\beta}$ from $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta}$, where the observation matrix $\boldsymbol{X}$ is a "partial" Toeplitz matrix of the form

$$\boldsymbol{X} = \begin{bmatrix} x_n & x_{n-1} & \dots & x_2 & x_1 \\ x_{n+1} & x_n & \dots & x_3 & x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n+k-1} & x_{n+k-2} & \dots & x_{k+1} & x_k \end{bmatrix}. \tag{2.142}$$

The second setting we consider corresponds to the case when the training sequence is immediately preceded and succeeded by $n-1$ zeros (i.e., a guard interval of length $n-1$ exists between the data and the training sequence). In this setting, the channel estimation problem corresponds to obtaining an estimate of the channel impulse response from the "full" set of observations described by (2.141). Notice that when $p \geq n$, the partial Toeplitz matrix described in (2.142) above is a submatrix of the observation matrix in this setting. In contrast, when $p < n$, every row of the observation matrix in this setting has at least one zero entry, and in the limiting case when $p = 1$, the observation matrix is just a scaled version of the $n \times n$ identity matrix.

The question we address in this section for both of the aforementioned settings is whether random binary probing, along with the use of a nonlinear Dantzig selector based estimator, can be employed to efficiently estimate a sparse channel, quantified by the condition $\|\boldsymbol{\beta}\|_{\ell_0} = S \ll n$. Note that initial theoretical analysis of CS systems that utilized random observation matrices relied inherently upon statistical independence among observations. The problem considered here is significantly more challenging—the Toeplitz structure of the (partial and full) observation matrices introduces statistical dependencies among observations and hence, existing techniques can no longer be employed. Instead, we develop a novel technique in Section 2.2.3 that facilitates analysis in the presence of such (structured) dependencies.

#### 2.2.2.1 MSE of Least-Squares Channel Estimates

Estimation of an unknown vector $\boldsymbol{\beta}$ from linear observation models of the form $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta}$ is a well-studied problem in the area of estimation theory—see, e.g., [38]. Traditionally, channel estimates are usually obtained from $\boldsymbol{y}$ by solving the least-squares (LS) problem (or a variant of it). Note that in the case that the observation matrix $\boldsymbol{X}$ is given by (2.142), LS solution requires that $k \geq n$ so as to obtain a meaningful channel estimate [38]. Under this assumption, the LS channel estimate is given by

$$\widehat{\boldsymbol{\beta}}_{LS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{2.143}$$

where the observation matrix $\boldsymbol{X}$ corresponds to (2.142) in the case of training without guard intervals and to the full Toeplitz matrix in the case of training with guard intervals. Below, we lower bound the MSE performance of an LS channel estimate corresponding to random binary probing (binary phase shift keying signaling).

**Theorem 2.19** *Let the training sequence $\{x_i\}_{i=1}^{p}$ be given by a sequence of i.i.d. binary random variables taking values $+1$ or $-1$ with probability $1/2$ each. Further, let $p = n + k - 1$ for some $k \geq n$ for the case of training without guard intervals. Then the MSE of the LS channel estimate $\widehat{\boldsymbol{\beta}}_{LS}$ is lower bounded by*

$$\mathbb{E}\left[\|\widehat{\boldsymbol{\beta}}_{LS} - \boldsymbol{\beta}\|_{\ell_2}^2\right] \geq \frac{n\sigma^2}{p} \tag{2.144}$$

*and*

$$\mathbb{E}\left[\|\widehat{\boldsymbol{\beta}}_{LS} - \boldsymbol{\beta}\|_{\ell_2}^2\right] \geq \frac{n\sigma^2}{k} \tag{2.145}$$

*for training with and without guard intervals, respectively. Further, the equality in the above two expressions hold if and only if the corresponding observation matrices have orthogonal columns.*

**Proof:** To establish this theorem, note that for both the cases of training with or without guard intervals

$$\mathbb{E}\left[\|\widehat{\boldsymbol{\beta}}_{LS} - \boldsymbol{\beta}\|_{\ell_2}^2\right] = \text{trace}\left\{(\boldsymbol{X}'\boldsymbol{X})^{-1}\right\}\sigma^2. \tag{2.146}$$

Further, let $\{\lambda_i\}_{i=1}^{n}$ denote the $n$ eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$. Then, from elementary linear algebra, we have

$$\begin{aligned}\text{trace}\left\{(\boldsymbol{X}'\boldsymbol{X})^{-1}\right\} &= \sum_{i=1}^{n}\frac{1}{\lambda_i} = n\left(\frac{\sum_{i=1}^{n}\frac{1}{\lambda_i}}{n}\right) \\ &\overset{(a)}{\geq} n\left(\frac{n}{\sum_{i=1}^{n}\lambda_i}\right) = \frac{n^2}{\text{trace}\left\{\boldsymbol{X}'\boldsymbol{X}\right\}}\end{aligned} \tag{2.147}$$

where $(a)$ follows from the arithmetic-harmonic means inequality. Also, from the arithmetic-harmonic means inequality, the equality in $(a)$ holds if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_n$, resulting in the condition that $\boldsymbol{X}$ must have orthogonal columns for the equality to hold in $(a)$. Finally, note that $\text{trace}\left\{\boldsymbol{X}'\boldsymbol{X}\right\} = np$ for the case of training with guard intervals, while $\text{trace}\left\{\boldsymbol{X}'\boldsymbol{X}\right\} = nk$ for the case of training without guard intervals and this completes the proof of the theorem. ∎

#### 2.2.2.2 MSE of Dantzig Selector Channel Estimates

We have seen in the previous section that the MSE of an LS channel estimate is lower bounded by (ambient dimension of $\boldsymbol{\beta}$) $\cdot\, \sigma^2/$(# of *effective* observations). Conventional channel learning techniques based on the LS criterion, however, fail to take into account the anticipated sparsity of the channel impulse response. To get an idea of the potential MSE gains that are possible when incorporating the sparsity assumption into the channel estimation strategy, we compare the performance of an LS based channel estimator to that of a channel estimation strategy that has been equipped with an *oracle*. The oracle does not reveal the true $\boldsymbol{\beta}$, but does inform us of the indices of nonzero entries of $\boldsymbol{\beta}$. Clearly this represents an ideal estimation strategy and one cannot expect to attain its performance level. Nevertheless, it is the target that one should consider.

To begin with, let $T_* \subset \{1, \ldots, n\}$ be the set of indices of the $S$ nonzero entries of $\boldsymbol{\beta}$ and suppose that an oracle provides us with the sparsity pattern $T_*$. Then an ideal channel estimate $\boldsymbol{\beta}^*$ can be obtained for both the cases of training with or without guard intervals by first forming a *restricted* LS estimator from $\boldsymbol{y}$

$$\boldsymbol{\beta}_{T_*} = (\boldsymbol{X}'_{T_*}\boldsymbol{X}_{T_*})^{-1}\boldsymbol{X}'_{T_*}\boldsymbol{y} \tag{2.148}$$

where $\boldsymbol{X}_{T_*}$ is a submatrix obtained by extracting the $S$ columns of $\boldsymbol{X}$ corresponding to the indices in $T_*$, and then setting $\boldsymbol{\beta}^*$ to $\boldsymbol{\beta}_{T_*}$ on the indices in $T_*$ and zero on the indices in $T_*^c$. Appealing to the proof of Theorem 2.19, the MSE of this oracle channel estimator can be lower bounded as

$$\begin{aligned} \mathbb{E}\left[\|\boldsymbol{\beta}^* - \boldsymbol{\beta}\|_{\ell_2}^2\right] &= \operatorname{trace}\left\{(\boldsymbol{X}'_{T_*}\boldsymbol{X}_{T_*})^{-1}\right\}\sigma^2 \\ &\geq \frac{S^2\sigma^2}{\operatorname{trace}\left\{\boldsymbol{X}'_{T_*}\boldsymbol{X}_{T_*}\right\}} \end{aligned} \tag{2.149}$$

which results in the lower bound of $S\sigma^2/k$ for training without the guard intervals and $S\sigma^2/p$ for training with the guard intervals. In other words, the MSE of an oracle based channel estimate is lower bounded by (# of nonzero entries of $\boldsymbol{\beta}$) $\cdot\, \sigma^2/$(# of *effective* observations). Comparison of this lower bound with that for the MSE of an LS channel estimate shows that linear channel estimates based on the LS criterion may be at a significant disadvantage when it comes to estimating sparse channels. Finally, notice that in the case of training without guard intervals (corresponding to the observation matrix given by (2.142)), the oracle estimator only requires that $k \geq S$ as opposed to $k \geq n$ for an LS channel estimate.

While the ideal channel estimate $\boldsymbol{\beta}^*$ is impossible to construct in practice, we now show that it is possible to obtain a more reliable estimate of $\boldsymbol{\beta}$ as a solution to the Dantzig selector (DS). This is accomplished by readily adapting the results from Lemma 2.16 in Section 2.2.1.1 and Theorems 2.27 and 2.28 in Section 2.2.3.1. Below, we summarize these adapted results in terms of two theorems.

**Theorem 2.20 (Training Without Guard Intervals)** *Let the training sequence* $\{x_i\}_{i=1}^p$ *be given by a sequence of i.i.d. binary random variables taking values* $+1$ *or* $-1$ *with probability* $1/2$ *each. Further, let*

*$p = n + k - 1$ for some $k \geq 4c_2 S^2 \log n$. Choose $\lambda_n = \sqrt{2(1+a)\log n}$ for any $a \geq 0$. Then the DS channel estimate*

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{z} \in \mathbb{R}^n}{\arg\min} \|\boldsymbol{z}\|_{\ell_1} \text{ subject to } \|\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{z})\|_{\ell_\infty} \leq \sigma \lambda_n \sqrt{k} \tag{2.150}$$

*satisfies*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq 2c_0'(1+a)\log n \cdot \left(\frac{S\sigma^2}{k}\right) \tag{2.151}$$

*with probability at least $1 - 2\max\left\{\left(\sqrt{\pi(1+a)\log n}\cdot n^a\right)^{-1}, \exp(-c_1 k/4S^2)\right\}$. Here, the observation matrix $\boldsymbol{X}$ corresponds to the partial Toeplitz matrix given in (2.142), $c_0'$ is as defined in Lemma 2.16, and $c_1$ and $c_2$ are positive constants that depend only on $S$ and are given in Theorem 2.27.*

**Theorem 2.21 (Training With Guard Intervals)** *Let the training sequence $\{x_i\}_{i=1}^p$ be given by a sequence of i.i.d. binary random variables taking values $+1$ or $-1$ with probability $1/2$ each. Further, let $p \geq 4c_2 S^2 \log n$ and choose $\lambda_n = \sqrt{2(1+a)\log n}$ for any $a \geq 0$. Then the DS channel estimate*

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{z} \in \mathbb{R}^n}{\arg\min} \|\boldsymbol{z}\|_{\ell_1} \text{ subject to } \|\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{z})\|_{\ell_\infty} \leq \sigma \lambda_n \sqrt{p} \tag{2.152}$$

*satisfies*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq 2c_0'(1+a)\log n \cdot \left(\frac{S\sigma^2}{p}\right) \tag{2.153}$$

*with probability at least $1 - 2\max\left\{\left(\sqrt{\pi(1+a)\log n}\cdot n^a\right)^{-1}, \exp(-c_1 p/4S^2)\right\}$. Here, the observation matrix $\boldsymbol{X}$ corresponds to the full Toeplitz matrix given in (2.141), $c_0'$ is as defined in Lemma 2.16, and $c_1$ and $c_2$ are positive constants that depend only on $S$ and are given in Theorem 2.28.*

The proofs of these theorems are essentially a direct application of Lemma 2.16 and Theorems 2.27 and 2.28, and are therefore omitted for the sake of brevity. These two theorems show that the DS channel estimate achieves squared error (roughly) within a factor of $\log n$ of the oracle based MSE lower bound of (# of nonzero entries of $\boldsymbol{\beta}$) $\cdot \sigma^2$/(# of *effective* observations).

The appeal of the DS channel estimator, however, goes beyond the estimation of truly sparse channels. Indeed, it is to be expected that physical channels in certain scattering environments happen to be only "approximately" sparse [36]. Specifically, rearrange (and reindex) the entries of the channel impulse response $\boldsymbol{\beta}$ by decreasing order of magnitude: $|\beta_{(1)}| \geq |\beta_{(2)}| \geq \cdots \geq |\beta_{(n)}|$. We term a wireless channel *approximately sparse* if the ordered entries $\{\beta_{(j)}\}$ of its impulse response decay with the index $j$. The following theorem, which focuses on the case of training with guard intervals and basically follows from Lemma 2.16 and Theorem 2.28, quantifies the reconstruction performance of the DS channel estimator in this setting.

**Theorem 2.22 (Estimating Approximately Sparse Channels)** *Let the training sequence $\{x_i\}_{i=1}^p$ be given by a sequence of i.i.d. binary random variables taking values $+1$ or $-1$ with probability $1/2$ each. Fix*

*any* $S \in \mathbb{N}$*, and choose* $p \geq 4c_2 S^2 \log n$ *and* $\lambda_n = \sqrt{2(1+a)\log n}$ *for any* $a \geq 0$*. Then the DS channel estimate* (2.152) *satisfies*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq c_0' \min_{1 \leq m \leq S} \left( \frac{\sigma \lambda_n \sqrt{m}}{\sqrt{p}} + \frac{\sum_{j=m+1}^{n} |\beta_{(j)}|}{\sqrt{m}} \right)^2 \tag{2.154}$$

*with probability at least* $1 - 2\max\left\{ \left( \sqrt{\pi(1+a)\log n} \cdot n^a \right)^{-1}, \exp(-c_1 p/4S^2) \right\}$*. Here, the observation matrix* $\boldsymbol{X}$ *corresponds to the full Toeplitz matrix given in* (2.141)*, and the constants* $c_0', c_1$ *and* $c_2$ *are as in Theorem 2.21.*

While the immediate significance of this result is obscured by the minimization over $m$ in (2.154), its implications can be better understood by focusing on a specific decay structure of the ordered entries of $\boldsymbol{\beta}$. One such decay structure, which is widely-studied in the literature [39], assumes that the $j$-th largest entry of $\boldsymbol{\beta}$ obeys

$$|\beta_{(j)}| \leq R \cdot j^{-\alpha - 1/2} \tag{2.155}$$

for some $R > 0$ and $\alpha > 1/2$. The parameter $\alpha$ here controls the rate of decay of the magnitudes of the ordered entries. Under this decay condition, the summation in (2.154) can be explicitly written as $\sum_{j=m+1}^{n} |\beta_{(j)}| \leq C_\alpha R\, m^{-\alpha+1/2}$, where $C_\alpha > 0$ is a constant that depends only on $\alpha$. We then have the following corollary of Theorem 2.22.

**Corollary 2.23** *Suppose that the channel impulse response* $\boldsymbol{\beta} \in \mathbb{R}^n$ *obeys* (2.155) *and let* $\{x_i = \pm 1\}_{i=1}^{p}$ *be the random binary sequence used to probe the channel for the case of training with guard intervals. Choose* $p \geq C_2 (\log n)^{\frac{2\alpha-3}{2\alpha-1}} (\sigma^2)^{-\frac{2}{2\alpha-1}}$ *and* $\lambda_n = \sqrt{2(1+a)\log n}$ *for any* $a \geq 0$*. Then the reconstruction error of the DS channel estimate* (2.152) *is upper bounded by*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_{\ell_2}^2 \leq C_0 (\log n)^{\frac{2\alpha}{2\alpha+1}} \cdot \left( \frac{\sigma^2}{p} \right)^{\frac{2\alpha}{2\alpha+1}} \tag{2.156}$$

*with probability at least* $1 - 2\max\left\{ \left( \sqrt{\pi(1+a)\log n} \cdot n^a \right)^{-1}, \exp\left( -C_1 (\log n \cdot \sigma^2)^{\frac{2}{2\alpha+1}} p^{\frac{2\alpha-1}{2\alpha+1}} \right) \right\}$*. Here, the absolute constants* $C_0(a, \alpha, c_0', R), C_1(a, \alpha, c_1, R)$*, and* $C_2(a, \alpha, c_2, R)$ *are strictly positive and depend only on the parameters* $a, \alpha, c_0', c_1, c_2$*, and* $R$*.*

It is instructive at this point to compare the reconstruction error performance of the DS channel estimate (given in (2.156)) with that of the LS channel estimate. Notice that since the MSE lower bound of $O(n\sigma^2/p)$ (given in (2.144)) holds for the LS channel estimate for all $\beta \in \mathbb{R}^n$, it remains valid under the decay condition (2.155). On the other hand, ignoring the $\log n$ factor in (2.156), we see that the reconstruction error of the

DS solution essentially behaves like $O\left((\sigma^2/p)^{\frac{2\alpha}{2\alpha+1}}\right)$. Thus, even in the case of an approximately sparse channel impulse response, the DS channel estimate shows an MSE improvement by a factor of (roughly) $O(n \cdot (\sigma^2/p)^{1/(2\alpha+1)})$ over the LS MSE of $n\sigma^2/p$. In fact, it can also be shown that $O\left((\sigma^2/p)^{\frac{2\alpha}{2\alpha+1}}\right)$ is the minimax MSE rate for the class of channels exhibiting the decay (2.155) and hence, the performance of the DS channel estimator comes within a $\log n$ factor of a minimax estimator. Finally, note that performance guarantees similar to the ones provided in Theorem 2.22 and Corollary 2.23 can also be obtained from Lemma 2.16 and Theorem 2.27 for the case of training without guard intervals.

### 2.2.3 Random Toeplitz Matrices Satisfy RIP

Because of the ubiquity of binary phase shift keying signaling in wireless communications, the channel estimation results in the previous section were stated in terms of random binary ($\pm 1$) probe sequences. However, the results also hold in settings where the probe sequence consists of realizations of random variables drawn from any bounded zero-mean distribution. In fact, the same results hold if the probe sequence is drawn from certain unbounded zero-mean distributions, such as the Gaussian distribution.

More generally, Toeplitz CS matrices have some additional benefits compared to completely independent (i.i.d.) random CS matrices. First, Toeplitz matrices are more efficient to generate and store. A $k \times n$ (random) partial Toeplitz matrix only requires the generation and storage of $k + n$ independent realizations of a random variable, while a fully-random matrix of the same size requires the generation and storage of $kn$ random quantities. In addition, the use of Toeplitz matrices in CS applications leads to a general reduction in computational complexity. Performing a matrix-vector multiplication between a fully-random $k \times n$ matrix and an $n \times 1$ vector requires $kn$ operations. In contrast, multiplication by a Toeplitz matrix can be performed in the frequency domain, because of the convolutional nature of Toeplitz matrices. Using fast Fourier transforms, the complexity of the multiplication can be reduced to $O(n \log n)$ operations, resulting in a significant speedup of the mixed-norm optimizations that are essential to several commonly-utilized CS reconstruction procedures such as GPSR [40] and SpaRSA [41]. Depending on the computational resources available, this speedup can literally be the difference between *intractable* and *solvable* problems.

In this section we establish the main claim of this paper, that Toeplitz matrices with entries drawn from either zero-mean bounded distributions or the zero-mean Gaussian distribution satisfy the restricted isometry property (RIP). Recall the definition of the Restricted Isometry Property from (2.12) of Section 2.2.1.1. The RIP statement is essentially a statement about singular values, and to establish RIP for a given matrix it suffices to bound the extremal eigenvalues of the Gram matrices of all column submatrices (having no more than $S$ columns) in the range $(1 - \delta_S, 1 + \delta_S)$. We will use this interpretation in our proofs, and the main results will be obtained using *Geršgorin's Disc Theorem*, which is an elegant result in classical eigenanalysis. We state this result here as a lemma, without proof. There are many valid references—see, for example, [42].

**Lemma 2.24 (Geršgorin)** *The eigenvalues of an $m \times m$ matrix $\boldsymbol{M}$ all lie in the union of $m$ discs $d_i = d_i(c_i, r_i)$, $i = 1, 2, \ldots, m$, centered at $c_i = M_{i,i}$, and with radius*

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^{m} |M_{i,j}|. \tag{2.157}$$

To begin, we consider any subset of column indices $T \subset \{1, \ldots, n\}$ of size $|T| \leq S$, and let $\boldsymbol{X}_T$ be the submatrix formed by retaining the columns of $\boldsymbol{X}$ indexed by the entries of $T$. The singular values of $\boldsymbol{X}_T$ are the eigenvalues of its $|T| \times |T|$ Gram matrix $\boldsymbol{G}(\boldsymbol{X}, T) = \boldsymbol{X}_T' \boldsymbol{X}_T$. Suppose that, for some integer $S \geq 1$ and some positive values $\delta_d$ and $\delta_o$ chosen such that $\delta_d + \delta_o = \delta_S \in (0, 1)$, every diagonal element of $\boldsymbol{G}(\boldsymbol{X}, T)$ satisfies $|G_{i,i}(\boldsymbol{X}, T) - 1| < \delta_d$ and every off-diagonal element $G_{i,j}(\boldsymbol{X}, T)$, $i \neq j$, satisfies $|G_{i,j}(\boldsymbol{X}, T)| < \delta_o / S$. Then the center of each Geršgorin disc associated with the matrix $\boldsymbol{G}(\boldsymbol{X}, T)$ deviates from 1 by no more than $\delta_d$ and the radius of each disc is no larger than $(S-1)\delta_o / S < \delta_o$. By Lemma 2.24, the eigenvalues of $G(\boldsymbol{X}, T)$ are all in the range $(1 - \delta_d - \delta_o, 1 + \delta_d + \delta_o) = (1 - \delta_S, 1 + \delta_S)$.

Now, notice that every Gram matrix $\boldsymbol{G}(\boldsymbol{X}, T)$ is a submatrix of the full Gram matrix $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{X}, \{1, \ldots, n\})$. Thus, instead of considering each submatrix separately, we can instead establish the above conditions on the elements of the full Gram matrix $\boldsymbol{G}$, and that suffices to ensure that the eigenvalues of *all* submatrices (formed by any choice of $T$, $|T| \leq S$) are controlled simultaneously. In the proofs that follow, we will show that every diagonal element of $\boldsymbol{G}$ is close to one (with high probability), and every off-diagonal element is bounded in magnitude (again, with high probability), and the final result will follow from a simple union bound.

It is instructive to note that because of the convolutional structure imposed by the linear, time-invariant observation model we consider here, the sufficient conditions to establish on the diagonal and off-diagonal elements of the Gram matrix of the resulting observation matrix essentially amount to properties of the autocorrelation function of the probe sequence. For the full observation matrix shown in (2.141), for example, each diagonal element is identical and equal to the autocorrelation of the probe sequence at lag zero. Similarly, each off-diagonal element corresponds to the autocorrelation at different nonzero lags (as stated in Section 2.2.2, the probe sequence is assumed to be zero outside of the specified range). For the partial observation matrix of (2.142), the diagonal and off-diagonal elements correspond to windowed versions of the autocorrelation function at different lags. In the following subsections we quantify these autocorrelations for certain random input sequences. However, we note that the proof technique described above can be used to establish RIP for *any* input sequence (including possibly deterministic sequences); one would only need to verify that the autocorrelation function of the sequence satisfies the required conditions.

#### 2.2.3.1 Bounded Entries

First we establish RIP for random Toeplitz matrices, for both the full observation matrices as shown in (2.141) as well as the partial matrices like (2.142), when the probe sequence $\{x_i\}$ consists of i.i.d. realizations of any bounded zero-mean random variable. We scale the distributions on $x_i$ appropriately so that columns of the observation matrices are unit-normed in expectation. Suitable distributions are

- $x_i \sim \text{unif}\left[ -\sqrt{3/\xi}, \sqrt{3/\xi} \right]$,
- $x_i \sim \begin{cases} 1/\sqrt{\xi} & \text{with prob. } 1/2 \\ -1/\sqrt{\xi} & \text{w.p. } 1/2 \end{cases}$,
- For $q \in (0,1)$, $x_i \sim \begin{cases} 1/\sqrt{\xi q} & \text{w.p. } q/2 \\ 0 & \text{w.p. } 1-q \\ -1/\sqrt{\xi q} & \text{w.p. } q/2 \end{cases}$,

where $\xi = k$ for partial matrices and $\xi = p$ for full matrices.

Before we state the first main results of the paper, we provide two lemmas that will be useful in the proofs. First, we describe the concentration of a sum of squares of bounded random variables.

**Lemma 2.25** *Let $x_i$, $i = 1, \ldots, k$ be a sequence of i.i.d., zero-mean bounded random variables such that $|x_i| \leq a$, and with variance $\mathbb{E}\left[x_i^2\right] = \sigma^2$. Then,*

$$\Pr\left( \left| \sum_{i=1}^{k} x_i^2 - k\sigma^2 \right| \geq t \right) \leq 2\exp\left( -\frac{2t^2}{ka^4} \right) \tag{2.158}$$

**Proof:** Recall Hoeffding's inequality, which states that a sequence of $k$ independent bounded random variables $z_i$ satisfying $a_i \leq z_i \leq b_i$ with probability one, satisfies

$$\Pr\left( |s_k - \mathbb{E}\left[s_k\right]| \geq t \right) \leq 2\exp\left( -\frac{2t^2}{\sum_{i=1}^{k}(b_i - a_i)^2} \right), \tag{2.159}$$

where $s_k = \sum_{i=1}^{k} z_i$. In our case, we let $z_i = x_i^2$, so $z_i \in [0, a^2]$ with probability one, and since $s_k = \sum_{i=1}^{k} x_i^2$, $\mathbb{E}\left[s_k\right] = k\sigma^2$. The result follows. ∎

Next, we describe how the inner product between vectors whose entries are bounded random variables concentrates about its mean.

**Lemma 2.26** *Let $x_i$ and $y_i$, $i = 1, \ldots, k$ be sequences of i.i.d., zero-mean, bounded random variables satisfying $|x_i| \leq a$ (and thus $|x_i y_i| \leq a^2$). Then,*

$$\Pr\left( \left| \sum_{i=1}^{k} x_i y_i \right| \geq t \right) \leq 2\exp\left( -\frac{t^2}{2ka^4} \right). \tag{2.160}$$

**Proof:** Again we apply Hoeffding's inequality to the sum $s_k = \sum_{i=1}^{k} z_i$, this time with $z_i = x_i y_i$. In this case we have $-a^2 \leq z_i \leq a^2$ and since the elements are independent and have zero mean, $\mathbb{E}[s_k] = 0$. The result follows. ∎

We are now in a position to state and prove the first main result of the paper.

**Theorem 2.27** *Let $\{x_i\}_{i=1}^{n+k-1}$ be a sequence whose entries are i.i.d. realizations of bounded zero-mean random variables with variance $\mathbb{E}\left[x_i^2\right] = 1/k$, satisfying $|x_i| \leq \sqrt{c/k}$ for some $c \geq 1$ (several such distributions are given above). Let*

$$\boldsymbol{X} = \begin{bmatrix} x_n & x_{n-1} & \dots & x_2 & x_1 \\ x_{n+1} & x_n & \dots & x_3 & x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n+k-1} & x_{n+k-2} & \dots & x_{k+1} & x_k \end{bmatrix}, \tag{2.161}$$

*be the $k \times n$ Toeplitz matrix generated by the sequence, and assume $n > 2$. Then, for any $\delta_S \in (0,1)$, there exist constants $c_1$ and $c_2$ depending only on $\delta_S$ and $c$, such that whenever $k \geq c_2 S^2 \log n$, $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability exceeding $1 - \exp(-c_1 k/S^2)$.*

**Proof:** Following the discussion of Geršgorin's Theorem, we need to establish conditions on the diagonal and off-diagonal elements of the Gram matrix $\boldsymbol{G} = \boldsymbol{X}'\boldsymbol{X}$. Applying Lemma 2.25 we see that each diagonal element $G_{i,i} = \sum_{j=1}^{k} x_j^2$ satisfies

$$\Pr\left(|G_{i,i} - 1| \geq \delta_d\right) \leq 2\exp\left(-\frac{2k\delta_d^2}{c^2}\right) \tag{2.162}$$

and by the union bound

$$\Pr\left(\bigcup_{i=1}^{n} \{|G_{i,i} - 1| \geq \delta_d\}\right) \leq 2n\exp\left(-\frac{2k\delta_d^2}{c^2}\right). \tag{2.163}$$

This establishes the required condition on the diagonal elements of the Gram matrix.

Next we treat the off-diagonal elements. Notice that entry $G_{i,j}$ is simply the inner product between columns $i$ and $j$ of the matrix $\boldsymbol{X}$. For example, one such term for the matrix specified in Theorem 2.27 is given by

$$G_{n-1,n} = x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_5 + \cdots + x_k x_{k+1}. \tag{2.164}$$

One issue is immediately apparent—the entries of the sum are not independent, so standard concentration inequalities cannot be applied directly. In the example here, the first two terms are dependent (they both depend on $x_2$), as are the second and third (both depend on $x_3$), and the third and fourth (both depend on $x_4$). But notice that the first and third terms are independent, as are the second and fourth, etc. Overall the sum may be split into two sums of i.i.d. random variables, where each component sum is formed simply

by grouping alternating terms. The number of terms in each sum is either the same (if $k$ is even) or differs by one if $k$ is odd.

In fact this decomposition into two sums over independent entries is possible for every $G_{i,j}$, and this observation is the key to tolerating the dependencies that arise from the structure in the sensing matrix. Note that the terms in any such sum are each dependent with *at most two* other terms in the sum. Each sum can be rearranged such that the dependent terms are "chained"—that is, the $\ell$-th (rearranged) term is dependent with (at most) the $(\ell-1)$-st term and the $(\ell+1)$-st terms. This rearranged sum has the same structure as the example above, and can be split in a similar fashion simply by grouping alternating terms.

Rewrite the sum $G_{i,j} = \sum_{i=1}^{k} z_i$, where the $z_i$'s are identically distributed zero-mean random variables that satisfy $-c/k \leq z_i \leq c/k$. When $k$ is even, the sum can be decomposed as

$$G_{i,j} = \sum_{i=1}^{t_1=k/2} z_{\pi_1(i)} + \sum_{i=1}^{t_2=k/2} z_{\pi_2(i)} \tag{2.165}$$

where $t_1$ and $t_2$ denote the number of terms in each sum, and $z_{\pi_1(i)}$ and $z_{\pi_2(i)}$ denote the rearranged and reindexed terms. The permutation operators $\pi_1$ and $\pi_2$ need not be known explicitly—it is enough to simply know such operators exist. When $k$ is odd, we can write

$$G_{i,j} = \sum_{i=1}^{t_1=(k-1)/2} z_{\pi_1(i)} + \sum_{i=1}^{t_2=(k+1)/2} z_{\pi_2(i)}. \tag{2.166}$$

Generically, we write $G_{i,j} = G^1_{i,j} + G^2_{i,j}$. Applying Lemma 2.26 with $a^2 = c/k$ to the component sums having $t_1$ and $t_2$ terms gives

$$\begin{aligned}
&\Pr\left(|G_{i,j}| \geq \frac{\delta_o}{S}\right) \\
&\quad\leq \Pr\left(\left\{|G^1_{i,j}| > \frac{\delta_o}{2S}\right\} \text{ or } \left\{|G^2_{i,j}| > \frac{\delta_o}{2S}\right\}\right) \\
&\quad\leq 2\max\left\{\Pr\left(|G^1_{i,j}| > \frac{\delta_o}{2S}\right), \Pr\left(|G^2_{i,j}| > \frac{\delta_o}{2S}\right)\right\} \\
&\quad\leq 2\max\left\{2\exp\left(-\frac{k^2\delta_o^2}{8t_1c^2S^2}\right), 2\exp\left(-\frac{k^2\delta_o^2}{8t_2c^2S^2}\right)\right\}.
\end{aligned} \tag{2.167}$$

It is easy to see that larger values of $t_1$ and $t_2$ decrease the error exponent, resulting in bounds that decay more slowly. For our purposes, to obtain a uniform bound independent of the parity of $k$, we use the (loose) upper bound $t_1 \leq t_2 < k$ to obtain

$$\Pr\left(|G_{i,j}| \geq \frac{\delta_o}{S}\right) \leq 4\exp\left(-\frac{k\delta_o^2}{8c^2S^2}\right). \tag{2.168}$$

To establish the condition for every off-diagonal element, we first note that, by symmetry, $G_{i,j} = G_{j,i}$. Thus, the total number of *unique* off-diagonal elements $G_{i,j}$ is $(n^2-n)/2 < n^2/2$, and we can apply the union of

events bound to obtain

$$\Pr\left(\bigcup_{i=1}^{n}\bigcup_{\substack{j=1\\ j\neq i}}^{n}\left\{|G_{i,j}|\geq\frac{\delta_o}{S}\right\}\right)\leq 2n^2\exp\left(-\frac{k\delta_o^2}{8c^2S^2}\right). \tag{2.169}$$

This establishes the required condition on the off-diagonal elements of the Gram matrix.

Now, recall that RIP of order $S$ holds with a prescribed $\delta_S\in(0,1)$ where $\delta_S=\delta_d+\delta_o$, when every diagonal element deviates from 1 by no more than $\delta_d$, and every off-diagonal element is less than $\delta_o/S$ in magnitude. To obtain the result claimed in Theorem 2.27, we assume $n\geq 3$, let $\delta_d=\delta_o=\delta_S/2$ and use the union bound to obtain

$$\begin{aligned}\Pr\left(\boldsymbol{X}\text{ does not satisfy }RIP(S,\delta_S)\right) &\leq 2n^2\exp\left(-\frac{k\delta_o^2}{8c^2S^2}\right)+2n\exp\left(-\frac{2k\delta_d^2}{c^2}\right)\\ &\leq 3n^2\exp\left(-\frac{k\delta_S^2}{32c^2S^2}\right).\end{aligned} \tag{2.170}$$

For $c_1<\delta_S^2/32c^2$, the upper bound

$$\Pr\left(\boldsymbol{X}\text{ does not satisfy }RIP(S,\delta_S)\right)\leq\exp\left(-\frac{c_1k}{S^2}\right) \tag{2.171}$$

holds whenever

$$k\geq\left(\frac{96c^2}{\delta_S^2-32c_1c^2}\right)S^2\log n, \tag{2.172}$$

which proves the theorem. ∎

The same technique can be applied to the full observation matrices as in (2.141). This leads to the second main result of the paper.

**Theorem 2.28** *Let $\{x_i\}_{i=1}^p$ be a sequence whose entries are i.i.d. realizations of bounded zero-mean random variables with variance $\mathbb{E}\left[x_i^2\right]=1/p$, satisfying $|x_i|\leq\sqrt{c/p}$ for some $c\geq 1$ (the example distributions listed at the start of the section again suffice). Let*

$$\boldsymbol{X}=\begin{bmatrix} x_1 & & 0\\ x_2 & \ddots & \\ \vdots & \ddots & x_1\\ x_p & & x_2\\ & \ddots & \vdots\\ 0 & & x_p\end{bmatrix} \tag{2.173}$$

*be the $(n+p-1)\times n$ full Toeplitz matrix generated by the sequence, and assume $n>2$. Then, for any $\delta_S\in(0,1)$ there exist constants $c_1$ and $c_2$ depending only on $\delta_S$ and $c$, such that for any sparsity level $S\leq c_2\sqrt{p/\log n}$, $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability exceeding $1-\exp(-c_1p/S^2)$.*

**Remark 2.29** *Notice the difference in the statements of results in Theorems 2.27 and 2.28, which highlight an inherent difference in the respective observation models. In the setting of Theorem 2.27, the user is allowed the flexibility to obtain more measurements "on the fly," and the resulting (rescaled) matrices satisfy RIP with higher orders $S$ (or smaller parameters $\delta_S$). Contrast that with the setting of Theorem 2.28, where the number of observations is fixed a priori. This effectively imposes an upper limit on the order $S$ (or a lower limit on the parameter $\delta_S$) for which RIP is satisfied.*

**Proof:** The proof proceeds in a similar fashion to the proof of Theorem 2.27. Each column of the "full" observation matrix now contains $p$ entries of the probe sequence, and is identical modulo an integer shift. From Lemma 2.25, the diagonal elements of the Gram matrix satisfy

$$\Pr\left(\bigcup_{i=1}^{n}\{|G_{i,i}-1|\geq\delta_d\}\right)\leq 2\exp\left(-\frac{2p\delta_d^2}{c^2}\right). \tag{2.174}$$

The off-diagonal elements are still composed of sums of dependent random variables, however, in this case the number of nonzero terms comprising each sum varies. At most (when $i$ and $j$ differ by 1), $G_{i,j}$ will consist of a sum of $p-1$ terms. On the other extreme, if $p\leq|j-i|$, each term of the inner product is zero trivially. In any event, we can still apply the results of Lemma 2.26 and upper-bound the error for each term by the worst-case behavior. This gives

$$\begin{aligned}
&\Pr\left(|G_{i,j}|\geq\frac{\delta_o}{S}\right)\\
&\quad\leq\quad\Pr\left(\left\{|G_{i,j}^1|>\frac{\delta_o}{2S}\right\}\text{ or }\left\{|G_{i,j}^2|>\frac{\delta_o}{2S}\right\}\right)\\
&\quad\leq\quad 2\max\left\{\Pr\left(|G_{i,j}^1|>\frac{\delta_o}{2S}\right),\Pr\left(|G_{i,j}^2|>\frac{\delta_o}{2S}\right)\right\}\\
&\quad\leq\quad 2\max\left\{2\exp\left(-\frac{p^2\delta_o^2}{8t_1c^2S^2}\right),2\exp\left(-\frac{p^2\delta_o^2}{8t_2c^2S^2}\right)\right\}.
\end{aligned}\tag{2.175}$$

Notice that now, regardless of the parity of $p$, the number of terms in each partial sum ($t_1$ and $t_2$) is no greater than $p/2$. The bound

$$\Pr\left(\bigcup_{i=1}^{n}\bigcup_{\substack{j=1\\j\neq i}}^{n}\left\{|G_{i,j}|\geq\frac{\delta_o}{S}\right\}\right)\leq 2n^2\exp\left(-\frac{p\delta_o^2}{4c^2S^2}\right). \tag{2.176}$$

follows. As before, we let $\delta_d=\delta_o=\delta_S/2$ and assume $n\geq 3$, to obtain

$$\Pr\left(\boldsymbol{X}\text{ does not satisfy }RIP(S,\delta_S)\right)\leq 3n^2\exp\left(-\frac{p\delta_S^2}{16c^2S^2}\right). \tag{2.177}$$

For any $c_1<\delta_S^2/16c^2$ and

$$S\leq\sqrt{\frac{\delta_S^2-16c_1c^2}{48c^2}}\cdot\sqrt{\frac{p}{\log n}}, \tag{2.178}$$

the matrix $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability at least $1-\exp(-c_1 p/S^2)$, proving the theorem. ∎

#### 2.2.3.2 Gaussian Entries

Results analogous to those of Theorems 2.27 and 2.28 can also be obtained if the entries of the probe sequence are drawn independently from certain unbounded distributions. For example, probe sequences consisting of i.i.d. Gaussian entries also generate Toeplitz matrices that satisfy RIP.

Following the proof techniques above, we first need to establish that the sum of squares of i.i.d. Gaussian random variables concentrates about its mean. For that, we utilize the following result from [43, Sec. 4, Lem. 1].

**Lemma 2.30** *Let $\{x_i\}_{i=1}^k$ be i.i.d. Gaussian variables with mean 0 and variance $\sigma^2$. The sum of squares of the $x_i$'s satisfies*

$$\Pr\left(\sum_{i=1}^{k} x_i^2 - k\sigma^2 \geq 2\sigma^2\sqrt{kt} + 2\sigma^2 t\right) \leq \exp(-t) \tag{2.179}$$

*and*

$$\Pr\left(\sum_{i=1}^{k} x_i^2 - k\sigma^2 \leq -2\sigma^2\sqrt{kt}\right) \leq \exp(-t). \tag{2.180}$$

*For $0 \leq t \leq 1$, the symmetric bound*

$$\Pr\left(\left|\sum_{i=1}^{k} x_i^2 - k\sigma^2\right| \geq 4\sigma^2\sqrt{kt}\right) \leq 2\exp(-t) \tag{2.181}$$

*follows.*

In addition, we can quantify the concentration of inner products between zero-mean Gaussian random vectors as follows.

**Lemma 2.31** *Let $x_i$ and $y_i$, $i = 1, \ldots, k$ be sequences of i.i.d., zero-mean Gaussian random variables with variance $\sigma^2$. Then,*

$$\Pr\left(\left|\sum_{i=1}^{k} x_i y_i\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{4\sigma^2(k\sigma^2 + t/2)}\right). \tag{2.182}$$

**Proof:** The proof basically follows the derivation of Bernstein's Inequality. Using the Chernoff bound, we obtain

$$\Pr\left(\sum_{i=1}^{k} x_i y_i \geq t\right) \leq \exp(-st) \prod_{i=1}^{k} \mathbb{E}\left[\exp\left(s x_i y_i\right)\right], \tag{2.183}$$

which holds for all $s \geq 0$ and all $t > 0$. Fix a term inside the product and expand the exponential in a Taylor Series, which gives

$$\begin{aligned} \mathbb{E}\left[\exp\left(sx_iy_i\right)\right] &= \mathbb{E}\left[1 + (sx_iy_i) + \frac{(sx_iy_i)^2}{2!} + \frac{(sx_iy_i)^3}{3!} + \frac{(sx_iy_i)^4}{4!} + \dots\right] \\ &\leq \mathbb{E}\left[1 + \frac{(sx_iy_i)^2}{2!} + \frac{|sx_iy_i|^3}{3!} + \frac{(sx_iy_i)^4}{4!} + \frac{|sx_iy_i|^5}{5!} + \dots\right] \end{aligned} \tag{2.184}$$

Now, since the $x_i$'s and $y_i$'s are Gaussian and independent, it is easy to verify that $\mathbb{E}\left[|x_iy_i|^p\right] = \mathbb{E}\left[|x_i|^p\right] \cdot \mathbb{E}\left[|y_i|^p\right] \leq p!\ \sigma^{2p}$ for $p \geq 2$, and so the expectation can be bounded by

$$\begin{aligned} \mathbb{E}\left[\exp\left(sx_iy_i\right)\right] &\leq 1 + s^2\sigma^4 + s^3\sigma^6 + s^4\sigma^8 + \dots \\ &= 1 + s^2\sigma^4 \sum_{j=0}^{\infty} (s\sigma^2)^j. \end{aligned} \tag{2.185}$$

Now, assume $s\sigma^2 = \nu < 1$ to obtain

$$\mathbb{E}\left[\exp\left(sx_iy_i\right)\right] \leq 1 + \frac{s^2\sigma^4}{1-\nu} \leq \exp\left(\frac{s^2\sigma^4}{1-\nu}\right). \tag{2.186}$$

Combining results, we have

$$\Pr\left(\sum_{i=1}^{k} x_iy_i \geq t\right) \leq \exp\left(-st + \frac{ks^2\sigma^4}{1-\nu}\right), \tag{2.187}$$

or equivalently,

$$\Pr\left(\sum_{i=1}^{k} x_iy_i \geq \frac{\gamma}{s} + \frac{ks\sigma^4}{1-\nu}\right) \leq \exp\left(-\gamma\right). \tag{2.188}$$

Now substitute $s = \nu/\sigma^2$, let $\alpha = k\nu\sigma^2/(1-\nu)$ and $\beta = \gamma\sigma^2/\nu$, and simplify to obtain

$$\Pr\left(Z \geq \alpha + \beta\right) \leq \exp\left(-\frac{\alpha\beta}{\sigma^2(k\sigma^2 + \alpha)}\right). \tag{2.189}$$

Letting $\alpha = \beta = t/2$, for $t < 2$, we obtain

$$\Pr\left(Z \geq t\right) \leq \exp\left(-\frac{t^2}{4\sigma^2(k\sigma^2 + t/2)}\right). \tag{2.190}$$

The other half of the bound can be obtained similarly using the fact that

$$\Pr\left(Z \leq -t\right) \leq \Pr\left(-sZ \geq st\right) \leq \exp\left(-st\right)\mathbb{E}\left[\exp\left(-sZ\right)\right], \tag{2.191}$$

and

$$\mathbb{E}\left[\exp\left(-sZ\right)\right] \leq \mathbb{E}\left[1 + \frac{(sx_iy_i)^2}{2!} + \frac{|sx_iy_i|^3}{3!} + \frac{(sx_iy_i)^4}{4!} + \frac{|sx_iy_i|^5}{5!} + \dots\right] \tag{2.192}$$

as above, making the bounds symmetric and identical. The result follows. ∎

Leveraging the above lemmas, we can establish the following.

**Theorem 2.32** *Let $\{x_i\}_{i=1}^{n+k-1}$ be a sequence whose entries are i.i.d. Gaussian random variables with mean zero and variance $\mathbb{E}\left[x_i^2\right] = 1/k$. Let*

$$\boldsymbol{X} = \begin{bmatrix} x_n & x_{n-1} & \dots & x_2 & x_1 \\ x_{n+1} & x_n & \dots & x_3 & x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n+k-1} & x_{n+k-2} & \dots & x_{k+1} & x_k \end{bmatrix}, \tag{2.193}$$

*be the $k \times n$ Toeplitz matrix generated by the sequence, and assume $n > 2$. Then, for any $\delta_S \in (0,1)$ there exist constants $c_1$ an $c_2$ depending only on $\delta_S$, such that whenever $k \geq c_2 S^2 \log n$, $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability exceeding $1 - \exp(-c_1 k/S^2)$.*

**Proof:** Following the proof method used in the previous subsection, we first use the symmetric bound in Lemma 2.30 to establish that

$$\Pr\left(\bigcup_{i=1}^{n} \{|G_{i,i} - 1| \geq \delta_d\}\right) \leq 2n \exp\left(-\frac{k\delta_d^2}{16}\right). \tag{2.194}$$

The off-diagonal elements exhibit the same dependencies treated in the proofs of Theorems 2.27 and 2.28. Again splitting each sum into two sums over independent entries, we leverage Lemma 2.31 to obtain

$$\Pr\left(|G_{i,j}| \geq \frac{\delta_o}{S}\right) \leq 2 \max\left\{2\exp\left(-\frac{k\delta_o^2}{4S^2(t_1/k + 1/2)}\right), 2\exp\left(-\frac{k\delta_o^2}{4S^2(t_2/k + 1/2)}\right)\right\} \tag{2.195}$$

for any $0 \leq \delta_d \leq 1$, where again $t_1$ and $t_2$ are the number of terms in each sum. Using the conservative upper bound $t_1 \leq t_2 \leq k$ we obtain

$$\Pr\left(\bigcup_{i=1}^{n}\bigcup_{\substack{j=1 \\ j\neq i}}^{n}\left\{|G_{i,j}| \geq \frac{\delta_o}{S}\right\}\right) \leq 2n^2 \exp\left(-\frac{k\delta_o^2}{6S^2}\right). \tag{2.196}$$

Now, let $\delta_d = 2\delta_S/3$ and $\delta_o = \delta_S/3$ and assume $n \geq 3$, to obtain

$$\Pr\left(\boldsymbol{X} \text{ does not satisfy } RIP(S, \delta_S)\right) \leq 3n^2 \exp\left(-\frac{k\delta_S^2}{54S^2}\right). \tag{2.197}$$

For any $c_1 < \delta_S^2/54$ and

$$k \geq \left(\frac{162}{\delta_S^2 - 54c_1}\right) S^2 \log n, \tag{2.198}$$

the matrix $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability at least $1 - \exp(-c_1 k/S^2)$, proving the theorem. ∎

For the full observation matrix, composed of entries from a Gaussian sequence, the following is true.

**Theorem 2.33** *Let $\{x_i\}_{i=1}^{p}$ be a sequence whose entries are i.i.d. realizations of zero-mean Gaussian random variables with variance $1/p$. Let*

$$\boldsymbol{X} = \begin{bmatrix} x_1 & & 0 \\ x_2 & \ddots & \\ \vdots & \ddots & x_1 \\ x_p & & x_2 \\ & \ddots & \vdots \\ 0 & & x_p \end{bmatrix} \tag{2.199}$$

*be the $(n+p-1) \times n$ full Toeplitz matrix generated by the sequence, and assume $n > 2$. Then, for any $\delta_S \in (0,1)$ there exist constants $c_1$ and $c_2$ depending only on $\delta_S$, such that for any sparsity level $S \leq c_2\sqrt{p/\log n}$ $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability exceeding $1 - \exp(-c_1 p/S^2)$.*

**Proof:** The proof is analogous to the proof of Theorem 2.28. The columns of $\boldsymbol{X}$ are identical (modulo an integer shift), so

$$\Pr\left(\bigcup_{i=1}^{n}\{|G_{i,i}-1| \geq \delta_d\}\right) \leq 2\exp\left(-\frac{p\delta_d^2}{16}\right). \tag{2.200}$$

and now,

$$\Pr\left(\bigcup_{i=1}^{n}\bigcup_{\substack{j=1\\ j\neq i}}^{n}\left\{|G_{i,j}| \geq \frac{\delta_o}{S}\right\}\right) \leq 2n^2\exp\left(-\frac{p\delta_o^2}{4S^2}\right). \tag{2.201}$$

Letting $\delta_d = 2\delta_S/3$ and $\delta_o = \delta_S/3$ and assuming $n \geq 3$, we have that for any $c_1 < \delta_S^2/36$ and

$$S \leq \sqrt{\frac{\delta_S^2 - 36c_1}{108}} \cdot \sqrt{\frac{p}{\log n}}, \tag{2.202}$$

the matrix $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability at least $1-\exp(-c_1 p/S^2)$, proving the theorem. ∎

## 2.2.4 Discussion

### 2.2.4.1 Generalizations and Dependency Tolerance using Graph Coloring

It is easy to see that the results of Theorems 2.27, 2.28, 2.32, and 2.33 also apply directly to Hankel matrices, which are Toeplitz-like matrices whose entries are identical along anti-diagonals. In addition, the proof techniques utilized to obtain the results of Theorems 2.27 and 2.32 also can be used to establish RIP

for (left- or right-shifted) partial circulant matrices of the form

$$\boldsymbol{X} = \begin{bmatrix} x_n & x_{n-1} & \dots & \dots & \dots & x_3 & x_2 & x_1 \\ x_1 & x_n & \dots & \dots & \dots & x_4 & x_3 & x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{k-1} & x_{k-2} & \dots & x_1 & x_n & \dots & x_{k+1} & x_k \end{bmatrix}, \tag{2.203}$$

generated by a random sequence of length $n$.

The techniques developed here can also be applied in more general settings where the observation matrices exhibit structured statistical dependencies. Recall that, in the above proofs, dependencies were tolerated by partitioning sums of dependent random variables into two component sums of fully independent random variables. The actual partitioning was not performed directly, rather the only facts required in the proof were that such partitions exist, and that the number of terms in each component sum was specified. If, for a given observation matrix, similar partitioning can be established, analogous results will follow.

We generalize the approach utilized in this paper using techniques from graph theory. See, for example, [44] for basic reference. Let

$$\Sigma = \sum_{i=1}^{k} x_i \tag{2.204}$$

be a sum of identically distributed random variables. We associate the sum $\Sigma$ with an undirected graph $g(\Sigma) = (V, E)$ of degree $\Delta_g$, by associating a vertex $i \in V = \{1, 2, \dots, k\}$ to each term $x_i$ in the sum and creating an edge set $E$ such that an edge $e = (i, j)$ between vertices is in the set if and only if $x_i$ and $x_j$ are statistically dependent. The degree of the graph $\Delta_g$ is defined to be the maximum number of edges originating from any of the vertices. Notice that any fully-disconnected subgraph of $g(\Sigma)$, by definition, represents a collection of i.i.d. random variables.

The goal, then, is to partition $g(\Sigma)$ into some number of fully-disconnected subgraphs. In graph theory terminology, any such partitioning—essentially a labeling of each vertex such that vertices sharing an edge are labeled differently—is called a (proper) coloring of the graph. Given a coloring of $g(\Sigma)$, the concentration behavior of each partial sum associated with each subgraph can be obtained in a straightforward manner by standard concentration inequalities, and the contribution of several such subgraphs can be quantified using the union bound. Note, however, that trivial partitions exist (let each subgraph contain only one vertex, for example), leading to particularly poor concentration bounds. We seek to partition $g(\Sigma)$ into as few fully-disconnected subgraphs as possible while ensuring that each subgraph contains as many vertices as possible.

To achieve this, we consider *equitable coloring* of $g(\Sigma)$. An equitable coloring is a proper graph coloring where the difference in size between the smallest and largest collections of vertices sharing the same color is at most one. Proving a conjecture of Paul Erdös, Hajnal and Szemerédi showed that equitable colorings of

a graph with degree $\Delta$ exist for any number of colors greater or equal to $(\Delta+1)$ [45]. Along with the above argument, this shows that the concentration behavior of any sum $\Sigma$ exhibiting limited statistical dependence, as defined by the degree $\Delta_g$ of the associated dependency graph $g(\Sigma)$, can be controlled using equitable graph coloring. This procedure was also used to extend Hoeffding's inequality to such graph-dependent random variables in [46].

Utilizing this framework, we can obtain results that apply to observation matrices with more general dependency structures. The following result is representative.

**Theorem 2.34** *Let $\boldsymbol{X}$ be a $k \times n$ matrix whose entries are identically distributed realizations of bounded zero-mean random variables with variance $\mathbb{E}\left[x_i^2\right] = 1/k$, satisfying $x_i^2 \leq c/k$ for some $c \geq 1$. Assume that the dependency degree among elements in any column of $\boldsymbol{X}$ is no greater than some integer $\Delta_d \geq 0$, and each inner product between columns exhibits dependency degree no greater than some integer $\Delta_o \geq 0$. Then, for any $\delta_S \in (0,1)$, there exist constants $c_1$ and $c_2$ depending on $\delta_S$, the dependency degrees $\Delta_d$ and $\Delta_o$, and $c$, such that whenever $k \geq c_2 S^2 \log n$, $\boldsymbol{X}$ satisfies RIP of order $S$ with parameter $\delta_S$ with probability exceeding $1 - \exp(-c_1 k/S^2)$.*

**Proof:** First consider the diagonal elements of the Gram matrix of $\boldsymbol{X}$, each of which satisfies

$$\Pr\left(|G_{i,i} - 1| \geq \delta_d\right) \leq 2(\Delta_d + 1)\exp\left(-\frac{2k\delta_d^2}{c^2}\left\lfloor\frac{k}{\Delta_d+1}\right\rfloor\right) \tag{2.205}$$

and by the union bound

$$\Pr\left(\bigcup_{i=1}^{n}\left\{|G_{i,i} - 1| \geq \delta_d\right\}\right) \leq 2n(\Delta_d + 1)\exp\left(-\frac{2\delta_d^2}{c^2}\left\lfloor\frac{k}{\Delta_d+1}\right\rfloor\right), \tag{2.206}$$

where $\lfloor\cdot\rfloor$ is the floor function, which returns the largest integer less than the argument. Similarly, the off-diagonal elements satisfy

$$\Pr\left(\bigcup_{i=1}^{n}\bigcup_{\substack{j=1\\ j\neq i}}^{n}\left\{|G_{i,j}| \geq \frac{\delta_o}{S}\right\}\right) \leq 2n^2(\Delta_o + 1)\exp\left(-\frac{\delta_o^2}{8c^2S^2}\left\lfloor\frac{k}{\Delta_o+1}\right\rfloor\right). \tag{2.207}$$

The result follows from suitable bounding of the overall error probability. ∎

#### 2.2.4.2 Connections with Other Works

To the best of our knowledge, this paper is the first work to establish the restricted isometry property for random Toeplitz matrices with bounded or Gaussian entries. Here we briefly describe connections between this paper and several related existing works.

In the compressed sensing literature, the first work to propose Toeplitz-structured observation matrices was [47], where observations were obtained by convolving the incoming unknown signal with a random

filter—a filter whose taps were generated as realizations of certain random variables—followed by periodic downsampling of the output stream. While this approach was shown to be effective in practice, no theoretical guarantees were given. Without downsampling this initial approach is identical to the full observation model analyzed here, and in fact the techniques presented here could be utilized to establish conditions under which RIP would be satisfied for certain downsampled random filtering systems.

The first theoretical results for using random Toeplitz matrices in compresed sensing were established in [33]. Using an equitable graph coloring approach applied to the RIP proof of [32], we showed that $k \times n$ partial Toeplitz, Hankel, and left- or right-shifted circulant random matrices satisfy RIP of order $S$ with high probability, provided $k = O\left(S^3 \log n\right)$. This sufficient condition is more restrictive than what we establish here, where we reduce the exponent on $S$ by one order of magnitude.

In [48], techniques in sparse representation are applied to recover matrices that can be expressed as the superposition of a small number of component matrices—the so-called sparse matrix identification problem. When the component matrices are time-frequency shift matrices, the matrix identification problem becomes similar to the convolutional sparse signal recovery problem considered here. However, this work differs from our own in several significant ways. First, the work in [48] considers noise-free observations, and requires circular convolution with certain deterministic (Alltop) probe sequences. Second, and perhaps most notably, the theoretical analysis in [48] focuses on the coherence property of the dictionary (maximum absolute inner product between any two distinct columns), instead of RIP. Consequently, while we can establish that, for a given probe sequence, *any* sparse signal can be recovered with high probability from a collection of convolutional observations, the results in [48] only guarantee recovery of *most* signals defined on a randomly chosen signal support.

The problem of matrix identification was also studied in [49]. While this work did consider noisy observations, the recovery procedure proposed was the "bounded-noise" optimization (2.137). As explained in Section 2.2.1.1 the error bounds resulting from this optimization can be sufficiently weaker than the bounds that can be obtained using the Dantzig selector, given in (2.133). In fact, in addition to being restricted to circularly convolutional observations, and theoretical guarantees only on coherence (instead of RIP), the work in [49] provides no theoretical analysis to quantify the MSE of reconstructions obtained from noisy observations.

Our own previous work [31] was the first to use Geršgorin's Theorem to establish RIP for Toeplitz random matrices, achieving the less restrictive sufficient condition on the number of observations required, $k = O\left(S^2 \log n\right)$. While that work only treated matrices whose entries were drawn from a symmetric Bernoulli distribution, here we extend the results to random matrices whose entries are bounded or Gaussian-distributed.

In addition to the directly-related contributions described above, we briefly mention several tangentially-related contributions. While this paper was in preparation we became aware of the work of [50], which also examines random convolution in compressed sensing. There are several significant differences between this work and our own. First, in [50], the probe sequence (while randomly generated) is assumed to have orthogonal shifts, while we can tolerate probe sequences that simply consist of independent realizations of certain random variables. Second, our results establish RIP for the observation matrices resulting from convolution with a random sequence, from which it follows that *any* sparse (or nearly sparse) signal can be recovered with high probability from the same set of observations. In contrast, for a given random probe, the results in [50] only guarantee recoverability of *most* sparse signals defined on a fixed support, albeit with a potentially less restrictive condition on the number of observations required, $k = O\left(\max\left\{S \log n, \log^3 n\right\}\right)$ compared to our requirement of $k = O(S^2 \log n)$. Perhaps the most significant difference, however, is that our observation model prescribes collecting a consecutive set of samples of the linear convolution between the unknown signal and a random probe sequence, while the observation model in [50] requires circular convolution of the unknown sparse signal with a random probe sequence, followed by either random subsampling or randomized "block averaging" of the full set of observations. Thus, while our observation model occurs naturally in the context of linear system identification, the methods proposed in [50] do not.

Finally, in [51] it was established that, subject to a similar condition as what we obtain here—namely that the number of rows of the matrix must be on the order of the square of the sparsity level of the target signal—certain deterministic matrices satisfy RIP. Among these was a special type of block-circulant matrix generated by a collection of $\ell > 1$ columns, where the elements of the matrix satisfy $X_{i+1,j+\ell} = X_{i,j}$, and the arithmetic on the indices is done modulo the signal length $n$. In contrast, the generalization of our Toeplitz results apply to "true" circulant matrices that are generated by a single (random) column.

#### 2.2.4.3 Eigenvalues by Geršgorin's Theorem

The theory of sparse representation was an active area of research even before the advent of compressed sensing. The techniques that were developed in early works relied on the notion of coherence of a matrix, which is quantified by the largest (in magnitude) inner product between distinct columns of the matrix. The interesting point to note is that the notion of coherence can be parlayed into statements about RIP, the connection coming by way of Geršgorin's Theorem. Reminiscent constructs can be found, for example, in [52]. In addition, Geršgorin-like techniques arise in the proof of RIP for the deterministic constructions of [51], and are mentioned in [53] in the context of determining the eigenvalues of randomly chosen submatrices of a given dictionary matrix.

Using Geršgorin's Theorem to establish eigenvalues for general dictionaries is not without its limitations. For example, as noted in [53], the work of [54] shows that the minimum coherence between columns of any

(generally overcomplete) finite Grassmanian frame cannot be too small. For large $k$ and $n$, the coherence scales like $\sqrt{1/k}$, which would essentially imply a $k = O(S^2)$ requirement on the number of observations, similar to what we obtain in our proofs. Applying Geršgorin's theorem to fully-independent random matrices leads to similar restrictions. For example, a simple application of Lemma 2.26 (analogous to the approach in the proof of Theorem 2.27, but without the dependency tolerance steps) shows that Geršgorin's Theorem leads to the requirement that $O(S^2 \log n)$ rows are needed in order for a fully random observation matrix to satisfy RIP of order $S$ with some fixed success probability. On the other hand, we know from [1–3, 14, 32] that $k = O(S \log n)$ requirements suffice to establish RIP with the same probability of success.

Thus, while it is tempting to claim that the presence of dependencies in the Toeplitz-structured matrices amounts to an increase in the number of observations required for RIP to be satisfied, such a claim does not follow from the work presented here. Indeed, it is fully possible that the random matrices considered in this work do satisfy RIP when $k = O(S \log n)$, but the proof techniques utilized here are insufficient to establish that stronger result. The takeaway message here is that Geršgorin's Theorem provides a straightforward, but possibly suboptimal, approach to establishing RIP for general observation matrices.

### 2.2.5 Acknowledgment

The authors wish to thank Phil Schniter for pointing out a few minor errors in the initial version of the dependency tolerance arguments.

## 2.3 On the Restricted Isometry of Deterministically Subsampled Fourier Matrices

Matrices satisfying the Restricted Isometry Property (RIP) are central to the emerging theory of compressive sensing (CS). Initial results in CS established that the recovery of sparse vectors $\mathbf{x}$ from a relatively small number of linear observations of the form $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be achieved, using a tractable convex optimization, whenever $\mathbf{A}$ is a matrix that satisfies the RIP; similar results also hold when $\mathbf{x}$ is nearly sparse or the observations are corrupted by noise. In contrast to random constructions prevalent in many prior works in CS, this paper establishes a collection of deterministic matrices, formed by deterministic selection of rows of Fourier matrices, which satisfy the RIP. Implications of this result for the recovery of signals having sparse spectral content over a large bandwidth are discussed.

### 2.3.1 Introduction

The emerging theory of compressive sensing (CS) establishes that sparse vectors $\mathbf{x} \in \mathbb{R}^p$ can be recovered exactly from a relatively small number of special, non-adaptive linear measurements of the form $\mathbf{y} = \mathbf{A}\mathbf{x}$, where the number of rows of the measurement matrix $\mathbf{A}$, denoted by $m$, can be far fewer than ambient signal dimension $p$. Further, and perhaps most remarkably, this recovery can be accomplished by solving a tractable convex optimization [1–3]. A particularly concise way of identifying for which measurement matrices this exact recovery is possible is through the use of the Restricted Isometry Property (RIP) [14].

**Definition 2.35 (Restricted Isometry Property)** *The matrix $\mathbf{A}$ satisfies the Restricted Isometry Property of order $s$ with parameter $\delta_s \in [0, 1)$ if*

$$(1 - \delta_s)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta_s)\|\mathbf{x}\|_2^2 \tag{2.208}$$

*holds simultaneously for all sparse vectors $\mathbf{x}$ having no more than $s$ nonzero entries.*

**Remark 2.36** *In the above, the function $\|\mathbf{z}\|_2^2$ denotes the $\ell_2$ norm of the vector $\mathbf{z} \in \mathbb{R}^n$, given by*

$$\|\mathbf{z}\|_2^2 = \sum_{i=1}^{n} z_i^2. \tag{2.209}$$

*In the sequel, we will also make use of the $\ell_1$ norm, which is given by*

$$\|\mathbf{z}\|_1 = \sum_{i=1}^{n} |z_i|, \tag{2.210}$$

*and we will use the $\ell_0$ quasi-norm, $\|\mathbf{z}\|_0$, to denote the number of nonzero entries of the vector $\mathbf{z}$.*

For matrices satisfying the RIP, the following recovery result, established in [14] and refined in [30], is representative.

**Lemma 2.37 (Exact Recovery Using RIP Matrices)** *Let* $\mathbf{A}$ *be a matrix satisfying RIP of order* $2s$ *with parameter* $\delta_{2s} < \sqrt{2}-1$*, and let* $\mathbf{y} = \mathbf{A}\mathbf{x}$ *be a vector of observations of any sparse signal* $\mathbf{x} \in \mathbb{R}^p$ *having no more than* $s$ *nonzero entries. Then, the estimate*

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{z}\in\mathbb{R}^p} \|\mathbf{z}\|_1 \quad \textit{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z}, \tag{2.211}$$

*is unique and equal to* $\mathbf{x}$*.*

A variety of additional CS recovery results that leverage the RIP, including those that guarantee approximate recovery of nearly-sparse signals and stable recovery when the observations are corrupted by noise, can be found in the CS literature. See, for example, [16, 30, 55].

While there is currently no known polynomial-time algorithm to test whether a given matrix satisfies the RIP, certain randomly-constructed matrices have been shown to satisfy the RIP with high probability. For example, $m \times p$ random matrices whose entries are independent and identically distributed (iid) realizations of certain zero-mean random variables have been shown to satisfy RIP of order $s$ with parameter $\delta_s$ with high probability for any integer $s$ satisfying

$$s \leq c(\delta_s) \cdot \frac{m}{\log p}, \tag{2.212}$$

where $c(\delta_s)$ is a constant that depends only on $\delta_s$ but not on $m$ or $p$ [2, 3, 14, 32]. In such cases, the results of Lemma 2.37 hold with high probability.

Random matrices with more structure have also been shown to satisfy the RIP. Define the $p \times p$ Discrete Fourier Transform (DFT) matrix, denoted here as $\boldsymbol{\mathcal{F}}^p$, to be the complex-valued matrix whose $(j,k)$th entry is given by

$$\{\mathcal{F}^p\}_{j,k} = \exp\left(\frac{2\pi i j k}{p}\right), \tag{2.213}$$

where $i = \sqrt{-1}$ is the complex element and $j, k = 0, 1, \ldots, p-1$. We denote by $\boldsymbol{\mathcal{F}}^p_T$ the $|T| \times p$ submatrix formed by selecting only the rows of $\boldsymbol{\mathcal{F}}^p$ indexed by elements in a set $T$ whose entries come from the set $\{0, 1, \ldots p-1\}$. Recent results in [56] improve upon initial results in [1], establishing that when $T \subset \{0, 1, \ldots, p-1\}$ is a subset of size $|T| = m$, selected uniformly at random without replacement, the (scaled) submatrix $m^{-1/2}\boldsymbol{\mathcal{F}}^p_T$ satisfies the RIP of order $s$ with parameter $\delta_s$ with high probability when

$$s \leq c'(\delta_s) \cdot \frac{m}{\log^4 p}, \tag{2.214}$$

where $c'(\delta_s)$ is a constant that does not depend on $m$ or $p$.

In this paper we identify a collection of matrices formed by *deterministic* selection of rows of the Fourier matrix that satisfy the RIP. Our results leverage ideas from number theory, specifically the analysis of certain structured exponential sums, and a result from classical eigenanalysis known as Geršgorin's Disc Theorem.

Our main results are stated in the next section. Applications of the result, in the context of recovering time-varying signals whose spectral representations are sparse over a large bandwidth, are discussed in Section 2.3.3. Proofs are relegated to the appendix.

### 2.3.2 Main Results

Our main results concern submatrices formed by determinstically selecting a set of rows of $\mathcal{F}^p$. When the indices of the rows selected correspond to the (integer) outputs of certain polynomial functions mod $p$, the resulting submatrices satisfy the restricted isometry property (RIP). Analogous results also hold for submatrices formed by selecting rows of the Fourier inverse matrix $(\boldsymbol{\mathcal{F}}^p)^{-1}$, which is the $p \times p$ matrix whose $(j,k)$th element is given by

$$\left\{(\mathcal{F}^p)^{-1}\right\}_{j,k} = \frac{1}{p}\ \exp\left(\frac{-2\pi ijk}{p}\right), \tag{2.215}$$

$j,k = 0,1,\ldots,p-1$. The first main result is stated below as a theorem.

**Theorem 2.38** *Let $p > 2$ be a prime integer and let $f(n)$ be any polynomial of degree $d \geq 2$ of the form $f(n) = a_1 n + \cdots + a_d n^d$, with real integer coefficients $a_j \in \{0,1,\ldots,p-1\}$ for $j = 1,2,\ldots,d-1$, and $a_d \in \{1,2,\ldots,p-1\}$. For any $\epsilon_1 \in (0,1)$, choose $m$ to be an integer satisfying*

$$p^{1/(d-\epsilon_1)} \leq m \leq p. \tag{2.216}$$

*Let $T = \{f(n) \mod p : n = 1,2,\ldots,m\}$ (note that $|T| = m$ and $T$ may contain duplicate entries). Then, for any $\delta_s \in (0,1)$ and $\epsilon_2 \in (0,\epsilon_1)$, the matrix $m^{-1/2}\boldsymbol{\mathcal{F}}_T^p$ satisfies the RIP of order $s$ with parameter $\delta_s$ whenever*

$$s \leq \delta_s \cdot C(d,\epsilon_2) \cdot m^{(\epsilon_1-\epsilon_2)/2^{d-1}}, \tag{2.217}$$

*where $C(d,\epsilon_2)$ is a constant that does not depend on $m$ or $p$. The same result also holds for the subsampled Fourier inverse matrix, $pm^{-1/2}\,(\boldsymbol{\mathcal{F}}^p)_T^{-1}$, for the same choice of $T$.*

When the degree of the polynomial is large, the following result provides a slightly stronger statement compared with the above.

**Theorem 2.39** *Let $p > 2$ be a prime integer and let $f(n)$ be any polynomial of degree $d > 2$ of the form $f(n) = a_1 n + \cdots + a_d n^d$, with real integer coefficients $a_j \in \{0,1,\ldots,p-1\}$ for $j = 1,2,\ldots,d-1$, and $a_d \in \{1,2,\ldots,p-1\}$. Choose $m$ to be an integer satisfying*

$$p^{1/(d-1)} \leq m \leq p. \tag{2.218}$$

*Let $T = \{f(n) \mod p : n = 1,2,\ldots,m\}$ (note that $|T| = m$ and $T$ may contain duplicate entries). Then, for any $\delta_s \in (0,1)$, the matrix $m^{-1/2}\boldsymbol{\mathcal{F}}_T^p$ satisfies the RIP of order $s$ with parameter $\delta_s$ whenever*

$$s \leq \delta_s \cdot C'(d) \cdot m^{1/(9d^2 \log d)}, \tag{2.219}$$

*where $C'(d)$ is a constant that does not depend on $m$ or $p$. The same result also holds for the subsampled Fourier inverse matrix, $pm^{-1/2}\,(\boldsymbol{\mathcal{F}}^p)_T^{-1}$, for the same choice of $T$.*

We briefly note the comparison between these results and the existing results for similar random constructions. As stated in the introduction, matrices $m^{-1/2}\boldsymbol{\mathcal{F}}_T^p$ formed by selecting a set of $m$ rows of the $p$-dimensional Fourier matrix uniformly at random satisfy the RIP of order $s$ with parameter $\delta_s$ with high probability provided

$$s \leq c'(\delta_s) \cdot \frac{m}{\log^4 p}. \tag{2.220}$$

In contrast, here we identify an analogous class of *deterministic* matrices that satisfy RIP of order $s$ with parameter $\delta_s$ for $s \leq C(d,\epsilon_2) \cdot m^{(\epsilon_1-\epsilon_2)/2^{d-1}}$ for $d = 2$ and

$$s \leq \max\left\{\delta_s\ C(d,\epsilon_2) \cdot m^{(\epsilon_1-\epsilon_2)/2^{d-1}}\ ,\ \delta_s\ C'(d) \cdot m^{1/(9d^2\log d)}\right\}, \tag{2.221}$$

for $d > 2$, where $0 < \epsilon_2 < \epsilon_1 < 1$, $d$ is the degree of the polynomial used to generate the set of rows to be selected, and the constants $C$ and $C'$ do not depend on the signal dimension $p$ or the number of measurements $m$.

Relative to the results for random constructions, the deterministic results presented here exhibit a less-favorable scaling relationship with respect to the parameter $m$. Specifically, while a randomly constructed matrix with $m$ rows allows for recovery of signals with at most $s$ nonzero entries as given in (2.220), our deterministic results only guarantee recovery if the number of nonzero entries is $o(\sqrt{m})$. An open question remains as to whether there exists a (deterministic) sampling strategy that can bridge the gap. In addition, the results we present here also require that the number of measurements $m$ exceeds some fractional power of the signal dimension $p$. This minimum sampling requirement is much more restrictive than what exists in the CS literature for random sampling methods, where the dependence of the minimum number of samples on the signal dimension is typically some power of $\log p$. The minimum sampling conditions obtained here are artifacts of the analysis techniques, and another open question remains as to the minimum number of samples required for deterministic sampling strategies.

### 2.3.3 Applications

In this section we describe how the results established above can be utilized to analyze radio frequency (RF) receiver systems that acquire time-varying signals having sparse frequency-domain representations using structured, deterministic, non-uniform sampling. The approach we consider here entails collecting samples at times that are related to the outputs of certain polynomial functions, and is motivated (in part) by the **Ny**quist **F**olding **R**eceiver (NyFR) architecture proposed in [57]. Upon casting the sampling processes in the canonical CS framework, where the target signal is the (sparse) frequency-domain representation of the time-varying signal of interest, the results of the previous section guarantee that the "effective" observation matrix

satisfies the restricted isometry property (RIP). Consequently, any of a number of CS recovery algorithms can then be utilized to recover the target signal.

Consider signals $x(t)$ defined on $t \in (0, t_d]$, for some time duration $t_d > 0$. Let $p$ be a large prime, and for $\tau = t_d/p$, define the *sampled signal vector* $\boldsymbol{x} \in \mathbb{R}^p$ to be the vector whose $k$th entry is given by $x_k = x(k\tau)$, for $k = 1, 2, \ldots, p$. The class of signals we will be interested in here are those having sparse frequency-domain representations—specifically, signals for which $\|\boldsymbol{\mathcal{F}}^n \boldsymbol{x}\|_0 \leq s \ll p$.

Motivated by the results in the previous section, we suggest two methods for acquiring samples of the signals of interest according to the outputs of certain polynomial functions. The first method corresponds to a "streaming" sampling process, where the output of the polynomial function determines a set of actual sampling times. The second method is a process where the sample times are determined a priori, and can be collected over a relatively small time duration. Each of the methods is briefly described below.

First, in the case where the polynomial function determines actual sampling times, we let $f(n)$ be a degree-$d$ polynomial of the form $f(n) = a_1 n + \cdots + a_d n^d$, where $a_i \in \{0, 1, \ldots, p-1\}$ for $i = 1, \ldots, d-1$, and $a_d \in \{1, 2, \ldots, p-1\}$. A total of $m$ samples of $x(t)$ are collected, at sampling times described by the entries of the set $\{\tau \cdot f(n) : n = 1, 2, \ldots, m\}$. Notice that this acquisition scheme can be modeled by a matrix-vector product corresponding to the action of a *sampling matrix* $\boldsymbol{S}$ on the sampled signal vector, where each row of the sampling matrix contains only one nonzero (unit) entry corresponding to the time at which the sample is collected. More generally, since $x(t)$ (and its corresponding sampled signal vector) is periodic, this model generalizes naturally to describe sampling of the *periodic extension* of $x(t)$. Specifically, for some integer $r \geq 1$, the acquisition process can be modeled by the operation of an $m \times rp$ sampling matrix (having a single unit entry in each row corresponding to the sampling time) on the $rp \times 1$ vector constructed by stacking $r$ identical copies of the vector of samples $x$. We can express the collection of samples $\boldsymbol{y} \in \mathbb{R}^m$ as

$$\overbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}^{\boldsymbol{y}} = \overbrace{\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}}^{\boldsymbol{S}} \begin{bmatrix} x_1 \\ \vdots \\ x_p \\ \vdots \\ x_1 \\ \vdots \\ x_p \end{bmatrix}. \tag{2.222}$$

$$\underbrace{\phantom{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}}_{m \times 1} \qquad \underbrace{\phantom{\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}}}_{m \times rp} \qquad \underbrace{\phantom{\begin{bmatrix} x_1 \\ \vdots \\ x_p \\ \vdots \\ x_1 \\ \vdots \\ x_p \end{bmatrix}}}_{rp \times 1}$$

Now, to simplify this expression, we note that

$$\underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_p \\ \vdots \\ x_1 \\ \vdots \\ x_p \end{bmatrix}}_{rp\times 1} = \frac{1}{r} \cdot \underbrace{\begin{bmatrix} (\boldsymbol{\mathcal{F}}^p)^{-1} \\ \vdots \\ (\boldsymbol{\mathcal{F}}^p)^{-1} \end{bmatrix}}_{rp\times p} \underbrace{\begin{bmatrix} \boldsymbol{\mathcal{F}}^p & \dots & \boldsymbol{\mathcal{F}}^p \end{bmatrix}}_{p\times rp} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_p \\ \vdots \\ x_1 \\ \vdots \\ x_p \end{bmatrix}}_{rp\times 1}. \tag{2.223}$$

Substituting this into the above, we can rewrite the observation process as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \underbrace{\boldsymbol{S} \begin{bmatrix} (\boldsymbol{\mathcal{F}}^p)^{-1} \\ \vdots \\ (\boldsymbol{\mathcal{F}}^p)^{-1} \end{bmatrix}}_{m\times p} \cdot \underbrace{\boldsymbol{\mathcal{F}}^p \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}}_{p\times 1}. \tag{2.224}$$

Now, the problem has been recast into the canonical CS framework, where the observation model comprises $m$ samples of a length-$p$ sparse (complex) vector. In addition, the effective observation matrix is formed by selecting a set of rows of the Fourier inverse matrix, which is the setting examined theoretically in the previous section. In other words, the results of Theorems 2.38 and 2.39 can be used to identify conditions on the number of samples $m$ (as a function of the degree of the polynomial) that must be satisfied in order for the effective observation matrix to satisfy the RIP. Any of a number of convex optimization schemes can then be employed to recover the spectral representation of the signal being observed.

Notice that in the above discussion, the total observation time required to obtain $m$ samples is no less than $\tau \cdot f(m) = t_d \cdot f(m)/p$. In some applications where the available observation time window is limited, or where the signal of interest is from a broader class of signals which are periodic but only for a short time duration (as is the case with frequency-hopping transmissions), this time requirement might be prohibitive. To overcome this, the second method we propose allows for the collection of the same number of samples over fewer periods of the signal (and consequently, less time) by designing a collection of sampling points in time a priori, based on the output of the polynomial function. Specifically, the results of Theorems 2.38 and 2.39 also apply when the sampling points are given by the entries of the set $T = \{\tau \cdot (f(n) \mod p) : n = 1, 2, \dots, m\}$.

Note that in this case, $T$ may contain duplicate entries, corresponding to indices $n_1 \neq n_2$ for which $f(n_1) = f(n_2) \mod p$. We suggest two possible ways to collect the required samples in this setting. First, for points in time where duplicate observations are required, a single observation can be collected and simply replicated as often as necessary in the overall vector of observations, prior to the implementation of a CS

recovery technique. Alternatively, the observations can be collected individually (at distinct times) over the periodic extension of the signal, similar in spirit to the above setting. The upshot here is that the observations can be collected over many fewer periods of the signal, and consequently, a shorter actual time.

Indeed, in order to collect distinct samples for each time prescribed in $T$, the number of periods of the signal that must be observed needs to be no greater than the largest multiplicity of any entry of $T$. By the Fundamental Theorem of Algebra, the number of solutions of the polynomial congruence $f(n) = 0 \mod p$ is at most the degree of the polynomial when $p$ is a prime and the polynomial has integer coefficients. It follows that the number of points $n_i$ for which the sample time indices $f(n_i)$ are equal (mod $p$) is no more than $d$, provided the number of measurements satisfies $m < p$. Consequently, all of the prescribed sampling points in time can be acquired by unique observations that occur over an *actual* observation time that need not exceed $t_d \cdot d$, which is much smaller than $t_d \cdot f(m)/p$ when $m$ satisfies the conditions prescribed in the theorems.

## 2.3.4 Appendix

### 2.3.4.1 Helpful Lemmata

The proofs will utilize a standard result in classical eigenanalysis known as *Geršgorin's Disc Theorem*, stated here as a lemma without proof (for more details see, for example, [42]).

**Lemma 2.40 (Geršgorin)** *The eigenvalues of a $p \times p$ complex matrix $\boldsymbol{M}$ all lie in the union of $p$ discs $d_j = d_j(c_j, r_j)$, $j = 1, 2, \ldots, p$, centered at $c_j = M_{j,j}$, and with radius*

$$r_j = \sum_{\substack{i=1 \\ i \neq j}}^{p} |M_{j,i}|. \tag{2.225}$$

When applying Geršgorin's Theorem, we will control the sums of off-diagonal elements using bounds on certain exponential sums. The following result is credited to H. Weyl [58], and appears in the present form in [59].

**Lemma 2.41 (Weyl)** *Let $d \geq 2$, and let $g(n) = a_1 n + \cdots + a_d n^d$, where $a_d = \alpha/p + \theta/p^2$, $|\theta| \leq 1$, and $\gcd(\alpha, p) = 1$. If, for $0 < \epsilon_2 < \epsilon_1 < 1$, the condition $m^{\epsilon_1} \leq p \leq m^{d-\epsilon_1}$ holds for some integer $m$, then*

$$\frac{1}{m} \left| \sum_{n=1}^{m} \exp\left(2\pi i g(n)\right) \right| \leq c(d, \epsilon_2) \cdot m^{(\epsilon_2 - \epsilon_1)/2^{d-1}}. \tag{2.226}$$

*For completeness, we note that the constant $c(d, \epsilon_2)$ is given by*

$$c(d, \epsilon_2) = 2 \left[ \left( \frac{64d}{\epsilon_2} \right) \left( \frac{d^2}{\epsilon_2 \log 2} \right)^{\exp(d^2/\epsilon_2)} d! \right]^{1/2^{d-1}}. \tag{2.227}$$

Theorem 2.39 utilizes an improvement of the above for large values of $d$, credited to I. Vinogradov [60], and appearing in its present form in [59].

**Lemma 2.42 (Vinogradov)** *Let $d > 2$, and let $g(n) = a_1 n + \cdots + a_d n^d$, where $a_d = \alpha/p + \theta/p^2$, $|\theta| \leq 1$, and $\gcd(\alpha, p) = 1$. If $m \leq p \leq m^{d-1}$, then*

$$\frac{1}{m}\left|\sum_{n=1}^{m} \exp\left(2\pi i g(n)\right)\right| \leq c(d) \cdot m^{1/(9d^2 \log d)}. \tag{2.228}$$

*For completeness, we note the constant $c(d) = \exp(3d)$.*

### 2.3.4.2 Proof of Theorem 2.38

Recall that the matrix of interest is the $|T| \times p$ matrix $m^{-1/2}\boldsymbol{\mathcal{F}}_T^p$, where the set $T$ indicates which rows of the DFT matrix $\boldsymbol{\mathcal{F}}^p$ are present. We begin by analyzing the Gram matrix $\boldsymbol{G} = m^{-1}(\boldsymbol{\mathcal{F}}_T^p)^H \boldsymbol{\mathcal{F}}_T^p$, where the superscript $H$ denotes the complex conjugate transpose (Hermitian) operator. Let $f(n)$ be any polynomial of degree $d > 2$ of the form $f(n) = a_1 n + \cdots + a_d n^d$, with real integer coefficients $a_j \in \{0, 1, \ldots, p-1\}$ for $j = 1, 2, \ldots, d-1$, and $a_d \in \{1, 2, \ldots, p-1\}$, as specified in the theorem. The entries of the Gram matrix $\boldsymbol{G}$ correspond to inner products between the columns of the matrix of interest, whose entries are given by

$$\begin{aligned}\left\{m^{-1/2}\mathcal{F}_T^p\right\}_{j,n} &= m^{-1/2}\exp\left(\frac{2\pi i j f(n) \mod p}{p}\right) \\ &= m^{-1/2}\exp\left(\frac{2\pi i j f(n)}{p}\right),\end{aligned} \tag{2.229}$$

for $j = 0, 1, \ldots, p-1$ and $n = 1, 2, \ldots, m$. Using this fact, we have that the entry of $\boldsymbol{G} = m^{-1}(\boldsymbol{\mathcal{F}}_T^p)^H \boldsymbol{\mathcal{F}}_T^p$ in row $y$ and column $z$ is given by

$$\{G\}_{y,z} = m^{-1}\sum_{n=1}^{m} \exp\left(\frac{2\pi i (z-y) f(n)}{p}\right), \tag{2.230}$$

for $y, z = 0, 1, \ldots, p-1$ and $n = 1, \ldots, m$.

The first point to note is that the diagonal elements of $\boldsymbol{G}$ (where $y = z$) are all equal to one since the argument of the exponential function is identically zero. For the off-diagonal elements, we can make the substitution $g(n) = (z-y)f(n)/p$ to obtain

$$\{G\}_{y,z} = m^{-1}\sum_{n=1}^{m} \exp\left(2\pi i g(n)\right), \tag{2.231}$$

for $y, z = 0, 1, \ldots, p-1$, $y \neq z$, and $n = 1, \ldots, m$.

Note that the coefficient of the highest order term in $g(n)$ is given by $(z-y)a_d/p$. For $z > y$, it is easy to see that the polynomial $g(n)$ satisfies the conditions of Lemma 2.41 because $p$ is prime and $(z-y)$ is a

positive integer less than $p$. For the case $z < y$, we can instead consider the polynomial $g'(n) = g(n) + n^d$, since in this case the periodicity of the complex exponential implies

$$\begin{aligned} \{G\}_{y,z} &= m^{-1} \sum_{n=1}^{m} \exp\left(2\pi i g(n)\right) && (2.232) \\ &= m^{-1} \sum_{n=1}^{m} \exp\left(2\pi i g'(n)\right). && (2.233) \end{aligned}$$

If $z < y$, then $(z - y)$ is negative but $(z - y + p)$ is positive since $y \leq p - 1$. Thus, the leading coefficient of $g'(n)$ is of the form of a positive integer (less than $p$) divided by $p$, and satisfies the conditions of Lemma 2.41 in this case as well. If, in addition, $0 < \epsilon_2 < \epsilon_1 < 1$ and $m$ satisfies the condition $p^{1/(d-\epsilon_1)} \leq m \leq p$, we can apply Lemma 2.41 to bound the magnitude of the off-diagonal elements of $\boldsymbol{G}$ by

$$\left|\{G\}_{y,z}\right| \leq c(d, \epsilon_2) \cdot m^{(\epsilon_2 - \epsilon_1)/2^{d-1}}, \tag{2.234}$$

for $y, z = 0, 1, \ldots, p-1$ and $y \neq z$.

Now, let $\Omega$ be a subset of $\{0, 1, \ldots, p-1\}$ satisfying $|\Omega| \leq s$, and denote by $\boldsymbol{\mathcal{F}}_{T,S}^p$ the $|T| \times |\Omega|$ submatrix of $m^{-1/2}\boldsymbol{\mathcal{F}}_T^p$ formed by retaining the columns indexed by the elements of $\Omega$. Notice that if for each unique choice of $\Omega$, the eigenvalues of the Gram matrix $\boldsymbol{G}_\Omega = m^{-1}(\boldsymbol{\mathcal{F}}_{T,\Omega}^p)^H \boldsymbol{\mathcal{F}}_{T,\Omega}^p$ are all within the range $(1 - \delta_s, 1 + \delta_s)$ then the RIP is satisfied. But, since each Gram matrix $\boldsymbol{G}_\Omega$ is a proper submatrix of $\boldsymbol{G}$, Geršgorin's Theorem guarantees that the eigenvalue bound is satisfied simultaneously for all choices of $\Omega$ whenever the sum of any $(s-1)$ off-diagonal elements of $\boldsymbol{G}$ does not exceed $\delta_s$. In other words, the RIP is satisfied whenever

$$\begin{aligned} s &\leq \frac{\delta_s}{c(d, \epsilon_2)} m^{(\epsilon_1 - \epsilon_2)/2^{d-1}} \\ &= \delta_s \cdot C(d, \epsilon_2) \cdot m^{(\epsilon_1 - \epsilon_2)/2^{d-1}}, && (2.235) \end{aligned}$$

as claimed.

The same method of proof also establishes the RIP for the matrix $pm^{-1/2}\left(\boldsymbol{\mathcal{F}}^p\right)_T^{-1}$.

### 2.3.4.3 Proof of Theorem 2.39

The proof proceeds along the same lines as above, with references to Lemma 2.42 instead of Lemma 2.41 where appropriate.

## 2.3.5 Acknowledgment

The authors would like to thank Jerry Fudge at L-3 Communications/Integrated Systems, whose structured non-uniform sampling architecture motivated this research direction.

# Chapter 3

# Advances in Adaptive Sensing

The results in this chapter provide significant new insight into the fundamental theoretical limits of sparse signal recovery. It is shown here that a simple adaptive sensing procedure, which directs subsequent observations based on the results of previous observations in an effort to focus on features of interest, dramatically outperforms the *best possible* nonadaptive sampling strategies in a variety of recovery tasks.

Section 3.1 is a reprint of a conference paper in which the adaptive sampling procedure now known as distilled sensing (DS) was first proposed and analyzed. In the context of sparse signal detection, this work established the first improvements in large-sample recoverability thresholds for adaptive sampling, relative to methods based on non-adaptive sensing. The original paper appeared at the 2008 Asilomar Conference on Signals, Systems, and Computers [6].

A follow-on effort, motivated by the problem of recovering the the location of the nonzero entries of a sparse signal (the sparsity pattern), is reprinted in Section 3.2. Here, it is shown that a refinement of the DS procedure enables even more dramatic improvements in the thresholds of recoverability. This work expands slightly upon the original paper, which appeared at the 12th International Conference on Artificial Intelligence and Statistics [7]. Namely, the version presented here contains a formal proof of the fundamental limits of sparsity pattern recovery using non-adaptive sampling—a result that was utilized in the original version of the paper, but was presented there without proof.

As in the previous chapter, each section essentially a stand-alone entity and there may be some notational differences across sections.

## 3.1 Adaptive Discovery of Sparse Signals in Noise

A multi-step adaptive resampling procedure is proposed, and shown to be an effective approach when detecting high-dimensional sparse signals in noise. Each step of the proposed procedure refines an estimate of the true signal subspace, allowing sensing energy to be focused more directly into the subspace of interest and significantly improving the performance of the final detection test. Large-sample analysis shows that for the sparse signal detection problems considered, the proposed adaptive sensing procedure outperforms the best possible detection methods based on non-adaptive sensing, allowing for the detection of signals that are exponentially weaker than what can be detected using non-adaptive samples.[1]

### 3.1.1 Introduction

In many high-dimensional inference problems, the task is to detect the presence of a signal from noisy measurements. For example, in sensor network monitoring applications, a large number of sensors might be deployed over a geographical region with the goal of determining whether a signal (perhaps a chemical or biological agent) is present anywhere in the region. Other applications in microarray analysis, astronomy, and covert communications are also easy to envision. Often, signals of interest in these applications are sparse (i.e., they exist in a very low-dimensional, but unknown, subspace).

Natural questions emerge regarding the performance of any detection method in such settings. For example, one is often concerned with the trade-offs among various problem parameters—namely, the sparsity level (signal subspace dimension), the amplitudes of each nonzero component, and the observation noise power. Consider the following $n$-dimensional signal plus noise observation model:

$$Y_i \ = x_i + z_i = \mu \ s_i \ + \ Z_i \ , \quad i = 1, \ldots, n, \tag{3.1}$$

where $Z_i \sim \mathcal{N}(0,1)$ denotes a Gaussian distribution with mean 0 and variance 1, the scalar $\mu$ denotes the signal strength and $s = [s_1, \ldots, s_n]$ is a binary vector whose non-zero entries indicate the signal subspace. In this setting, the detection problem amounts to determining whether $\mu \neq 0$. A simple detection scheme for this problem is based on linear data fusion. The data are averaged to obtain the test statistic $T = \frac{1}{n}\sum_{i=1}^{n} Y_i$, which is then compared to a threshold. Note that $\sqrt{n}T \ \sim \ \mathcal{N}(\mu d_n/\sqrt{n}, 1)$, where $d_n = \sum_{i=1}^{n} s_i$ (the signal subspace dimension) describes the relative the sparsity level. If $d_n/\sqrt{n}$ is bounded away from zero when $n$ tends to infinity (and $\mu$ is nonvanishing), then both the false-alarm and miss probabilities can be driven to zero using an appropriate threshold.

Note that this simple linear fusion approach can break down if the signal subspace dimension grows sublinearly in $n$. If, as $n \to \infty$, $d_n/\sqrt{n} \to 0$, then for any fixed $\mu$, $\sqrt{n}T \sim \mathcal{N}(\mu d_n/\sqrt{n}, 1) \to \mathcal{N}(0,1)$,

[1] The material in Section 3.1 is ©2008 IEEE. Reprinted, with permission, from *Proceedings of the 42nd Asilomar Conference on Signals, Systems, and Computers*, "Adaptive Discovery of Sparse Signals in Noise," J. Haupt, R. Castro, and R. Nowak.

implying that for any threshold the probability of miss does not tend to 0. An alternate approach is to test whether $\max_i Y_i$ exceeds a threshold. If $\mu = 0$, then in the large-sample limit $\max_i Y_i \leq \sqrt{2\log n}$ with probability 1, implying signals can be reliably detected when $\mu > \sqrt{2\log n}$. The limits of detectability using observations collected under the model (3.1) were established in [61, 62], where it was shown that more subtle testing procedures based on order statistics succeed at detecting slightly weaker signals in this setting ($\mu > \sqrt{2r\log n}$ for some $r < 1$).

The limits of detectability established in [61, 62] were based on the assumption that the observations of the signal are collected *non-adaptively*. This implies sensing/sampling resources must be equally allocated over the entire observation space (since one does not know the signal subspace a priori). In contrast, here we consider an adaptive procedure that utilizes feedback in sensing. Formally, adaptive sampling approaches allow the sampling procedure to be adjusted using previously collected samples. In the sparse signal detection problems considered here, adaptivity could allow subsequent samples to be focused more directly into subspaces where the signal of interest, if present, is most likely to exist, increasing the measurement signal to noise ratio (SNR) and ultimately improving the performance of any detection method.

In this paper we propose a multi-step adaptive resampling procedure that delivers such an improvement. In one setting we show that our adaptive sampling procedure (using a fixed number of resampling steps) effectively improves the range of problem parameters for which sparse signals are detectable. By letting the number of resampling steps tend to infinity (slowly as a function of the signal dimension $n$) we also establish that the proposed procedure succeeds in detecting signals that are *exponentially weaker* than any signal that can be detected using non-adaptive observations.

The remainder of this paper is organized as follows. In Section 3.1.2 we formalize the sparse signal and observation models that will be employed throughout. Known results on detectability using non-adaptive observations are summarized in Section 3.1.3. We propose our adaptive sampling procedure in Section 3.1.4, and describe its performance benefit relative to non-adaptive sampling in Section 3.1.5. A few brief conclusions are drawn in Section 3.1.6. Proofs of the main results are relegated to the Appendix.

### 3.1.2 Signal and Observation Models

Throughout this paper we will be concerned with the detection of signals $x \in \mathbb{R}^n$ which exhibit sublinear sparsity, and for simplicity we restrict our attention to signals in which each nonzero entry has the same amplitude. We will assume that the signals of interest are generated randomly, according to the following model. For a fixed $\beta \in (1/2, 1)$, we let $\epsilon(n) = n^{-\beta}$ and

$$s_i = \begin{cases} 1 & \text{with probability } \epsilon(n) \\ 0 & \text{with probability } 1 - \epsilon(n) \end{cases}, \; i = 1, 2, \ldots, n, \tag{3.2}$$

and we let $x_i = \mu s_i$ for some scalar amplitude $\mu$. The random signal model is chosen primarily to facilitate comparison between our adaptive procedure and prior work that uses non-adaptive observation models. Note, however, that for large $n$, signals so generated can be identified as having $m(n) = n^{1-\widetilde{\beta}}$ nonzero entries, where $\widetilde{\beta} \approx \beta$ with very high probability. This claim is easily formalized using standard concentration inequalities.

We assume that observations of $x$ come from multiple "looks," indexed by $j$, of the form

$$Y^{(j)} = \phi^{(j)} \cdot x \ + \ Z^{(j)} \ , \tag{3.3}$$

where each $\phi^{(j)} \in \mathbb{R}^n$ is a *sensing vector* with non-negative entries, $Z^{(j)} \overset{iid}{\sim} \mathcal{N}(0, 1_{n\times n})$, and the operation $\phi^{(j)} \cdot x$ denotes the Hadamard (or element-wise) product of the vectors. In addition, we impose the restriction $\sum_j \|\phi^{(j)}\|_2^2 = n$, limiting the total energy of the sensing vectors.

Using this model, an example of non-adaptive sampling arises when only one look at the signal is obtained. In this case, (3.3) can be written as

$$Y^{(1)} = Y = \phi \cdot x \ + \ Z \ , \tag{3.4}$$

and to satisfy the sensing energy condition, we can choose $\phi_i^{(1)} = \phi_i = 1$ for $i = 1, \ldots, n$. Another possibility is to make multiple iid observations, but each with only a fraction of the total sensing energy budget. For example, let $p \in \mathbb{N}$ denote the total number of looks, and let $\phi_i^{(j)} = 1/\sqrt{p}$, $i = 1, \ldots, n$, for $j = 1, \ldots, p$. Because of the independence of the noises, $\sum_{j=1}^{p} Y^{(j)}$ is statistically equivalent to $Y$ in the single-observation model (3.4) as well. There are obviously many other choices of $\{\phi^{(j)}\}_j$ that yield the same result. Furthermore, no non-adaptive sensing scheme exists that can produce better results than those obtained using observations from the standard model (3.4). Therefore, we are interested here in adaptive designs of $\{\phi^{(j)}\}_j$ that tend to focus in on the non-zero components of $x$. Specifically, in what follows we will allow $\phi^{(j)}$ to depend explicitly on $\{\phi^{(\ell)}, Y^{(\ell)}\}_{\ell<j}$, and we will show that information gleaned from previous observations can be used to effectively guide future sampling, yielding significant performance improvements.

### 3.1.3 Large-Sample Detection Thresholds from Non-Adaptive Observations

In [61], a fundamental large sample detectability threshold for sparse signals using non-adaptive observations was established. Let the fraction of nonzero entries be denoted by $\epsilon(n)$, and the amplitude of each be $\mu(n)$. When samples are drawn according to (3.4), the detection problem amounts to a hypothesis test between a joint null distribution and an alternative mixture distribution:

$$\begin{aligned} H_0 : &\quad Y_i \overset{iid}{\sim} \mathcal{N}(0,1) \\ H_1 : &\quad Y_i \overset{iid}{\sim} (1-\epsilon(n))\mathcal{N}(0,1) + \epsilon(n)\mathcal{N}(\mu(n),1) \end{aligned} \ , \tag{3.5}$$

for $i = 1, \ldots, n$. Carefully specifying how $\epsilon(n)$ and $\mu(n)$ evolve with $n$ allows for a concise statement about the detectability threshold behavior. Specifically, for a fixed $\beta \in (1/2, 1)$ let the fraction of nonzero entries

be given by $\epsilon(n) = n^{-\beta}$. In addition, suppose that the amplitude of each nonzero entry is $\mu(n) = \sqrt{2r \log n}$ for some $r \in (0, 1)$. The following result holds [61].

**Theorem 3.1** *Let*

$$\rho^*(\beta) = \begin{cases} \beta - \frac{1}{2}, & \frac{1}{2} < \beta \leq \frac{3}{4} \\ \left(1 - \sqrt{1-\beta}\right)^2, & \frac{3}{4} < \beta < 1 \end{cases} . \tag{3.6}$$

*When $r > \rho^*(\beta)$, then there exists a test for which the sum of Type I and Type II errors (the false alarm and miss errors, respectively) tends to zero as $n \to \infty$. Conversely, when $r < \rho^*(\beta)$, then the sum of Type I and Type II errors for* any *test tends to one as $n \to \infty$.*

Of course, the likelihood ratio test is one such test that achieves the performance described above. However, successfully implementing the test requires knowledge of the problem parameters $r$ and $\beta$. In [61] the authors proposed an alternative procedure called *Higher Criticism* that utilizes properties of the order statistics of the observed data to achieve the same performance as the likelihood ratio test but without requiring knowledge of the problem parameters.

### 3.1.4 Adaptive Testing and Resampling

In order to improve upon the threshold behavior specified in Theorem 3.1, we consider sampling methods that allow feedback. The key to our adaptive approach is a multi-step testing and resampling procedure, described as Algorithm 1, which proceeds as follows. Initially, all locations of $x$ are measured, but with only a fraction of the total sensing energy budget arising from an equal allocation of sensing energy among all observation steps. Each observation takes the form (3.3), where the sensing energy allocated to that step is distributed equally among a subset of promising entries of $x$. Following each of the first $k$ observation steps, a refinement test is performed, which identifies the subset of the promising locations where the current observation is positive. The rationale is that it is highly improbable that the signal (which is assumed to be positive) is present at locations where the current observation is negative. The algorithm terminates after the final observation, and the output consists of the final collection of observations and a set of possible signal component locations. A pseudocode description of the procedure appears as Algorithm 1.

To quantify the performance of this algorithm, we will show that the refinement test at each step retains *most* (in fact all, in the limit as $n$ tends to infinity) of the indices corresponding to nonzero signal components, but only about *half* of the indices corresponding to zero entries. When the signal is sparse, this implies that the effective dimension is roughly halved at each step, allowing about twice as much sensing energy to be allocated to the signal subspace in the next step. Applying Higher Criticism testing to the output observations results in significantly lower large-sample thresholds for consistent detectability, as described in the next section.

### 3.1.5 Main Results

The large-sample threshold for consistent sparse signal detection based on the proposed adaptive resampling procedure is examined in two settings; one where the number of observations, $k+1$, is finite, and another where $k$ is allowed to tend to infinity slowly as a function of the dimension $n$. The results are stated in this section. The proofs are relegated to the Appendix.

Our first main result quantifies the performance of adaptive resampling followed by Higher Criticism testing when the number of refinement steps is fixed. The result is an expanded region of large-sample detectability in the $(\beta, r)$ plane.

**Theorem 3.2** *For a fixed $\beta \in (1/2, 1)$, consider sparse signals generated according to the model (3.2), and suppose that the amplitude of each nonzero entry is given by $\mu(n) = \sqrt{2r \log n}$ for some $r \in (0,1)$. There exists an improved threshold function,*

$$\begin{aligned} \rho_a^*(\beta, k) &= \left(\frac{k+1}{2^k}\right) \rho^*(\beta) \\ &= \begin{cases} \frac{k+1}{2^k}\left(\beta - \frac{1}{2}\right), & \frac{1}{2} < \beta \leq \frac{3}{4} \\ \frac{k+1}{2^k}\left(1 - \sqrt{1-\beta}\right)^2, & \frac{3}{4} < \beta < 1 \end{cases}, \end{aligned} \tag{3.7}$$

*such that when $r > \rho_a^*(\beta, k)$, Higher Criticism testing applied to $Y^{(k+1)}$ (the output of the adaptive procedure with $k$ refinement testing steps) succeeds in detecting the sparse signal in the sense that the sum of Type I and Type II errors tends to zero as $n \to \infty$.*

This result shows that even a small number of refinement steps leads to a significant improvement in terms of the large-sample threshold for consistent recovery (since $(k+1)/2^k \ll 1$ for even modest values of $k$). Figure 3.1 shows the large-sample threshold of Theorem 3.1 (solid line) along with several of the improved thresholds for various values of $k$ (the threshold for $k = 3$ is shown as a dotted line, and the threshold for $k = 7$ is the dashed line).

Note that the result of Theorem 3.2 suggests increasing the number of refinement steps results in increasingly improved recovery thresholds. A natural question arises as to whether the number of observation steps can be large enough to detect signals whose amplitudes grow more slowly (as a function of the dimension $n$) than $\mu(n) = \sqrt{2r \log n}$. We address this question here by letting the number of observation steps tend to infinity slowly as a function of $n$. The result is the following theorem.

**Theorem 3.3** *For a fixed $\beta \in (1/2, 1)$, consider sparse signals generated according to the model (3.2), and suppose that for a fixed $\delta > 0$, the amplitude of each nonzero entry is given by $\mu(n) = \sqrt{2r\left(\log_2 \log n\right)^{1+\delta}}$ for some $r \in (0,1)$. Let the number of refinement steps be given by $k(n) = \log_2 \log n$. When $r > \rho^*(\beta)$ (as*

*defined in Theorem 3.1), Higher Criticism testing applied to $Y^{(k+1)}$ succeeds in detecting the sparse signal in the sense that the sum of Type I and Type II errors tends to zero as $n \to \infty$.*

In other words, the adaptive procedure followed by Higher Criticism testing succeeds at detecting sparse signals with non-zero amplitudes as small as $\sqrt{2r(\log_2 \log n)^{1+\delta}}$, for $\delta$ arbitrarily small, but fixed. For the same specification on $r$, non-adaptive sensing can only recover signals with amplitudes larger than $\sqrt{2r \log n}$. Thus, the proposed adaptive procedure is capable of detecting sparse signals that are essentially *exponentially weaker* than those recoverable using non-adaptive sensing.

### 3.1.6 Discussion and Conclusions

While we only considered the case where the amplitudes of the nonzero signal entries were equal and positive, the proposed adaptive procedure could also be applied to more general classes of signals, such as those for which the signal has both positive and negative values. In this case, one approach is to split the budget of sensing energy in half, and execute the procedure once assuming the nonzero entries are positive (keeping all entries that exceed zero at each step, as described above), and again assuming the entries are negative (retaining indices at each for which the corresponding observation is negative). The final sets of observations could then be independently subjected to Higher Criticism testing.

In general, we have shown that a simple adaptive procedure provably outperforms all non-adaptive sampling techniques when detecting sparse signals in additive noise. We note that analogous improvements can also be realized using the same procedure in estimation problems, and this setting will be treated in a future work.

### 3.1.7 Appendix

We first state two lemmata that will be useful in the proofs. The first describes a tail bound for the Binomial distribution, which follows from [63].

**Lemma 3.4** *Let $B$ be a Binomial$(n, p)$ random variable, and assume that $b < \mathbb{E}[B] = np$. Then,*

$$\Pr\left(B \leq b\right) \leq \left(\frac{n-np}{n-b}\right)^{n-b} \left(\frac{np}{b}\right)^{b}. \tag{3.8}$$

We will also need to bound tail probabilities for certain Gaussian random variables. For that, we will utilize the following standard result (see, for example, [64]).

**Lemma 3.5** *Let $Z \sim \mathcal{N}(z, 1)$ for $z > 0$. Then,*

$$\Pr\left(Z < 0\right) \leq \frac{1}{z\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right). \tag{3.9}$$

The proofs of both main theorems are similar in nature. To save space we formally present the proof of Theorem 3.2, then briefly describe the generalizations that are needed to establish Theorem 3.3.

### 3.1.7.1 Proof of Theorem 3.2

We first start with a lemma that quantifies the effect of each refinement step. Recall that $I^{(j)}$ is the index set of iteration $j$ (refer to Algorithm 1 for details).

**Lemma 3.6** *Let $\ell_j$ and $m_j$ denote, respectively, the number of indices corresponding to zero and non-zero elements of $x$ in $I^{(j)}$. Conditioning on $m_j$ and $\ell_j$ we have that, for $n$ sufficiently large,*

$$m_{j+1} \;\geq\; \left(1 - \frac{1}{\log n}\right) m_j, \tag{3.10}$$

*and*

$$\left(\frac{1}{2} - \frac{1}{\log n}\right) \ell_j \;\leq\; \ell_{j+1} \;\leq\; \left(\frac{1}{2} + \frac{1}{\log n}\right) \ell_j, \tag{3.11}$$

*hold simultaneously with probability at least $1 - \xi$, where*

$$\xi = \exp\left(-n^{-\alpha} m_j\right) + 2\exp\left(-2\ell_j / \left(\log n\right)^2\right), \tag{3.12}$$

*for any $\alpha > 0$.*

**Proof:** Note that in any step of the $(k+1)$-step procedure, the observed amplitude of non-zero signal components is no less than $\sqrt{2r \log n/(k+1)}$. It suffices to consider this worst case to establish a general bound.

Let $T_i = \mathbf{I}\{Y_i^{(j)} > 0\}$, $i \in I^{(j)}$, be the indicators of the indices retained in $I^{(j+1)}$. Let $\mathcal{I}_s \subseteq I^{(j)}$ be the collection of indices corresponding to non-zero components (therefore $|\mathcal{I}_s| = m_j$). From Lemma 3.5, we have for $i \in \mathcal{I}_s$

$$\Pr\left(T_i = 0\right) \leq \sqrt{\frac{k+1}{4\pi r \log n}}\, n^{-r/(k+1)}. \tag{3.13}$$

The number of signal components retained after the refinement step is simply given by $m_{j+1} = \sum_{i \in \mathcal{I}_s} T_i$. Let $\gamma$ be a small quantity that satisfies $\gamma > \Pr\left(T_i = 0\right)$, and apply Lemma 3.4 to obtain

$$\begin{aligned} \Pr\left(m_{j+1} \leq (1-\gamma) m_j \,|\, m_j\right) \;&\leq\; \left(\frac{\Pr\left(T_i = 0\right)}{\gamma}\right)^{\gamma m_j} \left(\frac{\Pr\left(T_i = 1\right)}{1-\gamma}\right)^{(1-\gamma)m_j} \\ &\leq\; \left(\frac{\Pr\left(T_i = 0\right)}{\gamma}\right)^{\gamma m_j} \left(\frac{1}{1-\gamma}\right)^{(1-\gamma)m_j}. \end{aligned} \tag{3.14}$$

Now, for $\alpha > 0$, notice that if

$$\gamma \log\left(\frac{\Pr\left(T_i = 0\right)}{\gamma}\right) + (1-\gamma)\log\left(\frac{1}{1-\gamma}\right) \leq -n^{-\alpha} \tag{3.15}$$

then the probability that (3.10) does not hold is given by

$$\Pr\left(m_{j+1} \leq (1-\gamma) m_j \,|\, m_j\right) \leq \exp\left(-n^{-\alpha} m_j\right). \tag{3.16}$$

Let $\gamma = (\log n)^{-1}$. We will now check that (3.15) holds, using (3.13) and the fact that

$$\left(\frac{1-\gamma}{\gamma}\right)\log\left(\frac{1}{1-\gamma}\right) \leq 1 \tag{3.17}$$

for any $\gamma \in (0,1)$. After trivial manipulation we see that (3.15) is satisfied if

$$\frac{-r\log n}{k+1} + \frac{\log\log n}{2} + \frac{\log n}{n^{\alpha}} + \left(1 + \frac{1}{2}\log\left(\frac{k+1}{4\pi r}\right)\right) \leq 0, \tag{3.18}$$

which is clearly the case for $n$ sufficiently large, since the first term dominates the second, the third decays to zero, and the remainder does not depend on $n$.

To verify the inequalities in (3.11), define $\mathcal{I}_0 \subseteq I^{(j)}$ to be the collection of indices corresponding to zero entries (therefore $|\mathcal{I}_0| = \ell_j$). Note that for $i \in \mathcal{I}_0$, $T_i$ is a Bernoulli random variable with parameter $1/2$ (since each corresponding observation $Y_i^{(j)}$ is a zero-mean Gaussian random variable). Applying Hoeffding's inequality to $\ell_{j+1} = \sum_{i\in\mathcal{I}_0} T_i$ we obtain

$$\Pr\left(\left|\ell_{j+1} - \frac{\ell_j}{2}\right| > \gamma\ell_j \middle| \ell_j\right) \leq 2\exp\left(-2\ell_j\gamma^2\right). \tag{3.19}$$

The probability that both claimed conditions hold simultaneously is obtained by applying the union bound, concluding the proof. ∎

The proof of Theorem 3.2 proceeds by iterating the result of Lemma 3.6. We begin by considering the set $I^{(2)}$. Let $\alpha = (1-\beta)/2$, and note that with probability at least $1 - \xi(2)$, where

$$\xi(2) = \exp\left(-n^{(1-\beta)/2}\right) + 2\exp\left(-\frac{2n(1-n^{-\beta})}{(\log n)^2}\right), \tag{3.20}$$

the event

$$E_2 = \left\{ \begin{array}{l} \left(1 - \frac{1}{\log n}\right) n^{1-\beta} \leq m_2 \leq n^{1-\beta} \\ \\ \left(\frac{1}{2} - \frac{1}{\log n}\right) n(1-n^{-\beta}) \leq \ell_2 \\ \\ \ell_2 \leq \left(\frac{1}{2} + \frac{1}{\log n}\right) n(1-n^{-\beta}) \end{array} \right\} \tag{3.21}$$

holds when $n$ is sufficiently large.

Now, conditioned on event $E_2$, we can apply Lemma 3.6 to $I^{(3)}$. Proceeding similarly for sets $I^{(4)}, \ldots, I^{(k+1)}$, and noting that $\xi(i) < \xi(j)$ when $i < j$, we have that for $n$ sufficiently large the event

$$E_{k+1} = \left\{ \begin{array}{l} \left(1 - \frac{1}{\log n}\right)^k n^{1-\beta} \leq m_{k+1} \leq n^{1-\beta} \\ \\ \left(\frac{1}{2} - \frac{1}{\log n}\right)^k n(1-n^{-\beta}) \leq \ell_{k+1} \\ \\ \ell_{k+1} \leq \left(\frac{1}{2} + \frac{1}{\log n}\right)^k n(1-n^{-\beta}) \end{array} \right\} \tag{3.22}$$

holds with probability at least $1 - k\xi(k+1)$ where

$$\xi(k+1) = \exp\left(-\left(1 - \frac{1}{\log n}\right)^{k-1} n^{(1-\beta)/2}\right) + 2\exp\left(-\frac{2n(1-n^{-\beta})\left(\frac{1}{2} - \frac{1}{\log n}\right)^{k-1}}{(\log n)^2}\right). \tag{3.23}$$

Note that, as $n \to \infty$ both of the terms in $\xi(k+1)$ tend to zero, implying that $E_{k+1}$ holds with probability tending to 1. The last step of the proof consists of showing what happens when Higher Criticism testing is applied after the last refinement. Conditionally on event $E_{k+1}$ we are left with an 'effective' signal vector with length $n_e = |I^{(k+1)}|$, where

$$\left(1 - \frac{1}{\log n}\right)^k n^{1-\beta} + \left(\frac{1}{2} - \frac{1}{\log n}\right)^k n(1-n^{-\beta}) \le n_e \le n^{1-\beta} + \left(\frac{1}{2} + \frac{1}{\log n}\right)^k n(1-n^{-\beta}). \tag{3.24}$$

The final observation, $Y_a^{(k+1)}$, allocates $n/(k+1)$ energy among the $n_e$ indices, resulting in an effective signal amplitude of

$$\sqrt{2r\frac{n\log n}{(k+1)n_e}}, \tag{3.25}$$

for each non-zero signal component. To specify the effective signal parameters $r_e$ and $\beta_e$ for this reduced dimension problem, we equate (3.25) with $\sqrt{2r_e \log n_e}$, which implies

$$r_e = \frac{r}{k+1}\frac{n\log n}{n_e \log n_e}. \tag{3.26}$$

Noting that when $n \to \infty$, $n/n_e \to 2^k$ and $\log n / \log n_e \to 1$, we conclude that

$$r_e \to r\frac{2^k}{k+1}. \tag{3.27}$$

We proceed similarly for the sparsity parameter $\beta_e$, to determine $\beta_e \to \beta$. Thus, the final observation $Y^{(k+1)}$, is equivalent to a single observation of a signal with length $n_e$, having $n_e^{1-\beta_e}$ nonzero entries of amplitude $\sqrt{2r_e \log n_e}$. Applying Higher Criticism testing to these observations, and noting that $n_e \to \infty$ as $n \to \infty$ results in the threshold behavior described in Theorem 3.1 (but with the new parameters $\beta_e$ and $r_e$).

### 3.1.7.2 Sketch of Proof of Theorem 3.3

The proof mimics that of Theorem 3.2 with differences as noted here. We first generalize the signal model by letting $\mu(n) = \sqrt{2rg(n)}$, and we let $k = k(n)$ denote the total number of refinements, which is now a function of $n$. An analogous claim to that of Lemma 3.6 can be obtained for a general tolerance $\gamma(n)$. Iterating this result under the following conditions:

- $g(n) = (\log_2 \log n)^{1+\delta}$, for any fixed $\delta > 0$,
- $k(n) = \log_2 \log n$,

- $\gamma(n) = (\log_2 \log n)^{-2}$,

it can be shown that the output of the adaptive procedure is equivalent to observations of a sparse signal of length $n_e \to \infty$ as $n \to \infty$, having $n_e^{1-\beta_e}$ nonzero entries of amplitude at least $\sqrt{2r_e \log n_e}$, where $r_e = r$ and $\beta_e = \beta$, establishing the claim.

**Algorithm 3.1**: Adaptive Testing and Resampling.

**Input**:

Number of refinements $k$

**Initialize:**

Initial index set $I^{(1)} \longleftarrow \{1, 2, \ldots, n\}$

Energy per observation $\mathcal{E} = n/(k+1)$

**Refinement:**

**for** $j = 1$ **to** $k$ **do**

$$Y_i^{(j)} = \left\{ \begin{array}{ll} \sqrt{\frac{\mathcal{E}}{|I^{(j)}|}} x_i + Z_i^{(j)}, & i \in I^{(j)} \\ Z_i^{(j)}, & i \in I^{(1)} \setminus I^{(j)} \end{array} \right\}$$

$$I^{(j+1)} \longleftarrow \{i \in I^{(j)} : Y_i^{(j)} > 0\}$$

**end**

**Final Observations:**

$$Y_i^{(k+1)} = \left\{ \begin{array}{ll} \sqrt{\frac{\mathcal{E}}{|I^{(k+1)}|}} x_i + Z_i^{(k+1)}, & i \in I^{(k+1)} \\ Z_i^{(k+1)}, & i \in I^{(1)} \setminus I^{(k+1)} \end{array} \right\}$$

**Output**:

Final index set $I^{(k+1)}$

Refined obs. $Y^{(k+1)} := \{Y_i^{(k+1)} : i \in I^{(k+1)}\}$

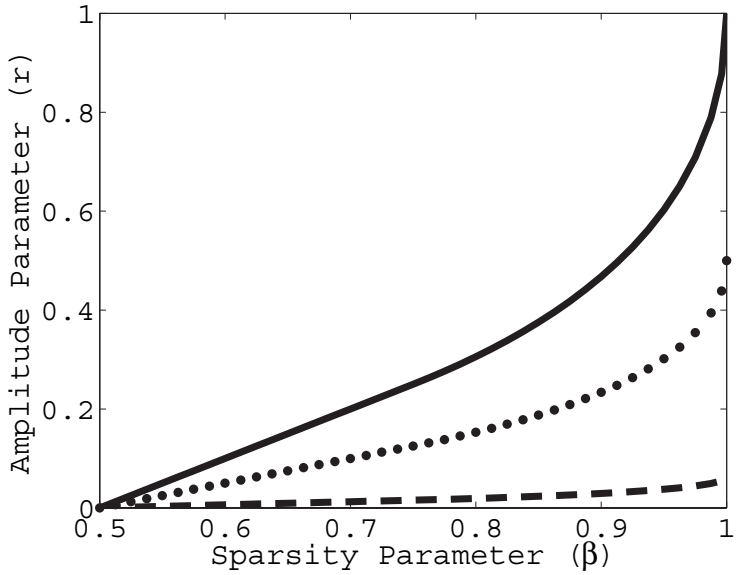



Figure 3.1 Large-Sample Thresholds for Consistent Detection. Signals whose sparsity and amplitude parameters are above the thresholds can be reliably detected. The solid line shows the threshold of Theorem 3.1 obtained using observations from the standard model (3.4), while the thresholds obtained using the adaptive procedure (for $k = 3$ and $k = 7$ refinement steps) are shown as dotted and dashed lines, respectively.

## 3.2 Distilled Sensing for Sparse Recovery

A selective sampling procedure called *distilled sensing* (DS) is proposed, and shown to be an effective method for recovering sparse signals in noise. Based on the notion that it is often easier to rule out locations that do not contain signal than it is to directly identify non-zero signal components, DS is a sequential method that systematically focuses sensing resources towards the signal subspace. This adaptivity in sensing results in rather surprising gains in sparse signal recovery—dramatically weaker sparse signals can be recovered using DS compared with conventional non-adaptive sensing procedures.[2]

### 3.2.1 Introduction

Consider the following canonical signal model.

$$X_i \sim \mathcal{N}(\mu_i, 1) \ , \ i = 1, \ldots, n, \tag{3.28}$$

where $\mathcal{N}(\mu_i, 1)$ denotes the normal distribution with mean $\mu_i$ and unit variance. The signal $\mu = (\mu_1, \ldots, \mu_n)$ is sparse if most of the components $\mu_i$ are zero. Identifying the locations of the non-zero components based on the data $X = (X_1, \ldots, X_n)$ when $n$ is very large is a fundamental problem arising in many applications, including fMRI [65], microarray analysis [66], and astronomical surveying [67]. A common approach in these problems entails coordinate-wise thresholding of the observed data $X$ at a given level, identifying the locations whose corresponding observation exceeds the threshold as signal components. Among such methods, false-discovery rate (FDR) analysis [68] tends to be the procedure of choice because it is less conservative than Bonferroni correction, making it more useful in practice, and because it enjoys asymptotically optimal performance characteristics [68–72].

Suppose that the number of non-zero components of $\mu$ grows sublinearly in $n$ according to $n^{1-\beta}$ for $\beta \in (0, 1)$, and that each non-zero component takes the same (positive) value $\sqrt{2r \log n}$ for $r > 0$. For a given recovery procedure, define the *false-discovery proportion* (FDP) to be the number of falsely discovered components relative to the total number of discoveries, and the *non-discovery proportion* (NDP) to be the number of non-zero components missed relative to the total number of non-zero components. The asymptotic limits of sparse recovery for data collected according to (3.28) are sharply delineated in the $(\beta, r)$ parameter space. Specifically, if $r < \beta$, no recovery procedure based on coordinate-wise thresholding of the observed data can drive the FDP and NDP to zero as $n \to \infty$. But, when $r > \beta$, there exists a recovery procedure based on coordinate-wise thresholding that drives both the NDP and FDP to zero as $n \to \infty$. Thus, the relation $r = \beta$ defines a sharp asymptotic boundary in the parameter space, identifying when sparse signals

[2]Material in Section 3.2 is reprinted from *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*, "Distilled Sensing: Selective Sampling for Sparse Signal Recovery," J. Haupt, R. Castro, and R. Nowak. Under the terms of the original publishing agreement, the authors retain copyright ownership of the work.

observed under the model (3.28) can be reliably recovered. Under similar sparse signal and noise models, several related works established sharp asymptotics in signal estimation and classification settings [69–72].

Suppose that instead of a single observation of a sparse signal in noise, one were able to take multiple 'looks,' possibly adjusting the focus in a sequential fashion. Similar adaptive methods have been proposed in the signal processing literature [73], and they certainly are conceivable in applications such as microarray analysis and astronomical surveying. Here we consider an approach based on the idea that after the first look, one might be able to rule out a large number of locations where the signal is probably not present, and then focus sensing resources into the remaining locations in question. Formalizing this idea, we show that for the same budget of sensing resources, sequential adaptive sensing procedures dramatically outperform non-adaptive procedures, resulting in different scaling laws in terms of estimability. Rather than requiring that the non-zero components obey $\mu_i > \sqrt{2\beta \log n}$, we show that a novel adaptive sensing procedure called *distilled sensing* (DS) guarantees that sparse signals at level $\mu_i = \alpha(n)$, where $\alpha(n)$ is any positive monotone diverging sequence in $n$ that exceeds the $m$-fold iterated logarithm function,

$$\log^{[m]} n = \underbrace{\log \log \ldots \log n}_{\text{repeated } m \text{ times}}, \tag{3.29}$$

for an arbitrary finite integer $m$, can be recovered in the sense that there exists a coordinate-wise thresholding procedure that sends the FDP and NDP to zero as $n \to \infty$. In other words, DS can detect dramatically weaker signals than non-adaptive methods.

The paper is organized as follows. In Section 3.2.2 we review the conventional non-adaptive approach to sparse recovery. We introduce our adaptive sensing technique (DS) in Section 3.2.3, and in Section 3.2.4 we state our main results, that DS enables the recoverability of significantly weaker signals than standard, non-adaptive methods. Section 3.2.5 provides numerical simulations of DS, and a short discussion appears in Section 3.2.6. Proofs of the main results are given in the Appendix.

### 3.2.2 Sparse Recovery by Non-adaptive Sensing

Suppose we observe a $n \times 1$ signal $\mu$ in noise according to the model (3.28). The signal $\mu$ is assumed to be sparse—that is, most of the components of the signal are equal to zero. Define $\mathcal{S} = \{i : \mu_i \neq 0,\ i = 1, \ldots, n\}$. The elements of $\mathcal{S}$ are called the signal components, and the elements in the complement set, $\mathcal{S}^c = \{1, \ldots, n\} \setminus \mathcal{S}$, are called null components. The goal of a signal recovery procedure is to identify the signal components (in other words, estimate $\mathcal{S}$) using the observed data $X$. Let $\widehat{\mathcal{S}}(X)$ be the outcome of a given signal recovery procedure. Define the *false-discovery proportion* (FDP) to be the ratio between the number of falsely-discovered signal components and the total number of discovered components, $\text{FDP} = \left|\widehat{\mathcal{S}}(X) \setminus \mathcal{S}\right| / \left|\widehat{\mathcal{S}}(X)\right|$, and define the *non-discovery proportion* to be the ratio between the number of

undiscovered signal components and the total number of signal components, NDP $= \left|\mathcal{S} \setminus \widehat{\mathcal{S}}(X)\right| / \left|\mathcal{S}\right|$ . An effective signal recovery procedure must be able to control both the FDP and NDP.

Consider sparse signals having $n^{1-\beta}$ signal components each of amplitude $\sqrt{2r \log n}$, for some $\beta \in (0,1)$ and $r > 0$, under the model (3.28). We consider a coordinate-wise thresholding procedure,

$$\widehat{\mathcal{S}}(X) = \{i \ : \ X_i > \tau\}, \ \tau > 0, \tag{3.30}$$

to estimate the locations of the signal components. The following lemma quantifies the recoverability of such signals by coordinate-wise thresholding. The proof is given in the appendix.

**Lemma 3.7** *If $r > \beta$, the procedure* (3.30)*, with a threshold $\tau$ that depends on $r$, $\beta$ and $n$, drives both the FDP and NDP to zero with probability one as $n \to \infty$. Conversely, if $r < \beta$, then no such coordinate-wise thresholding procedure can drive the FDP and NDP to zero simultaneously with probability tending to one as $n \to \infty$.*

In other words, for the specified signal parametrization and observation model, the $(\beta, r)$ parameter plane is partitioned into two disjoint regions. In the region $r > \beta$, sparse signal components can be reliably located using a coordinate-wise thresholding procedure. In the complementary region where $r < \beta$, no coordinate-wise thresholding procedure is reliable in the sense of controlling both the FDP and NDP. This establishes a sharp boundary in the parameter space, $r = \beta$, for large-sample consistent recovery of sparse signals.

### 3.2.3 Distilled Sensing

We generalize the observation model (3.28) to allow multiple observations, indexed by $j$, of the form

$$X_i^{(j)} = \sqrt{\phi_i^{(j)}}\mu_i + Z_i^{(j)} \ , \tag{3.31}$$

for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$, where each $\phi_i^{(j)}$ is non-negative, and $Z_i^{(j)} \stackrel{iid}{\sim} \mathcal{N}(0,1)$. In addition, we impose the restriction $\sum_{i,j} \phi_i^{(j)} \leq n$, limiting the total amount of sensing energy. Note that the standard observation model (3.28) takes the form (3.31) with $k = 1$ and $\phi_i^{(1)} = 1$ for $i = 1, \ldots, n$. Another possibility is to make multiple iid observations, but each with only a fraction of the total sensing energy budget. For example, set $\phi_i^{(j)} = 1/\sqrt{k}$, $i = 1, \ldots, n$, for $j = 1, \ldots, k$. Because of the independence of the $Z_i^{(j)}$, $\sum_{j=1}^k X^{(j)}$ is equivalent to $X$ in the standard model in this case as well. There are obviously many other non-adaptive choices of $\{\phi_i^{(j)}\}_{i,j}$ that yield the same result. Furthermore, no non-adaptive sensing scheme exists that can produce better results than those obtained using observations from the standard model (3.28).

Therefore, we are interested here in adaptive, sequential designs of $\{\phi_i^{(j)}\}_{i,j}$ that tend to focus on the signal components of $\mu$. In other words, we allow $\phi_i^{(j)}$ to depend explicitly on the past $\{\phi_i^{(\ell)}, X_i^{(\ell)}\}_{i,\ell<j}$. The principle upon which our procedure is based is simple—given a collection of noisy observations of the

components of a sparse vector, it is far easier to identify a *large set* of null components (where the signal is *absent*) than it is to identify a *small set* of signal components. When multiple observations of each component are allowed, this principle suggests a process for refining observations—iteratively allocate more sensing resources to locations that are most promising while ignoring locations that are unlikely to contain signal components. This is reminiscent of the purification that occurs in the process of distillation; hence, we refer to our procedure as *distilled sensing* (DS).

Let $k$ denote the number of observation steps in the DS process, and divide the total budget of sensing energy among the steps. Each observation takes the form (3.31), where the sensing energy allocated to that observation step is distributed equally among the set of locations of interest at that step. Following each of the first $k-1$ observation steps, a refinement or *distillation* is performed, identifying the subset of locations where the corresponding observation is positive. The rationale is that it is highly improbable that the signal (which is assumed to be positive) is present at locations where the observation is negative. The algorithm terminates after the final observation, and the output consists of the final observations and the set of locations that were measured in the last step. A pseudocode description of DS appears as Algorithm 2.

To quantify the performance of DS, we will show that each distillation step retains almost all of the locations corresponding to signal components, but only about *half* of the locations corresponding to null components. When the signal $\mu$ is sparse, this implies that the effective dimension is roughly halved at each step. A judicious allocation of sensing energy over observation steps provides increasing sensing energy per location in each subsequent step, resulting in a net exponential boost in the effective amplitude of each measured signal component. As a result, applying a coordinate-wise thresholding procedure to the output observations of DS results in significant improvements in recovery compared to procedures that utilize non-adaptive sensing, as shown in the next section.

### 3.2.4 Main Results

We use an energy allocation scheme designed to balance the probabilities of successful retention of signal components at each step, by allocating a larger portion of the sensing energy to the first steps and decreasing the energy used in later steps when there are fewer locations to observe. The exponential decrease in the number of observed locations at each step suggests that the sensing energy allocated to each step can also decrease exponentially. To accomplish this we allocate energy for the first $k-1$ steps according to the entries of a geometric progression, and put all remaining energy on the last step. For a fixed parameter $0 < \Delta < 1$, the energy allocation scheme is

$$\mathcal{E}^{(j)} = \left\{ \begin{array}{ll} \frac{\Delta n}{2}\left(1-\frac{\Delta}{2}\right)^{j-1}, & j = 1,\ldots,k-1 \\ \\ n\left(1-\frac{\Delta}{2}\right)^{k-1}, & j = k \end{array} \right\}, \tag{3.32}$$

and the total energy expended satisfies $\sum_{j=1}^{k} \mathcal{E}^{(j)} = n$.

Our first main result quantifies the performance gain provided by distilled sensing (DS) when the number of observation steps is fixed. The result, stated below as a theorem, establishes an expanded region of large-sample consistent recovery in the $(\beta, r)$ parameter space. The proof is given in the Appendix.

**Theorem 3.8** *Let $k$ be a positive integer, and consider applying the $k$-step DS procedure using the sensing energy allocation strategy described in* (3.32) *with fixed parameter $0 < \Delta < 1$, to sparse signals $\mu \in \mathbb{R}^n$ having $n^{1-\beta}$ signal components each of amplitude $\sqrt{2r \log n}$, for some $\beta \in (0,1)$ and $r > 0$. If $r > \beta/(2-\Delta)^{k-1}$, there exists a coordinate-wise thresholding procedure of $X_{DS}^{(k)}$ of the form* (3.30) *that drives both the FDP and NDP to zero with probability tending to one as $n \to \infty$.*

In other words, this result shows that a small number of observation steps leads to a significant improvement in terms of recoverability—the minimum signal amplitude required for consistent recovery decreases exponentially as a function of the number of observation steps in the DS procedure. A natural question arises as to whether the number of steps can be large enough, so that signals whose amplitudes grow (with the dimension $n$) slower than $\sqrt{\log n}$ can be recovered. We address this question here by letting the number of observation steps tend to infinity slowly as a function of $n$. The result is the following theorem, for which the proof appears in the Appendix.

**Theorem 3.9** *Using the sensing energy allocation strategy described in* (3.32) *with fixed parameter $0 < \Delta < 1$, let*

$$k = 1 + \left\lceil \frac{\log \log n}{\log (2 - \Delta)} \right\rceil, \tag{3.33}$$

*and apply the $k$-step DS procedure to sparse signals $\mu \in \mathbb{R}^n$ having $n^{1-\beta}$ signal components each of amplitude $\alpha = \alpha(n)$. If $\alpha(n)$ is any positive monotone diverging sequence in $n$ that exceeds the $m$-fold iterated logarithm $\log^{[m]} n$ for some finite integer $m$, then there exists a coordinate-wise thresholding procedure of $X_{DS}^{(k)}$ of the form* (3.30) *that drives both the FDP and NDP to zero with probability tending to one as $n \to \infty$.*

In other words, DS can result in dramatic improvements in recoverability, as it succeeds at recovering signals whose amplitudes are vanishingly small relative to those of signals that can be recovered using the best non-adaptive methods.

### 3.2.5 Experimental Evaluation

Our theoretical analysis establishes that DS is considerably more powerful than conventional non-adaptive sensing in the large-sample regime. In this section we examine the finite-sample performance of DS with a simulation experiment motivated by astronomical surveying [67]. For recovery, we apply (3.30), using the

data-dependent threshold identified by the Benjamini and Hochberg procedure at a specified FDR level, where FDR is defined as the expected value of the FDP [68]. We call this recovery BH thresholding.

Fig. 3.2(a) depicts a portion of a real radio telescope image collected by the Phoenix Deep Survey (`www.physics.usyd.edu.au/~ahopkins/phoenix/`). The image size is $256 \times 256$ pixels and 533 pixels have nonzero amplitudes of 2.98 (implying $\beta = 0.43$ and $r = 0.4$). Fig. 3.2(b) depicts a simulated noisy version of the image, equivalent to a collection of non-adaptive observations from the model (3.28), where to improve visualization, locations whose corresponding observations were negative are mapped to black (amplitudes of positive observations are unaltered). Fig. 3.2(c) shows the result of applying BH thresholding at FDR level 0.05 to the non-adaptive observations, and Fig. 3.2(d) shows the output of DS with $\Delta = 0.9$ and $k = 5$ after BH thresholding at the same level (we chose $\Delta$ here to be conservative—in practice, larger values of $\Delta$ allocate more sensing energy to earlier steps, resulting in fewer total non-discoveries). Note that considerably more "stars" are recovered by DS.

For the same experiment, we also examine the benefit of DS in terms of false and non-discovery proportions. For each method we computed the empirical NDP for a range of FDR levels. The curves in Fig. 3.3 show empirical NDP vs. empirical FDP for 10 trials of each procedure (solid lines for non-adaptive sampling, dashed lines for DS with $\Delta = 0.9$ and $k = 5$). For the same FDP, DS yields lower NDPs than non-adaptive sampling, except sometimes at very high FDP levels, quantifying the improvement that is observed when visually comparing Fig. 3.2(c-d).

### 3.2.6 Discussion

A fundamental difference between non-adaptive sensing and DS can be understood by comparing false and non-discovery criteria. Recovery procedures based on non-adaptive sampling methods must control the FDP and NDP simultaneously, while each step of DS only controls the the number of non-discoveries (keeping it near zero), and allows the number of false discoveries to remain large (nearly all discoveries are false when the signal is sparse). Simultaneous FDP and NDP control for DS is performed only after the last observation step, when sensing resources have been efficiently focused into the signal subspace.

An alternate way to evaluate DS is to compare the minimal sensing energy budget required to achieve the same large-sample performance as non-adaptive sensing methods (recoverability only when $r > \beta$). The results obtained here imply that procedures based on DS will be able to recover the same signals as non-adaptive sensing methods using an energy budget that grows only *sublinearly* with $n$, implying that DS can recover signals using less sensing energy (or in less time) than what is required by non-adaptive sensing methods.

While the theoretical results presented here are asymptotic in nature, the performance of DS can also be quantified in finite-dimensional settings using the intermediate result (Theorem 3.12) in the Appendix.

Rather than exhibiting sharp asymptotics, the efficacy of DS in finite-dimensional problems is quantified by probabilities of success that vary depending on $r$, $\beta$, $\Delta$, and $n$. In addition, in finite-dimensional applications, DS could be modified to improve the retention of true signal components, at the expense of rejecting fewer null components, by selecting a less aggressive (slightly negative) threshold at each step of the DS procedure.

The proposed DS procedure can also be applied to more general classes of signals, such as those for which $\mu$ has both positive and negative values. In this case, one approach would be to split the budget of sensing energy in half, and execute the DS procedure once assuming the signal components are positive as described above, and again assuming the signal components are negative (retaining locations at each step for which the corresponding observation is negative). The final (composite) set of observations could then be subjected to standard FDR controlling procedures.

Finally, we note that the results presented here (namely Theorem 3.12) also imply that DS followed by a recovery procedure of the form (3.30), but where selection of the threshold does not require prior knowledge of the signal amplitude or sparsity parameters, will recover signals that are potentially much more sparse than those described above. Specifically, signals exhibiting general sublinear sparsity having $s(n)$ signal components each with amplitude at least $\alpha(n)$, where $s(n)$ and $\alpha(n)$ are each positive monotone diverging sequences in $n$ and $\alpha(n)$ exceeds some finite iteration of the logarithm function, such that $\alpha(n) \cdot s(n) > c \log\log\log n$ for some $c$ depending on the energy allocation parameter $\Delta$, are recoverable using the adaptive DS procedure. In the interest of space, we relegate a thorough exposition of this to future work.

## 3.2.7 Appendix

### 3.2.7.1 Proof of Lemma 3.7

Consider observing a $n \times 1$ signal $\mu$ in noise according to the model (3.28), where $\mu$ has $n^{1-\beta}$ signal components each of amplitude $\sqrt{2r \log n}$, for some $\beta \in (0,1)$ and $r > 0$. We establish the FDP and NDP criteria directly by quantifying how many false and non-discoveries result from the coordinate-wise thresholding support estimation procedure, which produces the estimate $\widehat{\mathcal{S}}(X) = \{i \ : \ X_i > \tau\}$, for $\tau = \alpha\sqrt{2r \log n}$, for some $0 < \alpha < 1$.

Recall that the false discovery proportion (FDP) was defined as the ratio between the number of falsely-discovered signal components (indices in the set $\widehat{\mathcal{S}}$ corresponding to null components) and the total number of discoveries (the total number of indices in $\widehat{\mathcal{S}}$). For the sake of exposition, let $D_z$ denote the number of indices in $\widehat{\mathcal{S}}$ corresponding to null components, and $D_s$ be the number of indices in $\widehat{\mathcal{S}}$ corresponding to signal components. Then,

$$\text{FDP} = \frac{D_z}{D_z + D_s}. \tag{3.34}$$

Similarly, the NDP is defined as the ratio between the number of falsely-rejected (missed) signal components (number of indices corresponding to signal components that are not in $\widehat{\mathcal{S}}$) and the total number of signal components. Let $M_s$ denote the number of indices not in $\widehat{\mathcal{S}}$ corresponding to signal components, then

$$\text{NDP} = \frac{M_s}{n^{1-\beta}}. \tag{3.35}$$

To establish the lemma, we identify conditions under which FDP and NDP tend to zero. First, note that NDP tends to zero if and only if $M_s/n^{1-\beta}$ tends to zero. In addition, FDP tends to zero if $D_z$ tends to zero (and $D_s$ does not), but also if $D_z$ does not tend to zero but $D_z/D_s$ does, since in this latter case the FDP can be expressed as

$$\text{FDP} = \frac{D_z}{D_z + D_s} = \frac{1}{1 + D_s/D_z}, \tag{3.36}$$

and the denominator tends to infinity when $D_z/D_s$ tends to zero, driving the FDP to zero. Conversely, in any case, if $D_z/D_s$ does not tend to zero than neither does FDP.

In what follows we will make use of standard Gaussian tail bounds—namely, for $\gamma > 0$,

$$\Pr\left(\mathcal{N}(0,1) > \gamma\right) \leq \frac{1}{\sqrt{2\pi}\gamma} \exp\left(-\gamma^2/2\right), \tag{3.37}$$

and, for $\gamma > 1$,

$$\frac{1}{\sqrt{2\pi}\gamma}\left(1 - \frac{1}{\gamma^2}\right) \exp\left(-\gamma^2/2\right) \leq \Pr\left(\mathcal{N}(0,1) > \gamma\right). \tag{3.38}$$

We will also employ Chebyshev's inequality which states that for a random variable $V$ with mean $\mathbb{E}[V]$ and variance $\sigma_V^2$, for $t > 0$,

$$\Pr\left(|V - \mathbb{E}[V]| > t\right) \leq \frac{\sigma_V^2}{t^2}, \tag{3.39}$$

which, of course, implies the one-sided bound

$$\Pr\left(V > \mathbb{E}[V] + t\right) \leq \frac{\sigma_V^2}{t^2}. \tag{3.40}$$

We quantify the NDP first by counting the number of falsely rejected signal components. Let $p_{m,s}$ denote the probability of missing a signal component, an event that occurs when the observation of a signal component does not exceed the threshold. In other words,

$$\begin{aligned} p_{m,s} &= \Pr\left(\mathcal{N}(\sqrt{2r\log n}, 1) < \alpha\sqrt{2r\log n}\right) \\ &= \Pr\left(\mathcal{N}(0,1) > (1-\alpha)\sqrt{2r\log n}\right). \end{aligned} \tag{3.41}$$

Employing the upper and lower Gaussian tail bounds, we have that for $n \geq \exp\left(r^{-1}(1-\alpha)^{-2}\right)$, the quantity $p_{m,s}$ can be expressed as

$$p_{m,s} = \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}}\; n^{-r(1-\alpha)^2}, \tag{3.42}$$

for some $c = c(n, \alpha, r) \in [1/2, 1]$.

Notice that $M_s$ is a Binomial$(n^{1-\beta}, p_{m,s})$ random variable. Applying Chebyshev's inequality, we have for $t > 0$,

$$\begin{aligned} \Pr\left(\text{NDP} > p_{m,s} + \frac{t}{n^{1-\beta}}\right) &= \Pr\left(M_s > n^{1-\beta} p_{m,s} + t\right) \\ &\leq \frac{n^{1-\beta} p_{m,s}(1 - p_{m,s})}{t^2}. \end{aligned} \tag{3.43}$$

Let $t = n^b$ for some $b > 0$ to be specified shortly. Then, for $n \geq \exp\left(r^{-1}(1-\alpha)^{-2}\right)$,

$$\begin{aligned} &\Pr\left(\text{NDP} > \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}}\, n^{-r(1-\alpha)^2} + n^{\beta-1+b}\right) \\ &\leq n^{1-\beta-2b-r(1-\alpha)^2} \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}} \left(1 - \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}}\, n^{-r(1-\alpha)^2}\right). \end{aligned} \tag{3.44}$$

Notice that when $\beta - 1 + b < 0$, or $b < 1 - \beta$, both terms in the upper bound on NDP tend to zero as $n \to \infty$. in addition, the upper bound on the probability also tends to zero if $1 - \beta - r(1-\alpha)^2 - 2b \leq 0$, or $b > (1-\beta)/2 - r(1-\alpha)^2/2$. Both of these requirements on $b$ are satisfied by the choice $b = (1-\beta)/2$. Making this substitution, we have

$$\begin{aligned} &\Pr\left(\text{NDP} > \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}}\, n^{-r(1-\alpha)^2} + n^{-(1-\beta)/2}\right) \\ &\leq n^{-r(1-\alpha)^2} \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}} \left(1 - \frac{c}{2(1-\alpha)\sqrt{\pi r \log n}}\, n^{-r(1-\alpha)^2}\right), \end{aligned} \tag{3.45}$$

which guarantees that NDP tends to zero with probability tending to one as $n \to \infty$ for any $r > 0$.

Next we quantify the FDP, beginning with the analysis of the quantity $D_z/D_s$. For this, we consider simultaneous two-sided bounds on the quantities $D_z$ and $D_s$, as follows. First, as above, we let $p_{d,z}$ denote the probability that the observation of any one of the null components exceeds the threshold $\tau$ (and is discovered); that is,

$$p_{d,z} = \Pr\left(\mathcal{N}(0,1) > \alpha\sqrt{2r \log n}\right). \tag{3.46}$$

Employing the Gaussian tail bounds, we can simplify under the assumption that $n > \exp\left(r^{-1}\alpha^{-2}\right)$, to obtain

$$p_{d,z} = \frac{c}{2\alpha\sqrt{\pi r \log n}}\, n^{-\alpha^2 r}, \tag{3.47}$$

for some $c = c(n, \alpha, r) \in [1/2, 1]$. Similarly, let $p_{d,s}$ denote the probability that the observation of any one of the signal components exceeds the threshold $\tau$ and is discovered. Specifically,

$$\begin{aligned} p_{d,s} &= \Pr\left(\mathcal{N}(\sqrt{2r \log n}, 1) > \alpha\sqrt{2r \log n}\right) \\ &= 1 - \Pr\left(\mathcal{N}(0,1) > (1-\alpha)\sqrt{2r \log n}\right). \end{aligned} \tag{3.48}$$

Now, if $n > \exp\left(r^{-1}\pi^{-1}(1-\alpha)^{-2}\right)$, then

$$\begin{aligned}\Pr\left(\mathcal{N}(0,1) > (1-\alpha)\sqrt{2r\log n}\right) &\leq \frac{1}{2(1-\alpha)\sqrt{\pi r\log n}}\, n^{-(1-\alpha)^2 r} \\ &\leq \frac{1}{2(1-\alpha)\sqrt{\pi r\log n}} \\ &< \frac{1}{2}, \end{aligned} \tag{3.49}$$

and thus we have $p_{d,s} = c'$ for some $c' = c'(n, r, \alpha) \in [1/2, 1]$.

Now, we apply the two-sided form of Chebyshev's inequality to $D_z$ to obtain, for $t > 0$,

$$\Pr\left(n(1-n^{-\beta})p_{d,z} - t < D_z < n(1-n^{-\beta})p_{d,z} + t\right) \leq \frac{n(1-n^{-\beta})p_{d,z}(1-p_{d,z})}{t^2}. \tag{3.50}$$

Let $t = n(1-n^{-\beta})p_{d,z}/2$ and assume $n > \exp\left(r^{-1}\alpha^{-2}\right)$, then

$$\begin{aligned}&\Pr\left(\frac{c(1-n^{-\beta})}{4\alpha\sqrt{\pi r\log n}} n^{1-\alpha^2 r} < D_z < \frac{3c(1-n^{-\beta})}{4\alpha\sqrt{\pi r\log n}} n^{1-\alpha^2 r}\right) \\ &\qquad \leq \frac{8\alpha\sqrt{\pi r\log n}}{cn^{1-\alpha^2 r}}\left(\frac{1 - cn^{-\alpha^2 r}/(2\alpha\sqrt{\pi r\log n})}{1-n^{-\beta}}\right). \end{aligned} \tag{3.51}$$

Similarly, for $D_s$, we have, for $t' > 0$,

$$\Pr\left(n^{1-\beta}p_{d,s} - t' < D_z < n^{1-\beta}p_{d,s} + t'\right) \leq \frac{n^{1-\beta}p_{d,s}(1-p_{d,s})}{(t')^2}. \tag{3.52}$$

Let $t' = n^{1-\beta}p_{d,s}/2$ and assume $n > \exp\left(r^{-1}\pi^{-1}(1-\alpha)^{-2}\right)$ to obtain

$$\Pr\left(\frac{c'}{2}n^{1-\beta} < D_s < \frac{3c'}{2}n^{1-\beta}\right) \leq \frac{4(1-c')}{c'n^{1-\beta}}. \tag{3.53}$$

Combining these results, we have that with probability at least

$$1 - \frac{8\alpha\sqrt{\pi r\log n}}{cn^{1-\alpha^2 r}}\left(\frac{1 - cn^{-\alpha^2 r}/(2\alpha\sqrt{\pi r\log n})}{1-n^{-\beta}}\right) - \frac{4(1-c')}{c'n^{1-\beta}}, \tag{3.54}$$

for $n > \max\{\exp\left(r^{-1}\alpha^{-2}\right), \exp\left(r^{-1}\pi^{-1}(1-\alpha)^{-2}\right)\}$, the quantity of interest satisfies

$$\frac{c(1-n^{-\beta})}{6c'\alpha\sqrt{\pi r\log n}}\, n^{\beta-\alpha^2 r} < \frac{D_z}{D_s} < \frac{3c(1-n^{-\beta})}{2c'\alpha\sqrt{\pi r\log n}} n^{\beta-\alpha^2 r}. \tag{3.55}$$

Since, by assumption, $r < 1/\alpha^2$, we have that FDP tends to zero with probability tending to one as $n \to \infty$ if $r \geq \beta/\alpha^2$, and conversely FDP is bounded away from zero with probability tending to one as $n \to \infty$ if $r < \beta/\alpha^2$.

Our last part of the analysis addresses the FDP behavior in the regime $r \geq 1/\alpha^2$ by identifying conditions under which $D_z$ tends to zero (and $D_s$ does not). The one-sided form of Chebyshev's inequality guarantees that for $t > 0$,

$$\Pr\left(D_z > n\left(1-n^{-\beta}\right)p_{d,z} + t\right) \leq \frac{n\left(1-n^{-\beta}\right)p_{d,z}(1-p_{d,z})}{t^2}. \tag{3.56}$$

Let

$$t = \left(n\left(1-n^{-\beta}\right)p_{d,z}(1-p_{d,z})\right)^{1/2}\ n^{b/2}(\log\log n)^{1/4} \tag{3.57}$$

for some $b \geq 0$ (to be specified) to obtain

$$\begin{aligned}\Pr\Bigg(D_z > &\frac{c}{2\alpha\sqrt{\pi r\log n}}\ n^{1-\alpha^2 r}(1-n^{-\beta}) \\ &+ \sqrt{n^{1+b-\alpha^2 r}\frac{(1-n^{-\beta})c}{2\alpha}\left(\frac{\log\log n}{\pi r\log n}\right)^{1/2}\left(1-\frac{c}{2\alpha(\pi r\log n)^{1/2}}\ n^{-\alpha^2 r}\right)}\Bigg) \leq \frac{n^{-b}}{\sqrt{\log\log n}},\end{aligned} \tag{3.58}$$

for $n > \exp\left(r^{-1}\alpha^{-2}\right)$. Now, assume $\alpha^2 r \geq 1$ and set $b = (\alpha^2 r - 1)/4$, and simplify to obtain

$$\begin{aligned}\Pr\Bigg(D_z > &\frac{c}{2\alpha\sqrt{\pi r\log n}}\ n^{1-\alpha^2 r}(1-n^{-\beta}) \\ &+ n^{3(1-\alpha^2 r)/8}\sqrt{\frac{(1-n^{-\beta})c}{2\alpha}\ \left(\frac{\log\log n}{\pi r\log n}\right)^{1/2}\left(1-\frac{c}{2\alpha(\pi r\log n)^{1/2}}\ n^{-\alpha^2 r}\right)}\Bigg) \leq \frac{n^{(1-\alpha^2 r)/4}}{\sqrt{\log\log n}},\end{aligned} \tag{3.59}$$

for $n > \exp\left(r^{-1}\alpha^{-2}\right)$.

Combining this result with (3.53), we have that for $n \geq \max\{\exp\left(r^{-1}\alpha^{-2}\right), \exp\left(r^{-1}\pi^{-1}(1-\alpha)^{-2}\right)\}$, with probability at least

$$1 - \frac{n^{(1-\alpha^2 r)/4}}{\sqrt{\log\log n}} - \frac{4(1-c')}{c' n^{1-\beta}}, \tag{3.60}$$

$D_s$ is bounded away from zero ($D_s > c' n^{1-\beta}/2$) and

$$\begin{aligned}D_z < &\frac{c}{2\alpha\sqrt{\pi r\log n}}\ n^{1-\alpha^2 r}(1-n^{-\beta}) \\ &+ n^{3(1-\alpha^2 r)/8}\sqrt{\frac{(1-n^{-\beta})c}{2\alpha}\ \left(\frac{\log\log n}{\pi r\log n}\right)^{1/2}\left(1-\frac{c}{2\alpha(\pi r\log n)^{1/2}}\ n^{-\alpha^2 r}\right)}.\end{aligned} \tag{3.61}$$

This implies that $D_z$ tends to zero (and $D_s$ does not) with probability tending to one as $n \to \infty$ for $r \geq 1/\alpha^2$.

Putting it all together, we have shown that as $n \to \infty$, NDP tends to zero with probability tending to one for any $r > 0$. On the other hand, the FDP shows a phase transition. As $n \to \infty$, FDP tends to zero with probability tending to one if $r > \beta/\alpha^2$, while FDP remains bounded away from zero with probability tending to one if $r < \beta/\alpha^2$. By assumption, we have $\alpha \in (0,1)$, and so the most favorable thresholds are obtained when $\alpha \approx 1$, giving rise to the $r = \beta$ threshold claimed in the lemma.

#### 3.2.7.2 Proofs of Theorems 3.8 and 3.9

Establishing the main results of the paper amounts to counting how many locations corresponding to signal and null components are retained by thresholding the observations at level zero in each distillation step. We begin by considering the signal components.

**Lemma 3.10** *Consider a vector $\mu$ with $s$ components, each of amplitude $\alpha > 0$, observed according to the model $X_i = \sqrt{\phi}\mu_i + Z_i$, where $Z_i$ is a collection of independent $\mathcal{N}(0,1)$ noises and $\phi > 0$ denotes the amount of sensing energy allocated to each location. When the amplitude satisfies $\alpha \geq 2/\sqrt{\phi}$, the number of components retained by thresholding the observations at the level zero, denoted $\widetilde{s}$, satisfies $(1-\epsilon)\, s \leq \widetilde{s} \leq s$ with probability at least*

$$1 - \exp\left(-\frac{\alpha \cdot s}{4}\sqrt{\frac{\phi}{2\pi}}\right), \tag{3.62}$$

*where*

$$\epsilon = \frac{1}{\alpha}\sqrt{\frac{1}{2\pi\phi}} \leq \frac{1}{2\sqrt{2\pi}}. \tag{3.63}$$

*The upper bound on $\epsilon$ follows from the condition on $\alpha$, and ensures that the fraction of vector components retained is bounded away from zero.*

**Proof:** The proof amounts to counting the number of components retained by thresholding. To each component observation $X_i$, assign a Bernoulli random variable $T_i = \mathbf{I}_{\{X_i > 0\}}$, which is equal to one whenever the observation exceeds 0, and zero otherwise. Let $p = \Pr(T_i = 1)$. Using the Gaussian tail bound given above, we have that the number of vector components whose observations exceed the threshold, $\widetilde{s} = \sum_i T_i$, is a Binomial random variable, with

$$1 - p \leq \frac{1}{\alpha}\sqrt{\frac{1}{2\pi\phi}}\exp\left(-\frac{\alpha^2\phi}{2}\right). \tag{3.64}$$

Establishing the lemma amounts to quantifying the probability that $\widetilde{s} \geq (1-\epsilon)s$ for an appropriately chosen $\epsilon$. To that end, we use a bound on the tail probability of the Binomial distribution [63]. Namely, for a Binomial$(n,p)$ random variable $B$, whenever $b < \mathbb{E}[B] = np$,

$$\Pr\left(B \leq b\right) \leq \left(\frac{n-np}{n-b}\right)^{n-b}\left(\frac{np}{b}\right)^{b}. \tag{3.65}$$

In our context, this result implies

$$\Pr\left(\widetilde{s} \leq (1-\epsilon)s\right) \leq \left(\frac{1-p}{\epsilon}\right)^{\epsilon s}\left(\frac{p}{1-\epsilon}\right)^{(1-\epsilon)s}, \tag{3.66}$$

provided $\epsilon > 1 - p$, which is satisfied by the choice

$$\epsilon = \frac{1}{\alpha}\sqrt{\frac{1}{2\pi\phi}}. \tag{3.67}$$

Now, notice that when $\alpha > 2/\sqrt{\phi}$, the condition

$$-\frac{\alpha^2\phi\epsilon s}{2} + (1-\epsilon)s\log\left(\frac{1}{1-\epsilon}\right) \leq -\frac{\alpha s}{4}\sqrt{\frac{\phi}{2\pi}}, \tag{3.68}$$

obtained by upper-bounding the logarithm of the right-hand side of (3.66), holds for any $\epsilon \in (0,1)$. The result follows from exponentiating this last bound. ∎

Next, we quantify how many of the null components are retained by each thresholding step.

**Lemma 3.11** *Consider a vector $\mu$ with $z$ components, each of amplitude 0, observed according to the model $X_i = \sqrt{\phi}\mu_i + Z_i$, where $Z_i$ is a collection of independent $\mathcal{N}(0,1)$ noises and $\phi > 0$ denotes the amount of sensing energy allocated to each measured location. For any $\epsilon_0 < 1/2$, the number of components retained by thresholding the observations at the level zero, denoted $\widetilde{z}$, satisfies $(1/2 - \epsilon_0)\, z \leq \widetilde{z} \leq (1/2 + \epsilon_0)\, z$, with probability at least $1 - 2\exp\left(-2z\epsilon_0^2\right)$.*

**Proof:** To each observation $X_i$, assign a Bernoulli random variable $T_i = \mathbf{I}_{\{X_i > 0\}}$ which takes the value one when the corresponding observation exceeds 0 and zero otherwise. Since each observation is of noise only, the number of vector components whose corresponding observation exceeds the threshold, $\widetilde{z} = \sum_i T_i$, is a Binomial random variable with probability 1/2. Applying Hoeffding's inequality we obtain $\Pr\left(|\widetilde{z} - z/2| > \epsilon_0 z\right) \leq 2\exp\left(-2z\epsilon_0^2\right)$, which holds for any $\epsilon_0 > 0$. Imposing the restriction $\epsilon_0 < 1/2$ guarantees that the fraction of components retained is within $(0, 1)$. ∎

Taken together, the lemmata above establish that by thresholding at level zero, almost all of the signal components and about half of the zero components are retained with high probability. Incorporating the geometric allocation of sensing energy per step specified in (3.32), we obtain the following.

**Theorem 3.12** *Let $\mu \in \mathbb{R}^n$ be a sparse signal having $s > 0$ signal components each of amplitude $\alpha$, where $\alpha > 4/\sqrt{\Delta}$, and $z = n - s > s$ null components. In the DS procedure of Algorithm 2, let $1 < k \leq 1 + \log_2(z/s)$, let $\Delta$ be a fixed parameter satisfying $0 < \Delta < 1 - 2\epsilon_0$ for some $\epsilon_0 < 1/2$, and let the energy allocation $\mathcal{E}^{(j)}$ be as described in (3.32). For $j = 1, \ldots, k-1$, define*

$$\epsilon^{(j)} = \frac{1}{\alpha}\sqrt{\frac{1}{2\pi\xi^{(j)}}}, \tag{3.69}$$

*where*

$$\xi^{(j)} = \frac{\Delta}{4}\left(\frac{2-\Delta}{1+2\epsilon_0}\right)^{j-1}. \tag{3.70}$$

*Then, with probability at least*

$$1 - (k-1)\exp\left(-\frac{\alpha\cdot s}{8}\cdot\sqrt{\frac{\Delta}{2\pi}}\cdot\prod_{j=1}^{k-2}\left(1-\epsilon^{(j)}\right)\right) - 2(k-1)\exp\left(-2\cdot z\cdot\epsilon_0^2\cdot\left(\frac{1}{2}-\epsilon_0\right)^{k-2}\right), \tag{3.71}$$

*the output of the DS procedure, $X_{DS}^{(k)}$, is equivalent in distribution to a single collection of noisy observations of a vector $\mu_{\text{eff}} \in \mathbb{R}^{n_{\text{eff}}}$ according to the observation model (3.28). The number of signal components in $\mu_{\text{eff}}$ is denoted by $s_{\text{eff}}$ and satisfies*

$$s\prod_{j=1}^{k-1}\left(1-\epsilon^{(j)}\right) \leq s_{\text{eff}} \leq s, \tag{3.72}$$

*the effective signal length* $n_{\text{eff}}$ *satisfies*

$$s\prod_{j=1}^{k-1}\left(1-\epsilon^{(j)}\right)+z\left(\frac{1}{2}-\epsilon_0\right)^{k-1}\leq n_{\text{eff}}\leq s+z\left(\frac{1}{2}+\epsilon_0\right)^{k-1}, \tag{3.73}$$

*and the effective observed amplitude* $\alpha_{\text{eff}}$ *satisfies*

$$\alpha_{\text{eff}}\geq\alpha\sqrt{\frac{n\left(1-\Delta/2\right)^{k-1}}{s+z\left(1/2+\epsilon_0\right)^{k-1}}}. \tag{3.74}$$

**Proof:** We begin by applying the union bound to the result of Lemma 3.11 to enforce the condition for each of the first $k-1$ distillation steps. Using superscripts on $z$ and $s$ to index the observation step, such that $z^{(1)}=z$ and $s^{(1)}=s$, for $\epsilon_0<1/2$, the bounds

$$z^{(1)}\left(\frac{1}{2}-\epsilon_0\right)^{j-1}\leq z^{(j+1)}\leq z^{(1)}\left(\frac{1}{2}+\epsilon_0\right)^{j-1} \tag{3.75}$$

hold simultaneously for all $j=1,2,\ldots,k-1$ with probability exceeding

$$1-2(k-1)\exp\left(-2z^{(1)}\epsilon_0^2\left(\frac{1}{2}-\epsilon_0\right)^{k-2}\right). \tag{3.76}$$

As a result, with the same probability, the total number of locations in each set $I^{(j)}$ satisfies $|I^{(j)}|\leq s^{(1)}+z^{(1)}\left(\frac{1}{2}+\epsilon_0\right)^{j-1}$, for $j=1,2,\ldots,k$. Using these upper bounds and the energy allocation rule (3.32), we can lower bound the sensing energy per location at each step, $\phi_i^{(j)}=\mathcal{E}^{(j)}/|I^{(j)}|$, for $i\in I^{(j)}$—specifically,

$$\phi_i^{(j)}\geq\left\{\begin{array}{ll}\frac{\Delta n(1-\Delta/2)^{j-1}}{2\left(s^{(1)}+z^{(1)}(1/2+\epsilon_0)^{j-1}\right)}, & j=1,\ldots,k-1\\ \\ \frac{n(1-\Delta/2)^{k-1}}{s^{(1)}+z^{(1)}(1/2+\epsilon_0)^{k-1}}, & j=k\end{array}\right\}, \tag{3.77}$$

for $i\in I^{(j)}$ (and $\phi_i^{(j)}=0$ for $i\notin I^{(j)}$). Notice that when $k\leq 1+\log_2\left(z^{(1)}/s^{(1)}\right)$, for each $i\in I^{(j)}$,

$$\phi_i^{(j)}\geq\left\{\begin{array}{ll}\frac{\Delta}{4}\left(\frac{2-\Delta}{1+2\epsilon_0}\right)^{j-1}, & j=1,\ldots,k-1\\ \\ \frac{1}{2}\left(\frac{2-\Delta}{1+2\epsilon_0}\right)^{k-1}, & j=k\end{array}\right\}. \tag{3.78}$$

Since $\Delta<1-2\epsilon_0$, this shows that the amount of sensing energy allocated to each retained location *increases exponentially* with the number of observation steps.

Now, conditioned on the above event, we can invoke Lemma 3.10 and apply the union bound again so that with probability at least

$$1-(k-1)\exp\left(-\frac{\alpha\cdot s^{(1)}}{8}\cdot\sqrt{\frac{\Delta}{2\pi}}\cdot\prod_{j=1}^{k-2}\left(1-\epsilon^{(j)}\right)\right), \tag{3.79}$$

when $\alpha \geq 4/\sqrt{\Delta}$, the bounds

$$\left(1-\epsilon^{(j)}\right) s^{(j)} \leq s^{(j+1)} \leq s^{(j)} \tag{3.80}$$

hold simultaneously for all $j = 1, 2, \ldots, k-1$.

Applying the union bound to both composite events, we obtain that with the specified probability, the number of signal and null components present in the $k$th observation step are given by $s^{(k)}$ and $z^{(k)}$, respectively, as defined above. Thus, the final effective signal dimension is $s^{(k)} + z^{(k)}$, and the effective observed amplitude of each signal component is obtained using the lower bound on $\phi_i^{(j)}$, establishing the claim. ∎

Before we can prove the main results, we need one final lemma quantifying the (limiting) fraction of signal components retained throughout the DS procedure.

**Lemma 3.13** *Let $g = g(n)$ and $k = k(n)$ be positive monotone diverging sequences in $n$, where $g(n)$ exceeds the $m$-fold iteration $\log^{[m]} n$ for some finite integer $m$ and $k(n) \leq n$. Define $\epsilon^{(j)}(n) = c^{-j}/g(n)$ for some constant $c > 1$, and assume $\epsilon^{(k)}(n) < 1$. Then*

$$\lim_{n\to\infty} \prod_{j=1}^{k} \left(1-\epsilon^{(j)}(n)\right) = 1. \tag{3.81}$$

**Proof:** Since $c > 1$, $(1-\epsilon^{(j)}(n)) < (1-\epsilon^{(j+1)}(n))$ and thus the expression of interest satisfies

$$\left(1-\epsilon^{(1)}(n)\right)^k = \left(1-\frac{1}{g(n)c}\right)^k \leq \prod_{j=1}^{k} \left(1-\epsilon^{(j)}(n)\right). \tag{3.82}$$

If $k(n)/g(n) \to 0$ as $n \to \infty$, then the limit of the left hand side is easily seen to be 1 by a Taylor Series expansion about $\epsilon^{(1)}(n) = 0$, and the lemma is established. Suppose this is not the case. Then, notice that for any $1 \leq k' < k$,

$$\begin{aligned}
\prod_{j=1}^{k} \left(1-\epsilon^{(j)}(n)\right) &= \prod_{j=1}^{k'} \left(1-\epsilon^{(j)}(n)\right) \cdot \prod_{j=k'+1}^{k} \left(1-\epsilon^{(j)}(n)\right) \\
&\overset{(a)}{\geq} \prod_{j=1}^{k'} \left(1-\epsilon^{(j)}(n)\right) \cdot \prod_{j=k'+1}^{k} \left(1-\epsilon^{(k'+1)}(n)\right) \\
&\overset{(b)}{\geq} \prod_{j=1}^{k'} \left(1-\epsilon^{(j)}(n)\right) \cdot \left(1-\epsilon^{(k'+1)}(n)\right)^k,
\end{aligned} \tag{3.83}$$

where $(a)$ follows because $c > 1$, and $(b)$ is the result of multiplying by additional terms that are positive and less than one. Now, choosing $k'$ so that $c^{k'+1} > k^2$, say $k' = \max\{\lfloor 2\log_c k \rfloor, 0\} + 1$, is sufficient to ensure that $k\epsilon^{k'+1}(n) \to 0$, and thus $\lim_{n\to\infty}(1-\epsilon^{(k'+1)}(n))^k = 1$ by a Taylor Series argument. In this case we have

$$\lim_{n\to\infty} \prod_{j=1}^{k} \left(1-\epsilon^{(j)}(n)\right) \geq \lim_{n\to\infty} \prod_{j=1}^{k'} \left(1-\epsilon^{(j)}(n)\right). \tag{3.84}$$

Now, either $k'(n)/g(n) \to 0$ as $n \to \infty$ and the limit of the lower bound is 1 (by a Taylor Series argument), or we repeat the process above by introducing some $k'' < k'$ such that $c^{k''+1} > (k')^2$, say $k'' = \max\{\lfloor 2\log_c k' \rfloor, 0\} + 1 \sim \log_c \log_c k$. Since $g(n)$ exceeds a finite iteration of the log function and $k(n) \leq n$, this reduction process will eventually terminate in a finite number of steps, and the limit in this terminating case will be 1 by a Taylor Series argument, establishing the claim. ∎

#### 3.2.7.3 Proof of Theorem 3.8

Let $s = n^{1-\beta}$ for some $\beta \in (0,1)$, let $\alpha = \sqrt{2r\log n}$ for some $r > 0$, and let $k \in \mathbb{N}$ be a fixed integer. Choose $\epsilon_0 = n^{-1/3}$ and $\Delta < 1 - 2\epsilon_0$, and note that all of the conditions of Theorem 3.12 are satisfied when $n > \exp(8/\Delta r)$. By Lemma 3.13 we have that

$$\lim_{n\to\infty} \prod_{j=1}^{k-1} \left(1 - \epsilon^{(j)}\right) = 1, \tag{3.85}$$

which is easy to see by making the substitutions

$$g(n) = \sqrt{\frac{\pi r \Delta(1+2\epsilon_0)\log n}{2-\Delta}},\ c = \sqrt{\frac{2-\Delta}{1+2\epsilon_0}}. \tag{3.86}$$

Thus, we obtain that with probability tending to one as $n \to \infty$, $s_{\text{eff}} \to s$, $n_{\text{eff}} \to s + z \cdot 2^{-(k-1)}$, and $\alpha_{\text{eff}} \geq \sqrt{2r(2-\Delta)^{k-1}\log n} > \sqrt{2r(2-\Delta)^{k-1}\log n_{\text{eff}}}$. Leveraging the results in the non-adaptive setting, there exists a thresholding procedure of $X_{DS}^{(k)}$ of the form (3.30) that will drive the FDP and NDP to zero with probability tending to one as $n \to \infty$ whenever $r(2-\Delta)^{k-1} > \beta$, as claimed.

#### 3.2.7.4 Proof of Theorem 3.9

Let $s = n^{1-\beta}$ for some $\beta \in (0,1)$, and let $\alpha = \alpha(n)$ be any positive monotone diverging sequence in $n$ exceeding the $m$-fold iteration $\log^{[m]} n$ for an arbitrary finite integer $m$. Let the number of observation steps be

$$k = k(n) = 1 + \left\lceil \frac{\log\log n}{\log(2-\Delta)} \right\rceil. \tag{3.87}$$

Choose $\epsilon_0 = n^{-1/3}$ and $\Delta < 1 - 2\epsilon_0$. Applying Theorem 3.12 and Lemma 3.13, we obtain that with probability tending to one as $n \to \infty$, $s_{\text{eff}} \to s$,

$$n_{\text{eff}} \to s + \frac{z}{(\log n)^{\log(2)/\log(2-\Delta)}} \tag{3.88}$$

and $\alpha_{\text{eff}} \geq \alpha\sqrt{\log n} > \alpha\sqrt{\log n_{\text{eff}}}$. Since $\alpha = \alpha(n)$ diverges, for large enough $n$ the effective observed amplitude will exceed $\sqrt{2\beta\log n_{\text{eff}}}$ for any fixed $\beta$. Now, applying the non-adaptive results, there exists a thresholding procedure of $X_{DS}^{(k)}$ of the form (3.30) that will drive the FDP and NDP to zero with probability tending to one as $n \to \infty$.

### 3.2.8 Acknowledgement

The authors wish to thank Jiashun Jin for discussing the details of his work with them.

**Algorithm 3.2**: Distilled sensing.

**Input**:

Number of observation steps $k$

Energy allocation strategy: $\mathcal{E}^{(j)}$, $\sum_{j=1}^{k} \mathcal{E}^{(j)} \leq n$

**Initialize:**

Initial index set $I^{(1)} \longleftarrow \{1, 2, \ldots, n\}$

**Distillation:**

**for** $j = 1$ **to** $k$ **do**

$$X_i^{(j)} = \left\{ \begin{array}{ll} \sqrt{\frac{\mathcal{E}^{(j)}}{|I^{(j)}|}}\mu_i + Z_i^{(j)}, & i \in I^{(j)} \\ Z_i^{(j)}, & i \in I^{(1)} \setminus I^{(j)} \end{array} \right\}$$

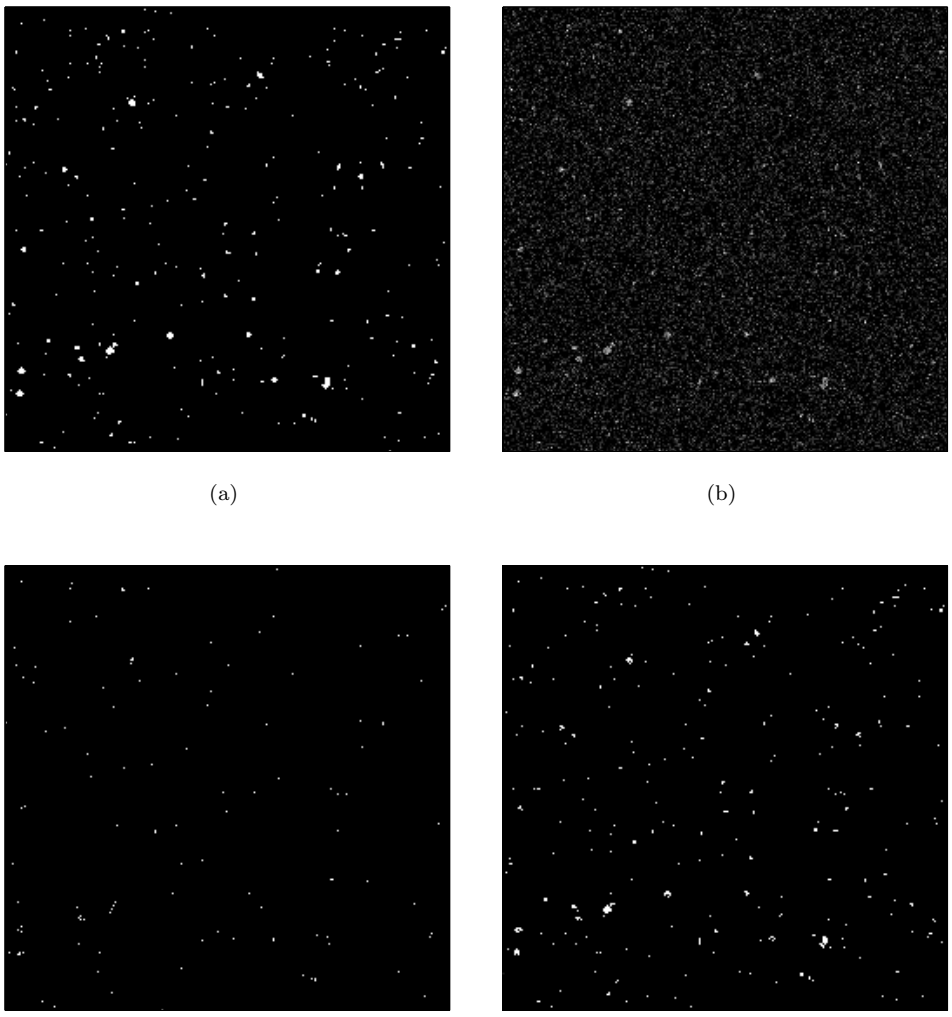
$$I^{(j+1)} \longleftarrow \{i \in I^{(j)} : X_i^{(j)} > 0\}$$

**end**

**Output**:

Final index set $I^{(k)}$

Distilled observations $X_{\mathrm{DS}}^{(k)} := \{X_i^{(k)} : i \in I^{(k)}\}$



(a) (b)

(c) (d)

Figure 3.2 Experimental validation of DS procedure. Panels (a) and (b) show a noiseless radio telescope image and a simulated noisy version of the image, respectively. Panel (c) shows the result of applying BH thresholding (at FDR level 0.05) to the data in (b), while applying BH thresholding at the same level to the output of the DS procedure with $\Delta = 0.9$ and $k = 5$ gives the result depicted in panel (d).

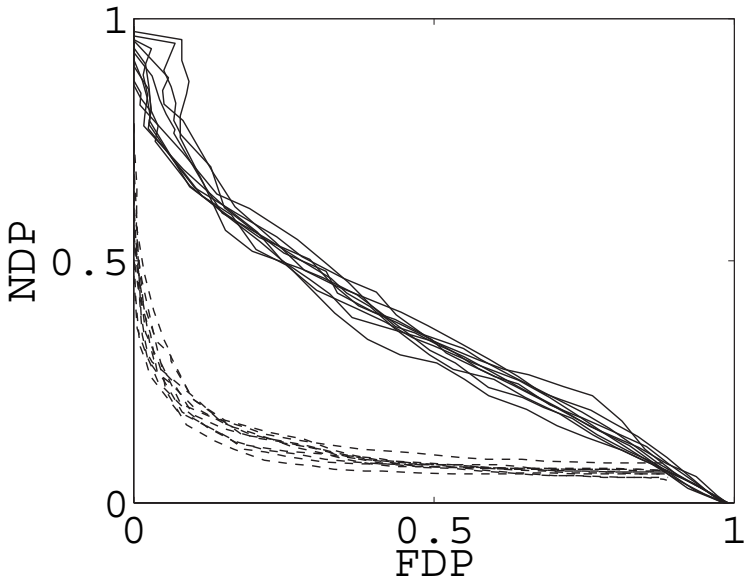



Figure 3.3 NDP vs. FDP for 10 independent trials of each procedure applied to the noisy star recovery task whose one-trial results are depicted graphically in Fig. 3.2. Curves for non-adaptive sampling followed by BH thresholding are solid lines, while curves for DS ($\Delta = 0.9$, $k = 5$) followed by BH thresholding are dashed lines. The separation of curve clusters illustrates the improvement of DS per trial and on average.

# Chapter 4

# Directions for Future Study

## Sequential Distilled Sensing

The results in Chapter 3 follow a general template: *if* the unknown signal of interest is sufficiently sparse, *then* the proposed adaptive sampling procedure (distilled sensing, or DS) outperforms traditional non-adaptive sampling methods. The salient feature of the DS procedure is that current observations are evaluated to determine the most promising locations for future measurements, while less-promising locations are ignored from that point on. Provided the sparsity assumption is valid, adaptive sampling permits the recovery (detection, or estimation of sparsity pattern) of signals for which comparable non-adaptive methods would fail. On the other hand, the results in Chapter 3 provide no specific guarantees if the adaptive sampling procedure is utilized to acquire signals that are not sufficiently sparse. A natural question arises—can the DS procedure be made universal, in the sense that the same dramatic improvements are obtained if the sparsity assumption is valid, but there is little or no performance degradation (relative to non-adaptive methods) if not?

Answering this question in the affirmative would entail the specification and analysis of a sequential distilled sensing (SDS) procedure. The primary contrast between DS and SDS would be that the SDS procedure would use current observations first to evaluate whether refinement should be performed, and if so, to determine what are the most promising locations for future measurements. In this sense, the number of refinement steps in SDS would also be adaptive to unknown sparsity level, and this additional flexibility would inevitably make it a very attractive option in a variety of practical applications.

## Compressive Distilled Sensing (Adaptive Compressive Sensing)

The bulk of the work in Chapter 2 leveraged the formal notion of restricted isometry (through the use of the restricted isometry property, or RIP). More qualitatively, the notion of *incoherence* can also provide intuition for understanding when projections of a signal onto a collection of test vectors are sufficient to allow recovery of the signal. In essence, incoherent projections (such as random projections) deposit an equal fraction of sensing energy in every dimension of the signal space. While this spreading of sensing

energy is the enabling property of CS when the observations are noise-free, in noisy settings the signal to noise ratio (SNR) *per dimension* is necessarily much lower using this approach than if one had used projections with all sensing power concentrated on a single coordinate (i.e., traditional point sampling).

If an oracle were to provide the coordinates of the sparse signal components a priori, then the optimal samples would be projections directly on to those coordinates, in effect maximizing the SNR per sample. Without an oracle this is not possible, but it is intuitively clear that an adaptive sampling and estimation procedure, where the projections are adjusted according to previous observations in order to maximize the SNR per sample, could be tremendously helpful.

The potential advantages of adaptive projection shaping schemes for sparse signal discovery were empirically demonstrated in recent works. The first published approach was the fully-Bayesian procedure proposed in [74]. While this procedure was shown to be effective in practice, the computational complexity of the projection shaping procedure limits its applications to problems of relatively small dimension. A more computationally efficient (myopic) Bayesian algorithm was proposed and evaluated in [75]. Absent from these prior works was a full theoretical evaluation of the performance of the algorithms, which was prohibitive due to the complicated statistical dependencies among the observations.

An interesting direction for future study is to extend the distilled sensing (DS) framework to CS observation models. Ideally, a hybrid procedure of CS and DS would share the benefits of each: recovery of a sparse signal from a relatively small number of observations (the salient feature of CS), along with the analytical framework of DS which would enable theoretical analysis of adaptive compressive sensing procedures. An initial investigation of such a scheme was performed in [76], where some very preliminary theoretical results were obtained. Extending this work to quantify the performance improvements (in signal detection, reconstruction, sparsity pattern recovery, etc.) resulting from adaptivity would be a significant contribution to the general theory of sparse recovery.

# REFERENCES


[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[4] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4036–4048, Sep. 2006.

[5] J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to channel sensing," *submitted manuscript*, Aug. 2008.

[6] J. Haupt, R. Castro, and R. Nowak, "Adaptive discovery of sparse signals in noise," in *Proceedings of the 42nd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2008, pp. 1727–1731.

[7] ——, "Distilled sensing: Selective sampling for sparse signal recovery," in *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*, Clearwater Beach, FL, Apr. 2009, pp. 216–223.

[8] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, Atlanta, GA, May 1996, pp. 1688–1691.

[9] M. Gastpar and Y. Bresler, "On the necessary density for spectrum-blind nonuniform sampling subject to quantization," in *Proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, Istanbul, Jun. 2000, pp. 348–351.

[10] A. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss, "Near-optimal sparse Fourier representations via sampling," in *Proceedings of the ACM Symposium on Theory of Computing (STOC)*, 2002, pp. 152–161.

[11] J. Zou, A. Gilbert, M. Strauss, and I. Daubechies, "Theoretical and experimental analysis of a randomized algorithm for sparse Fourier Transform analysis," *Journal of Computational Physics*, vol. 211, no. 2, pp. 572–595, Jan. 2006.

[12] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[13] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2788–2805, Aug. 2005.

[14] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[15] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[16] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when $p$ is much larger than $n$," *Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, Dec. 2007.

[17] A. Barron, "Complexity regularization with application to artificial neural networks," in *Nonparametric Functional Estimation and Related Topics*. Kluwer Academic Publishers, 1991, pp. 561–576.

[18] C. Craig, "On the Tchebychef inequality of Bernstein," *Annals of Statistics*, vol. 4, no. 2, pp. 94–102, May 1933.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[20] M. Figueiredo and R. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 906–916, Aug. 2003.

[21] I. Daubechies, M. Defrise, and C. DeMol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.

[22] M. Figueiredo and R. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *Proceedings of the 2005 International Conference on Image Processing*, vol. 2, Sep. 2005, pp. 782–785.

[23] D. Donoho, "For most large underdetermined systems of linear equations, the minimal l1-norm near-solution approximates the sparsest near-solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 7, pp. 907–934, Jul. 2006.

[24] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," *Lecture Notes in Computer Science*, vol. 2634, pp. 162–177, 2003.

[25] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[26] W. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor networks," in *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks*, Apr. 2005, pp. 332–339.

[27] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *Proceedings of the 5th International Symposium on Information Processing in Sensor Networks*, Nashville, TN, Apr. 2006.

[28] M. Rabbat, J. Haupt, A. Singh, and R. Nowak, "Decentralized compression and predistribution via randomized gossip," in *Proceedings of the 5th International Symposium on Information Processing in Sensor Networks*, Nashville, TN, Apr. 2006.

[29] A. Bonami, "Étude des coefficients de Fourier des fonctions de $L^p(g)$," *Annales de l'institut Fourier*, vol. 20, no. 2, pp. 335–402, 1970.

[30] E. Candès, "The restricted isometry property and its implications for compressed sensing," in *Compte Rendus de l'Academie des Sciences, Serie I*, Paris, 2008, vol. 346, pp. 589–592.

[31] W. Bajwa, J. Haupt, G. Raz, and R. Nowak, "Compressed channel sensing," in *Proceedings of the 42nd Annual Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 2008, pp. 5–10.

[32] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," in *Constructive Approximation*. New York: Springer, 2008.

[33] W. Bajwa, J. Haupt, G. Raz, S. Wright, and R. Nowak, "Toeplitz-structured compressed sensing matrices," in *Proceedings of the IEEE/SP 14th Workshop on Statistical Signal Processing*, Madison, WI, Aug. 2007, pp. 294–298.

[34] D. Kilfoyle and A. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 4–27, Jan. 2000.

[35] *Receiver Performance Guidelines*, ATSC Recommended Practices for Digital Television, 2004. [Online]. Available: http://www.atsc.org/standards/practices.html

[36] A. Molisch, "Ultrawideband propagation channels-Theory, measurement, and modeling," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1528–1545, Sep. 2005.

[37] J. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, 2001.

[38] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993.

[39] D. Donoho, M. Vetterli, R. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2435–2476, Oct. 1998.

[40] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing — Special Issue on Convex Optimization Methods for Signal Processing*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[41] S. Wright, R. Nowak, and M. Figueiredo, "Sparse reconstruction by separable approximation," in *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, NV, Apr. 2008, pp. 3373–3376.

[42] R. S. Varga, *Geršgorin and His Circles*. Berlin, Germany: Springer-Verlag, 2004, no. 36.

[43] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *Annals of Statistics*, vol. 28, no. 5, pp. 1302–1338, Oct. 2000.

[44] D. West, *Introduction to Graph Theory*. Upper Saddle River, NJ: Prentice Hall, 2000.

[45] A. Hajnal and E. Szemerédi, "Proof of a conjecture of P. Erdős," in *Combinatorial Theory and its Application*, P. Erdős, A. Rényi, and V. T. Sós, Eds., North-Holland, Amsterdam, 1970, pp. 601–623.

[46] S. Pemmaraju, "Equitable coloring extends Chernoff-Hoeffding bounds," *Lecture Notes in Computer Science*, vol. 2129, pp. 285–296, 2001.

[47] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France, May 2006, pp. 872–875.

[48] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Transactions on Signal Processing*, 2007, to appear.

[49] G. Pfander, H. Rauhut, and J. Tanner, "Identification of matrices having a sparse representation," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5376–5388, 2007.

[50] J. Romberg, "Compressive sensing by random convolution," *SIAM Journal of Imaging Science*, 2008, to appear.

[51] R. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, pp. 918–925, August 2007.

[52] D. Donoho and M. Elad, "Optimally sparse represetations in general (nonorthogonal) dictionaries via $\ell^1$ minimization," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 2197–2202, March 2003.

[53] J. Tropp, "On the conditioning of random subdictionaries," *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 1–24, July 2008.

[54] T. Strohmer and R. Heath, "Grassmanian frames with applications to coding and communication," *Applied and Computational Harmonic Analysis*, vol. 14, no. 3, pp. 257–275, May 2003.

[55] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, May 2009.

[56] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Communications on Pure and Applied Mathematics*, vol. 61, no. 8, pp. 1025–1045, Aug. 2008.

[57] G. Fudge, M. Chivers, S. Ravindran, R. Bland, P. Pace, and J. Haupt, "A Nyquist folding analog-to-information receiver," in *Proceedings of the 42nd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2008.

[58] H. Weyl, "Über die Gleichverteilung von Zahlen mod. Eins," *Mathematische Annalen*, vol. 77, pp. 312–352, 1916.

[59] N. Korobov, *Exponential sums and their applications*. Kluwer Academic Publishers, 1992.

[60] I. Vinogradov, *Selected Works*. Berlin-New York: Springer-Verlag, 1985.

[61] D. Donoho and J. Jin, "Higher criticsm for detecting sparse heterogenous mixtures," *Annals of Statistics*, vol. 32, no. 3, pp. 962–994, 2004.

[62] J. Jin, "Detecting a target in very noisy data from multiple looks," *Institute of Mathematical Statistics Lecture Notes*, vol. 45, pp. 255–286, 2004.

[63] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Annals of Statistics*, vol. 23, pp. 493–507, 1952.

[64] S. Verdu, *Multiuser Detection*. New York: Cambridge University Press, 1998.

[65] C. Genovese, N. Lazar, and T. Nichols, "Thresholding of statistical maps in functional neuroimaging using the false discovery rate," *NeuroImage*, vol. 15, pp. 870–878, 2002.

[66] Y. Pawitan, S. Michiels, S. Koscielny, A. Gusnanto, and A. Ploner, "False discovery rate, sensitivity and sample size for microarray studies," *Bioinformatics*, vol. 21, pp. 3017–3024, 2005.

[67] A. Hopkins, C. Miller, A. Connolly, C. Genovese, R. Nichol, and L. Wasserman, "A new source detection algorithm using the false-discovery rate," *Astronomical Journal*, vol. 123, pp. 1086–1094, 2002.

[68] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society: Series B*, vol. 57, pp. 289–300, 1995.

[69] F. Abramovich, Y. Benjamini, D. Donoho, and I. Johnstone, "Adapting to unknown sparsity by controlling the false discovery rate," *Annals of Statistics*, vol. 34, no. 2, pp. 584–653, 2006.

[70] D. Donoho and J. Jin, "Asymptotic minimaxity of false discovery rate thresholding for sparse exponential data," *Annals of Statistics*, vol. 34, no. 6, pp. 2980–3018, 2006.

[71] ——, "Feature selection by higher criticism thresholding: Optimal phase diagram," *submitted manuscript*, 2008.

[72] J. Jin, "Detecting and estimating sparse mixtures," Ph.D. dissertation, Department of Statistics, Stanford University, 2003.

[73] R. Rangarajan, R. Raich, and A. Hero, "Optimal sequential energy allocation for inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 67–78, 2007.

[74] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[75] R. Castro, J. Haupt, R. Nowak, and G. Raz, "Finding needles in noisy haystacks," in *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, NV, Apr. 2008, pp. 5133–5136.

[76] J. Haupt, R. Castro, and R. Nowak, "Adaptive sensing for sparse signal recovery," in *Proceedings of the 2009 IEEE Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop*, Marco Island, FL, Jan. 2009.