Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors*

Laurent Jacques[†], Jason N. Laska[‡], Petros T. Boufounos[§], and Richard G. Baraniuk[‡]

April 15, 2011

Abstract

The Compressive Sensing (CS) framework aims to ease the burden on analog-to-digital converters (ADCs) by reducing the sampling rate required to acquire and stably recover sparse signals. Practical ADCs not only sample but also quantize each measurement to a finite number of bits; moreover, there is an inverse relationship between the achievable sampling rate and the bit depth. In this paper, we investigate an alternative CS approach that shifts the emphasis from the sampling rate to the number of bits per measurement. In particular, we explore the extreme case of 1-bit CS measurements, which capture just their sign. Our results come in two flavors. First, we consider ideal reconstruction from noiseless 1-bit measurements and provide a lower bound on the best achievable reconstruction error. We also demonstrate that a large class of measurement mappings achieve this optimal bound. Second, we consider reconstruction robustness to measurement errors and noise and introduce the *Binary ϵ-Stable Embedding* (BϵSE) property, which characterizes the robustness measurement process to sign changes. We show the same class of matrices that provide optimal noiseless performance also enable such a robust mapping. On the practical side, we introduce the *Binary Iterative Hard Thresholding* (BIHT) algorithm for signal reconstruction from 1-bit measurements that offers state-of-the-art performance.

1 Introduction

Recent advances in signal acquisition theory have led to significant interest in alternative sampling methods. Specifically, conventional sampling systems rely on the Shannon sampling theorem that states that signals must be sampled uniformly at the Nyquist rate, i.e., a rate twice their bandwidth.

*L. J. is funded by the Belgian Science Policy, "Return Grant", BELSPO, IAP-VI BCRYPT. J. L. and R. B. were supported by the grants NSF CCF-0431150, CCF-0728867, CCF-0926127, CNS-0435425, and CNS-0520280, DARPA/ONR N66001-08-1-2065, N66001-11-1-4090, ONR N00014-07-1-0936, N00014-08-1-1067, N00014-08-1-1112, and N00014-08-1-1066, AFOSR FA9550-07-1-0301 and FA9550-09-1-0432, ARO MURI W911NF-07-1-0185 and W911NF-09-1-0383, and the Texas Instruments Leadership University Program. P.B. is funded by Mitsubishi Electric Research Laboratories.

[†]ICTEAM Institute, ELEN Department, Université catholique de Louvain (UCL), B-1348 Louvain-la-Neuve, Belgium. Email: laurent.jacques@uclouvain.be.

[‡]Department of Electrical and Computer Engineering, Rice University, Houston, TX, 70015 USA. Email: laska@rice.edu, richb@rice.edu.

[§]Mitsubishi Electric Research Laboratories (MERL) e-mail: petrosb@merl.com.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
15 APR 2011		**00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rice University,Department Electrical and Computer Engineering,Houston,TX,77005**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
submitted

14. ABSTRACT

The Compressive Sensing (CS) framework aims to ease the burden on analog-to-digital converters (ADCs) by reducing the sampling rate required to acquire and stably recover sparse signals. Practical ADCs not only sample but also quantize each measurement to a nite number of bits; moreover, there is an inverse relationship between the achievable sampling rate and the bit depth. In this paper, we investigate an alternative CS approach that shifts the emphasis from the sampling rate to the number of bits per measurement. In particular, we explore the extreme case of 1-bit CS measurements, which capture just their sign. Our results come in two avors. First, we consider ideal reconstruction from noiseless 1-bit measurements and provide a lower bound on the best achievable reconstruction error. We also demonstrate that a large class of measurement mappings achieve this optimal bound. Second, we consider reconstruction robustness to measurement errors and noise and introduce the Binary -Stable Embedding (B SE) property, which characterizes the robustness measurement process to sign changes. We show the same class of matrices that provide optimal noiseless performance also enable such a robust mapping. On the practical side, we introduce the Binary Iterative Hard Thresholding (BIHT) algorithm for signal reconstruction from 1-bit measurements that o ers state-of-the-art performance.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**37**	

However, the *compressive sensing* (CS) framework describes how to reconstruct a signal $\boldsymbol{x} \in \mathbb{R}^N$ from the linear measurements

$$\boldsymbol{y} = \Phi \boldsymbol{x}, \tag{1}$$

where $\Phi \in \mathbb{R}^{M \times N}$ with $M < N$ is an underdetermined measurement system [1, 2]. It is possible to design a physical sampling system $\bar{\Phi}$ such that $\boldsymbol{y} = \Phi \boldsymbol{x} = \bar{\Phi}(x(t))$ where $\boldsymbol{x}$ is a vector of Nyquist-rate samples of a bandlimited signal $x(t)$, $t \in \mathbb{R}$. In this case, (1) translates to low, sub-Nyquist sampling rates, providing the framework's axial significance: CS enables the acquisition and accurate reconstruction of signals that were previously out of reach, limited by hardware sampling rates [3] or number of sensors [4].

Although inversion of (1) seems ill-posed, it has been demonstrated that K-sparse signals, i.e., $\boldsymbol{x} \in \Sigma_K$ where $\Sigma_K := \{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_0 := |\mathrm{supp}(\boldsymbol{x})| \leq K\}$, can be reconstructed exactly [1, 2]. To do this, we could naïvely solve for the sparsest signal that satisfies (1),

$$\boldsymbol{x}^* \quad = \quad \underset{\boldsymbol{x} \in \mathbb{R}^N}{\mathrm{argmin}} \, \|\boldsymbol{x}\|_0 \quad \text{s.t.} \quad \boldsymbol{y} = \Phi \boldsymbol{x}; \tag{$\mathrm{R_{CS}}$}$$

however, this non-convex program exhibits combinatorial complexity in the size of the problem [5]. Instead, we solve *Basis Pursuit* (BP) by relaxing the objective in ($\mathrm{R_{CS}}$) to the ℓ_1-norm; the result is a convex, polynomial-time algorithm [6]. A key realization is that, under certain conditions on Φ, the BP solution will be equivalent to that of ($\mathrm{R_{CS}}$) [1]. This basic reconstruction framework has been expanded to include numerous fast algorithms as well as provably robust algorithms for reconstruction from noisy measurements [7–11]. Reconstruction can also be performed with iterative and greedy methods [12–14].

Reconstruction guarantees for BP and other algorithms are often demonstrated for Φ that are endowed with the *restricted isometry property* (RIP), the sufficient condition that the norm of the measurements is close to the norm of the signal for all sparse $\boldsymbol{x}$ [7].[1] This can be expressed, in general terms, as a *δ-stable embedding*. Let $\delta \in (0, 1)$ and $X, S \subset \mathbb{R}^N$. We say the mapping Φ is a *δ-stable embedding* of X, S if

$$(1 - \delta)\|\boldsymbol{x} - \boldsymbol{s}\|_2^2 \leq \|\Phi \boldsymbol{x} - \Phi \boldsymbol{s}\|_2^2 \leq (1 + \delta)\|\boldsymbol{x} - \boldsymbol{s}\|_2^2, \tag{2}$$

for all $\boldsymbol{x} \in X$ and $\boldsymbol{s} \in S$. The RIP requires that (2) hold for all $\boldsymbol{x}, \boldsymbol{s} \in \Sigma_K$; that is, it is a stable embedding of sparse vectors. A key result in the CS literature is that, if the coefficients of Φ are randomly drawn from a sub-Gaussian distribution, then Φ will satisfy the RIP with high probability as long as $M \geq C_\delta K \log(N/K)$, for some constant C_δ [15, 16]. Several hardware inspired designs with only a few randomized components have also been shown to satisfy this property [3, 17–19].

In practice, CS measurements must be quantized, i.e., each measurement is mapped from a real value (over a potentially infinite range) to a discrete value over some finite range. For example, in uniform quantization, a measurement is mapped to one of 2^B distinct values, where B denotes the number of bits per measurement. Quantization is an irreversible process that introduces error in the measurements. One way to account for quantization error is to treat it as bounded noise and employ robust reconstruction algorithms. Alternatively, we might try to reduce the error by choosing the most efficient quantizer for the distribution of the measurements. Several reconstruction techniques that specifically address CS quantization have also been proposed [20–25].

[1] The RIP is in fact not needed to demonstrate exact reconstruction guarantees in noiseless settings, however it proves quite useful for establishing robust reconstruction guarantees in noise.

While quantization error is a minor inconvenience, fine quantization invokes a more burdensome, yet often overlooked source of adversity: in hardware systems, it is the primary bottleneck limiting sample rates [26, 27]. In other words, the ADC is beholden to the quantizer. First, quantization significantly limits the maximum speed of the analog-to-digital converter (ADC), forcing an exponential decrease in sampling rate as the number of bits is increased linearly [27]. Second, the quantizer is the primary power consumer in an ADC. Thus, more bits per measurement directly translates to slower sampling rates and increased ADC costs. Third, fine quantization is more susceptible to non-linear distortion in the ADC electronics, requiring explicit treatment in the reconstruction [28]. As we have seen, the CS framework provides one mechanism to alleviate the quantization bottleneck by reducing the ADC sampling rate. Is it possible to extend the CS framework to mitigate this problem directly in the quantization domain by reducing the number of bits per measurement (bit-depth) instead?

In this paper we concretely answer this question in the affirmative. We consider an extreme quantization; just one bit per CS measurement, representing its sign. The quantizer is thus reduced to a simple comparator that tests for values above or below zero, enabling extremely simple, efficient, and fast quantization. A 1-bit quantizer is also more robust to a number of commonly encountered non-linear distortions in the input electronics, as long as they preserve the sign of the measurements.

It is not obvious that the signs of the CS measurements retain enough information for signal reconstruction; for example, it is immediately clear that the scale (absolute amplitude) of the signal is lost. Nonetheless, there is strong empirical evidence that signal reconstruction is possible [28–31]. In this paper we develop strong theoretical reconstruction and robustness guarantees, in the same spirit as neoclassical guarantees provided in CS by the RIP.

We briefly describe the 1-bit CS framework proposed in [29]. Measurements of a signal $\boldsymbol{x} \in \mathbb{R}^N$ are computed via

$$\boldsymbol{y}_s = A(\boldsymbol{x}) := \operatorname{sign}(\Phi \boldsymbol{x}). \tag{3}$$

Thus, the measurement operator $A(\cdot)$ is a mapping from $\mathbb{R}^N$ to the Boolean cube[2] $\mathcal{B}^M := \{-1, 1\}^M$. At best, we hope to recover signals $\boldsymbol{x} \in \Sigma_K^* := \{\boldsymbol{x} \in S^{N-1} : \|\boldsymbol{x}\|_0 \leq K\}$ where $S^{N-1} := \{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_2 = 1\}$ is the unit hyper-sphere of dimension N. We restrict our attention to sparse signals on the unit sphere since, as previously mentioned, the scale of the signal has been lost during the quantization process. To reconstruct, we enforce *consistency* on the signs of the estimate's measurements, i.e., that $A(\boldsymbol{x}^*) = A(\boldsymbol{x})$. Specifically, we define a general non-linear reconstruction algorithm $\Delta^{\text{1bit}}(\boldsymbol{y}_s, \Phi, K)$ such that, for $\boldsymbol{x}^* = \Delta^{\text{1bit}}(\boldsymbol{y}_s, \Phi, K)$, the solution $\boldsymbol{x}^*$ is

(i) sparse, i.e., satisfies $\|\boldsymbol{x}^*\|_0 \leq K = \|\boldsymbol{x}\|_0$; and

(ii) consistent, i.e., satisfies $A(\boldsymbol{x}^*) = \boldsymbol{y}_s = A(\boldsymbol{x})$.

With ($\mathrm{R_{CS}}$) from CS as a guide, one candidate program for reconstruction is of course

$$\boldsymbol{x}^* \quad = \quad \operatorname*{argmin}_{\boldsymbol{x} \in S^{N-1}} \|\boldsymbol{x}\|_0 \quad \text{s.t.} \quad \boldsymbol{y}_s = \operatorname{sign}(\Phi \boldsymbol{x}). \tag{$\mathrm{R_{1BCS}}$}$$

Although the parameter K is not explicit in ($\mathrm{R_{1BCS}}$), the property *(i)* above holds because $\boldsymbol{x}$ is a feasible point of the constraint.

[2]Generally, the m-dimensional Boolean cube is defined as $\{0, 1\}^M$. Without loss of generality, we use $\{-1, 1\}^M$ instead.

Since (R_{1BCS}) is computationally intractable, [29] proposes a relaxation that replaces the objective with the ℓ_1-norm and enforces consistency via a linear convex constraint. However, the resulting program remains non-convex due to the the unit-sphere requirement. Be that as it may, several optimization algorithms have been developed for the relaxation, as well as a greedy algorithm inspired by the same ideas [29–31]. While previous empirical results from these algorithms provide motivation for the validity of this 1-bit framework, there have been few analytical guarantees to date.

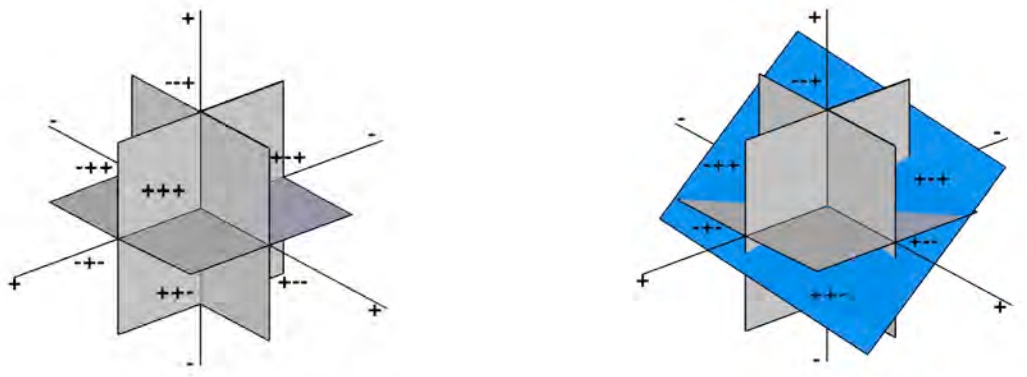
The primary contribution of this paper is a rigorous analysis of the 1-bit CS framework. We provide two flavors of results. First, we determine a lower bound on reconstruction performance from all possible mappings A with the reconstruction algorithm $\Delta^{1\text{bit}}$, i.e., the best achievable performance of this 1-bit CS framework. We further demonstrate that if the elements of Φ are drawn randomly from Gaussian distribution or its rows are drawn uniformly from the unit sphere, then the solution to $\Delta^{1\text{bit}}$ will have bounded error on the order of the optimal lower bound. Second, we provide conditions on A that enable us to characterize the reconstruction performance even when some of the measurement signs have changed (e.g., due to noise in the measurements). In other words, we derive the conditions under which robust reconstruction from 1-bit measurements can be achieved. We do so by demonstrating that A is a stable embedding of sparse signals, similar to the RIP. We apply these stable embedding results to the cases where we have noisy measurements and signals that are not strictly sparse. Our guarantees demonstrate that the 1-bit CS framework is on sound footing and provide a first step toward analysis of the relaxed 1-bit techniques used in practice.

To develop robust reconstruction guarantees, we propose a new tool, the *binary ϵ-stable embedding* (BϵSE), to characterize 1-bit CS systems. The BϵSE implies that the normalized angle between any sparse vectors in S^{N-1} is close to the normalized Hamming distance between their 1-bit measurements. We demonstrate that the same class of random A as above exhibit this property when $M \geq C_\epsilon K \log N$ (where C_ϵ is some constant). Thus remarkably, there exist A such that the BϵSE holds when both the number of measurements M is smaller than the dimension of the signal N and the measurement bit-depth is at minimum.

As a complement to our theoretical analysis, we introduce a new 1-bit CS reconstruction algorithm, *Binary Iterative Hard Thresholding* (BIHT). Via simulations, we demonstrate that BIHT yields a significant improvement in both reconstruction error as well as consistency, as compared with previous algorithms. To gain intuition about the behavior of BIHT, we explore the way that this algorithm enforces consistency and compare and contrast it with previous approaches. Perhaps more important than the algorithm itself is the discovery that the BIHT consistency formulation provides a significantly better feasible solution in noiseless settings, as compared with previous algorithms. Finally, we provide a brief explanation regarding why this new formulation achieves better solutions, and its connection with results in the machine learning literature.

In addition to benchmarking the performance of BIHT, our simulations demonstrate that many of the theoretical predictions that arise from our analysis (such as the error rate as a function of the number of measurements or the error rate as a function of measurement Hamming distance), are actually exhibited in practice. This implies that our theoretical analysis is accurately explaining the true behavior of the framework.

The remainder of this paper is organized as follows. In Section II, we develop performance results for 1-bit CS in the noiseless setting. Specifically we develop a lower bound on reconstruction



(a) (b)

Figure 1: (a) The 8 orthants in $\mathbb{R}^3$. (b) Intersection of orthants by a 2-dimensional subspace. At most 6 of the 8 available orthants are intersected.

performance as well as provide the guarantee that Gaussian matrices enable this performance. In Section III we introduce the notion of a BϵSE for the mapping A and demonstrate that Gaussian matrices facilitate this property. We also expand reconstruction guarantees for measurements with Gaussian noise (prior to quantization) and non-sparse signals. To make use of these results in practice, in Section IV we present the BIHT algorithm for practical 1-bit reconstruction. In Section V we provide simulations of BIHT to verify our claims. In Section VI we conclude with a discussion about implications and future extensions. To facilitate the flow and the description, most of our proofs are provided in the appendices.

2 Noiseless Reconstruction Performance

2.1 Reconstruction performance lower bounds

In this section, we seek to provide guarantees on the reconstruction error from 1-bit CS measurements. Before analyzing this performance from a specific mapping A with the consistent sparse reconstruction algorithm $\Delta^{1\text{bit}}(\boldsymbol{y}_s, \Phi, K)$, it is instructive to determine the best achievable performance from measurements acquired using any mapping. Thus, in this section we seek a lower bound on the reconstruction error.

We develop the lower bound on the reconstruction error based on how well the quantizer exploits the available measurement bits. A distinction we make in this section is that of measurement bits, which is the number of bits acquired by the measurement system, versus information bits, which represent the true amount of information carried in the measurement bits. Our analysis follows similar ideas to that in [32, 33], adapted to sign measurements.

We first examine how 1-bit quantization operates on the measurements. Specifically, we consider the orthants of the measurement space. An orthant in $\mathbb{R}^M$ is the set of vectors such that all the vector's coefficients have the same sign pattern

$$\mathcal{O}_{\boldsymbol{s}} = \{\boldsymbol{x} \mid \operatorname{sign} \boldsymbol{x} = \boldsymbol{s}\},$$

where $\boldsymbol{s}$ is a vector of ± 1. Any M-dimensional space is partitioned to 2^M orthants. Figure 1(a) shows the 8 orthants of $\mathbb{R}^3$ as an example. Since 1-bit quantization only preserves the signs of the measurements, it encodes in which measurement space orthant the measurements lie. Thus, each available quantization point corresponds to an orthant in the measurement space. Any unquantized measurement vector $\Phi\boldsymbol{x}$ that lies in an orthant of the measurement space will quantize to the corresponding *quantization point* of that orthant and cannot be distinguished from any other measurement vector in the same orthant. To obtain a lower bound on the reconstruction error, we begin by bounding the number of quantization points (or equivalently the number of orthants) that are used to encode the signal.

While there are generally 2^M orthants in the measurement space, the space formed by measuring all sparse signals occupies a small subset of the available orthants. We determine the number of available orthants that can be intersected by the measurements in the following lemma:

Lemma 1. *Let $\boldsymbol{x} \in \mathcal{S} := \bigcup_{i=1}^{L} \mathcal{S}_i$ belong to a union of L subspaces $\mathcal{S}_i \subset \mathbb{R}^N$ of dimension K, and let M 1-bit measurements $\boldsymbol{y}_s$ be acquired via the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ as defined in (3). Then the measurements $\boldsymbol{y}_s$ can effectively use at most $L\binom{M}{K}2^K$ quantization points, i.e., carry at most $K \log_2(eLM/K)$ information bits.*

Proof. A K-dimensional subspace in an M dimensional space cannot lie in all the 2^M available octants. For example, as shown in Fig. 1(b), a 2-dimensional subspace of a 3-dimensional space can intersect at most 6 of the available octants. In Appendix A, we demonstrate that one arbitrary K-dimensional subspace in an M-dimensional space intersects at most $\binom{M}{K}2^K$ orthants of the 2^M available. Since Φ is a linear operator, any K-dimensional subspace $\mathcal{S}_i$ in the signal space $\mathbb{R}^N$ is mapped through Φ to a subspace $\mathcal{S}_i' = \Phi\mathcal{S}_i \subset \mathbb{R}^M$ that is also at most K-dimensional and therefore follows the same bound. Thus, if the signal of interest belongs in a union $\mathcal{S} := \bigcup_{i=1}^{L} \mathcal{S}_i$ of L such K-dimensional subspaces, then $\Phi\boldsymbol{x} \in \mathcal{S}' := \bigcup_{i=1}^{L} \mathcal{S}_i'$, and it follows that at most $L\binom{M}{K}2^K$ orthants are intersected. This means that at most $L\binom{M}{K}2^K$ effective quantization points can be used, i.e., at most $K \log_2(eLM/K)$ information bits can be obtained. □

Since K-sparse signals in any basis $\Psi \in \mathbb{R}^{N\times N}$ belong to a union of at most $\binom{N}{K}$ subspaces in $\mathbb{R}^N$, using Lemma 1 we can obtain the following corollary.[3]

Corollary 1. *Let $\boldsymbol{x} = \Psi\boldsymbol{\alpha} \in \mathbb{R}^N$ be K-sparse in a certain basis $\Psi \in \mathbb{R}^{N\times N}$, i.e., $\boldsymbol{\alpha} \in \Sigma_K$. Then the measurements $\boldsymbol{y}_s = A(\boldsymbol{x})$ can effectively use at most $\binom{N}{K}\binom{M}{K}2^K$ 1-bit quantization points, i.e, carry at most $2K \log_2(e\sqrt{NM}/K)$ information bits.*

The set of signals of interest to be encoded is the set of unit-norm K-sparse signals Σ_K^*. Since unit-norm signals of a K-dimensional subspace form a K-dimensional unit sphere in that subspace, Σ_K^* is a union of $\binom{N}{K}$ such unit spheres. The $Q = \binom{N}{K}\binom{M}{K}2^K$ available quantization points partition Σ_K^* into Q smaller sets, each of which contains all the signals that quantize the same point.

To develop the lower bound on the reconstruction error we examine the optimal such partition, with respect to the worst-case error, given the number of quantization points used. The measurement and reconstruction process maps each signal in Σ_K^* to a finite set of quantized signals $\mathcal{Q} \subset \Sigma_K^*, |\mathcal{Q}| = Q$. At best this map ensures that the worst case reconstruction error is minimized,

[3]This corollary is easily adaptable to a redundant basis $\Psi \in \mathbb{R}^{N\times D}$ with $D \geq N$.

i.e.,

$$\epsilon_{\text{opt}} \;=\; \max_{\boldsymbol{x}\in\Sigma_K^*} \min_{\boldsymbol{q}\in\mathcal{Q}} \|\boldsymbol{x}-\boldsymbol{q}\|_2, \tag{4}$$

where ϵ_{opt} denotes the worst-case quantization error and $\boldsymbol{q}$ each of the available quantization points. The optimal lower bound is achieved by designing $\mathcal{Q}$ to minimize (4) without considering whether the measurement and reconstruction process actually achieve this design. Thus, designing the set $\mathcal{Q}$ becomes a set covering problem.

Using this intuition and Lemma 1, Appendix A proves the following statement about a set of unit-norm signals in a union of L, K-dimensional subspaces, specifically $\boldsymbol{x} \in \Sigma_K^*$.

Theorem 1. *Let the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ and measurements $\boldsymbol{y}_s$ be defined as in (3) and let $\boldsymbol{x} \in \Sigma_K^*$. Then the estimate from the reconstruction algorithm $\Delta^{\text{1bit}}(\boldsymbol{y}_s, \Phi, K)$ has error defined by (4) of at least*

$$\epsilon_{\text{opt}} \geq \frac{K}{2eM} = \Omega\left(\frac{K}{M}\right).$$

Thus, the worst-case error cannot decay at a rate faster than $\Omega(1/M)$ as a function of the number measurements, no matter what reconstruction algorithm is used. The bound in the theorem is independent of L, but similarly to the relation between Lemma 1 and Corollary 1, K-sparse signals are a special case with $L = \binom{N}{K}$.

This result assumes noiseless acquisition and provides no guarantees of robustness and noise resiliency. This is in line with existing results on scalar quantization in oversampled representations and CS that state that the distortion due to scalar quantization of noiseless measurements cannot decrease faster than the inverse of the measurement rate [32–36]. To improve the rate vs. distortion trade-off, alternative quantization methods must be used, such as Sigma-Delta quantization [37–43] or non-monotonic scalar quantization [44].

Theorem 1 bounds the best possible performance of a consistent reconstruction over all possible mappings A. However, it is straightforward to construct mappings A that do not behave as the lower bound suggests. In the next section we identify one class of matrices such that the mapping A admits an almost optimal upper bound on the reconstruction error from a general algorithm Δ^{1bit}.

2.2 Achievable performance via random projections

In this section we describe a class of matrices Φ such that the consistent sparse reconstruction algorithm $\Delta^{\text{1bit}}(\boldsymbol{y}_s, \Phi, K)$ can indeed achieve error decay rates of optimal order, described by Theorem 1, with the number of measurements growing linear in the sparsity K and logarithmically in the dimension N, as is required in conventional CS. We first focus our analysis on Gaussian matrices, i.e., Φ such that each element $\phi_{i,j}$ is randomly drawn i.i.d. from the standard Gaussian distribution, $\mathcal{N}(0,1)$. In the rest of the paper, we use the short notation $\Phi \sim \mathcal{N}^{M\times N}(0,1)$ for characterizing such matrices, and we write $\boldsymbol{\varphi} \sim \mathcal{N}^{N\times 1}(0,1)$ for describing equivalent random vectors in $\mathbb{R}^N$ (e.g., the rows of Φ). For these matrices Φ, we prove the following in Appendix A.

Theorem 2. *Let Φ be matrix generated as $\Phi \sim \mathcal{N}^{M\times N}(0,1)$, and let the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ be defined as in (3). Fix $0 \leq \eta \leq 1$ and $\epsilon_o > 0$. If the number of measurements is*

$$M \geq \tfrac{1}{\epsilon_o}\left(2K\,\log(N) + 4K\log(\tfrac{16}{\epsilon_o}) + \log\tfrac{1}{\eta}\right), \tag{5}$$

then for all $\boldsymbol{x}, \boldsymbol{s} \in \Sigma_K^*$ *we have that*

$$\|\boldsymbol{x} - \boldsymbol{s}\|_2 > \epsilon_o \;\Rightarrow\; A(\boldsymbol{x}) \neq A(\boldsymbol{s}), \tag{6}$$

or equivalently

$$A(\boldsymbol{x}) = A(\boldsymbol{s}) \;\Rightarrow\; \|\boldsymbol{x} - \boldsymbol{s}\|_2 \leq \epsilon_o,$$

with probability higher than $1 - \eta$.

The Theorem demonstrates that if we use Gaussian matrices in the mapping A, then, given a fixed probability level η, the reconstruction algorithm $\Delta^{1\text{bit}}(\boldsymbol{y}_s, \Phi, K)$ will recover signals with optimal error order

$$\epsilon_o \;=\; O\Big(\big(\tfrac{K}{M}\big)^{1-\alpha} \log N\Big),$$

for arbitrarily small $\alpha > 0$; the presence of the $\log(1/\epsilon_0)$ term in (5) prevents us from setting $\alpha = 0$.

A similar result has been very recently shown for sign measurements of non-sparse signals in the context of quantization using frame permutations [45]. Specifically, it has been shown that reconstruction from sign measurements of signals can be achieved (almost surely) with a $O((1/M)^{1-\alpha})$ error rate decay for arbitrarily small $\alpha > 0$. Our main contribution here is extending this result to K-sparse vectors in $\mathbb{R}^N$. Our results, in addition to introducing the almost linear dependence on K, also show that if the signal is sparse then we pay a logarithmic penalty in N. This is consistent with results in CS, but seems not to be necessary from the lower bound in the previous section. We will see in Section V that for Gaussian matrices, the optimal error behavior is empirically exhibited on average. Finally, we note that for a constant ϵ_0, the number of measurements required to guarantee (6) is $M = O(K \log N/K)$, nearly the same as order in conventional CS.

We note a few minor extensions of the Theorem. We can multiply the rows of Φ with a positive scalar without changing the signs of the measurements. By normalizing the rows of the Gaussian matrix, we obtain a matrix with rows drawn uniformly from the unit ℓ_2 sphere in $\mathbb{R}^N$. It is thus straightforward to extend the Theorem to such matrices with such rows as well. Furthermore, note that these projections are "universal," meaning that the theorem remains valid for sparse signals in Ψ, i.e., for $\boldsymbol{x}, \boldsymbol{s}$ belonging to $\Sigma_{\Psi,K}^* := \{\boldsymbol{u} = \Psi\boldsymbol{\alpha} \in \mathbb{R}^N : \boldsymbol{\alpha} \in \Sigma_K^*\}$. This is true since for any orthonormal basis $\Psi \in \mathbb{R}^{N\times N}$, $\Phi' = \Phi\Psi \sim \mathcal{N}^{M\times N}(0,1)$ when $\Phi \sim \mathcal{N}^{M\times N}(0,1)$.

We can also view the binary measurements as a *hash* or a *sketch* of the signal. With this interpretation of the result we guarantee with high probability that no sparse vectors with Euclidean distance greater than ϵ_o will "hash" to the same binary measurements. In fact, similar results play a key role in *locality sensitive hashing* (LSH), a technique that aims to efficiently perform approximate nearest neighbors searches from quantized projections [46–49]. Most LSH results examine the performance on point-clouds of a discrete number of signals instead of the infinite subspaces that we explore in this paper. Furthermore, the primary goal of the LSH is to preserve the structure of the nearest neighbors with high probability. Instead, in this paper we are concerned with the ability to reconstruct the signal from the hash, as well as the robustness of this reconstruction to measurement noise and signal model mismatch. To enable these properties, we require a property of the mapping A that preserves the structure (geometry) of the entire signal set. Thus, in the next section we seek an embedding property of A that preserves geometry for the set of sparse signals and thus ensures robust reconstruction.

3 Acquisition and Reconstruction Robustness

3.1 Binary ϵ-stable embeddings

In this section we establish an embedding property for the 1-bit CS mapping A that ensures that the sparse signal geometry is preserved in the measurements, analogous to the RIP for real-valued measurements. This robustness property enables us to upper bound the reconstruction performance even when some measurement signs have been changed due to noise. Conventional CS achieves robustness via the δ-stable embeddings of sparse vectors (2) discussed in Section I. This embedding is a restricted *quasi-isometry* between the metric spaces $(\mathbb{R}^N, d_X)$ and $(\mathbb{R}^M, d_Y)$, where the distance metrics d_X and d_Y are the ℓ_2-norm in dimensions N and M, respectively, and the domain is restricted to sparse signals.[4] We seek a similar definition for our embedding; however, now the signals and measurements lie in the different spaces S^{N-1} and $\mathcal{B}^M$, respectively. Thus, we first consider appropriate distance metrics in these spaces.

The Hamming distance is the natural distance for counting the number of unequal bits between two measurement vectors. Specifically, for $\boldsymbol{y}, \boldsymbol{v} \in \mathcal{B}^M$ we define the *normalized Hamming distance* as

$$d_H(\boldsymbol{y}, \boldsymbol{v}) = \frac{1}{M} \sum_{i=1}^{M} y_i \oplus v_i,$$

where $\oplus$ is the XOR operation such that $a \oplus b$ equals 0 if $a = b$ and 1 otherwise. The distance is normalized such that $d_H \in [0, 1]$. In the signal space we only consider unit-norm vectors, thus, a natural distance is the angle formed by any two of these vectors. Specifically, for $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$, we consider

$$d_S(\boldsymbol{x}, \boldsymbol{s}) := \frac{1}{\pi} \arccos\langle \boldsymbol{x}, \boldsymbol{s} \rangle.$$

As with the Hamming distance, we normalize the true angle $\arccos\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ such that $d_S \in [0, 1]$. Note that since both vectors have the same norm, the inner product $\langle \boldsymbol{x}, \boldsymbol{s} \rangle$ can easily be mapped to the ℓ_2-distance using the polarization identity.

Using these distance metrics we define the binary stable embedding.

Definition 1 (Binary ϵ-Stable Embedding)**.** *Let $\epsilon \in (0, 1)$. A mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ is a **binary ϵ-stable embedding** (BϵSE) of order K for sparse vectors if*

$$d_S(\boldsymbol{x}, \boldsymbol{s}) - \epsilon \leq d_H(A(\boldsymbol{x}), A(\boldsymbol{s})) \leq d_S(\boldsymbol{x}, \boldsymbol{s}) + \epsilon$$

for all $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$ with $\boldsymbol{x} \pm \boldsymbol{s} \in \Sigma_K$.

Our definition describes a specific quasi-isometry between the two metric spaces (S^{N-1}, d_S) and $(\mathcal{B}^M, d_H)$, restricted to sparse vectors. While this mirrors the form of the δ-stable embedding for sparse vectors, one important difference is that the sensitivity term ϵ is additive, rather than multiplicative, and thus the BϵSE is not bi-Lipschitz. This is a necessary side-effect of the loss of information due to quantization.

A stated in the next Lemma, the BϵSE enables robustness guarantees on any reconstruction algorithm extracting a sparse signal $\boldsymbol{x}^*$ from the mapping $A(\boldsymbol{x})$.

[4] A function $A : X \to Y$ is called a *quasi-isometry* between metric spaces (X, d_X) and (Y, d_Y) if there exists $C > 0$ and $D \geq 0$ such that $\frac{1}{C} d_X(\boldsymbol{x}, \boldsymbol{s}) - D \leq d_Y(A(\boldsymbol{x}), A(\boldsymbol{s})) \leq C d_X(\boldsymbol{x}, \boldsymbol{s}) + D$ for $\boldsymbol{x}, \boldsymbol{s} \in X$, and $E > 0$ such that $d_Y(y, A(\boldsymbol{x})) < E$ for all $y \in Y$ [50]. Since $D = 0$ for δ-stable embeddings, they are also called bi-Lipschitz mappings.

Lemma 2. *Let $A : \mathbb{R}^N \to \mathcal{B}^M$ be a BϵSE of order $2K$ for sparse vectors and let $\boldsymbol{x} \in \Sigma_K^*$. A sparse, unit norm estimate $\boldsymbol{x}^*$ of $\boldsymbol{x}$ with Hamming error $d_H\big(A(\boldsymbol{x}), A(\boldsymbol{x}^*)\big)$ from any reconstruction algorithm has angular error bounded by*

$$d_S(\boldsymbol{x}, \boldsymbol{x}^*) \ \leq d_H\big(A(\boldsymbol{x}), A(\boldsymbol{x}^*)\big) + \epsilon.$$

Proof. If $\boldsymbol{x}^*$ is K-sparse ($\|\boldsymbol{x}^*\|_0 \leq K$) and unit norm, then the result follows from the lower bound in Definition 1. □

In other words, the reconstruction error is bounded by a small quantity more than the Hamming error. Thus, if an algorithm returns a unit norm sparse solution with measurements that are not consistent (i.e., $d_H(A(\boldsymbol{x}), A(\boldsymbol{x}^*)) > 0$), as is the case with several algorithms [29–31], then the the worst-case angular reconstruction error is close to Hamming distance between the estimate's measurements' signs and the original measurements' signs. Section V verifies this behavior with simulation results. Furthermore, in Section III-C we use the BϵSE property to guarantee that if measurements are corrupted by noise or if signals are not exactly sparse, then the reconstruction error is bounded.

Note that if A is a BϵSE, then the angular error of any $\Delta^{\text{1bit}}(\boldsymbol{y}_s, \Phi, K)$ algorithm is bounded by ϵ since in that case $d_H\big(A(\boldsymbol{x}), A(\boldsymbol{x}^*)\big) = 0$. As we have seen earlier this is to be expected because, unlike conventional noiseless CS, quantization fundamentally introduces uncertainty and exact recovery cannot be guaranteed. This is an obvious consequence of the mapping of the infinite set Σ_K^* to a discrete set of quantized values.

We next identify a class of matrices Φ for which A is a BϵSE.

3.2 Binary ϵ-stable embeddings via random projections

As is the case for conventional CS systems with RIP, designing a Φ for 1-bit CS such that A has has the BϵSE property is a computationally intractable task. Fortunately, an overwhleming number of "good" matrices do exist. Specifically we again focus our analysis on Gaussian matrices, i.e., $\Phi \sim \mathcal{N}^{M\times N}(0, 1)$ such that each element $\phi_{i,j}$ is randomly drawn i.i.d. from $\mathcal{N}(0, 1)$, as in as in Section II-B. As motivation that this choice of Φ will indeed enable robustness, we begin with a classical concentration of measure result for binary measurements from a Gaussian matrix.

Lemma 3. *Let Φ be a matrix generated as $\Phi \sim \mathcal{N}^{M\times N}(0, 1)$, and let the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ be defined as in (3). Fix $\epsilon > 0$. For any $\boldsymbol{x}, \boldsymbol{s} \in S^{N-1}$, we have*

$$\mathbb{P}\Big(\ \Big| \, d_H\big(A(\boldsymbol{x}), A(\boldsymbol{s})\big) \ - \ d_S(\boldsymbol{x}, \boldsymbol{s}) \, \Big| \ \leq \ \epsilon \ \Big) \ \geq \ 1 - 2\, e^{-2\epsilon^2 M}, \tag{7}$$

where the probability is with respect to the generation of Φ.

Proof. This lemma is a simple consequence of Lemma 3.2 in [51] which shows that, for one measurement, $\mathbb{P}[A_j(\boldsymbol{x}) \neq A_j(\boldsymbol{s})] = d_S(\boldsymbol{x}, \boldsymbol{s})$. The result then follows by applying Hoeffding's inequality to the binomial random variable $M d_H\big(A(\boldsymbol{x}), A(\boldsymbol{s})\big)$ with M trials. □

In words, Lemma 3 implies that the Hamming distance between two binary measurement vectors $A(\boldsymbol{x}), A(\boldsymbol{s})$ tends to the angle between the signals $\boldsymbol{x}$ and $\boldsymbol{s}$ as the number of measurements M

increases. In [51] this fact is used in the context of randomized rounding for max-cut problems; however, this property has also been used in similar contexts as ours with regards to preservation of inner products from binary measurements [52, 53].

The expression (7) indeed looks similar to the definition of the BϵSE, however, it only holds for a fixed pair of arbitrary (not necessarily sparse) signals, chosen prior to drawing Φ. Our goal is to extend (7) to cover the entire set of sparse signals. Indeed, concentration results similar to Lemma 3, although expressed in terms of norms, have been used to demonstrate the RIP [15]. These techniques usually demonstrate that the cardinality of the space of all sparse signals is sufficiently small, such that the concentration result can be applied to demonstrate that distances are preserved with relatively few measurements.

Unfortunately, due to the non-linearity of A we cannot immediately apply Lemma 3 using the same procedure as in [15]. To briefly summarize, [15] proceeds by covering the set of all K-sparse signals Σ_K with a finite set of points (with covering radius $\delta > 0$). A concentration inequality is then applied to this set of points. Since any sparse signal lies in a δ-neighborhood of at least one such point, the concentration property can be extended from the finite set to Σ_K by bounding the distance between the measurements of the points within the δ-neighborhood. Such an approach cannot be used to extend (7) to Σ_K, because the severe discontinuity of our mapping does not permit us to characterize the measurements $A(\boldsymbol{x}+\boldsymbol{s})$ using $A(\boldsymbol{x})$ and $A(\boldsymbol{s})$ and obtain a bound on the distance between measurements of signals in a δ-neighborhood.

To resolve this issue, we extend Lemma 3 to include all points within Euclidean balls around the vectors $\boldsymbol{x}$ and $\boldsymbol{s}$ inside the (sub) sphere $\Sigma^*(T) = \{\boldsymbol{u} \in S^{N-1} : \operatorname{supp} \boldsymbol{u} \subset T\}$ for some fixed support set $T \subset \{1, \cdots, N\}$ of size $|T| = D$. Define the δ-ball $B_\delta(\boldsymbol{x}) := \{\boldsymbol{a} \in S^{N-1} : \|\boldsymbol{x}-\boldsymbol{a}\|_2 < \delta\}$ to be the ball of Euclidean distance δ around $\boldsymbol{x}$, and let $B^*_\delta(\boldsymbol{x}) = B_\delta(\boldsymbol{x}) \cap \Sigma^*(T)$.

Lemma 4. *Given $T \subset \{1, \cdots, N\}$ of size $|T| = D$, let Φ be a matrix generated as $\Phi \sim \mathcal{N}^{M\times N}(0,1)$, and let the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ be defined as in (3). Fix $\epsilon > 0$ and $0 \le \delta \le 1$. For any $\boldsymbol{x}, \boldsymbol{s} \in \Sigma^*(T)$, we have*

$$\mathbb{P}\Big(\Big| d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;-\; d_S(\boldsymbol{x}, \boldsymbol{s}) \Big| \;\le\; \epsilon + \sqrt{\tfrac{\pi}{2}D}\,\delta \Big) \;\ge\; 1 - 2\,e^{-2\epsilon^2 M}$$

for all $\boldsymbol{u} \in B^*_\delta(\boldsymbol{x})$ and $\boldsymbol{v} \in B^*_\delta(\boldsymbol{s})$.

The proof of this result is given in Appendix A.

In words, if the width δ is sufficiently small, then the Hamming distance between the 1-bit measurements $A(\boldsymbol{u})$, $A(\boldsymbol{v})$ of any points $\boldsymbol{u}$, $\boldsymbol{v}$ within the balls $B^*_\delta(\boldsymbol{x})$, $B^*_\delta(\boldsymbol{s})$, respectively, will be close to the angle between the centers of the balls.

Lemma 4 is key for providing a similar argument to that in [15]. We now simply need to count the number of pairs of K-sparse signals that are euclidean distance δ apart. The Lemma can then be invoked to demonstrate that the angles between all of these pairs will be approximately preserved by our mapping.[5] Thus, with Lemma 4 under our belt, we demonstrate in Appendix A the following result.

Theorem 3. *Let Φ be a matrix generated as $\Phi \sim \mathcal{N}^{M\times N}(0,1)$ and let the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ be defined as in (3). Fix $0 \le \eta \le 1$ and $\epsilon > 0$. If the number of measurements is*

$$M \;\ge\; \tfrac{4}{\epsilon^2}\big(K\,\log(N) + 2K\,\log(\tfrac{50}{\epsilon}) + \log(\tfrac{2}{\eta})\big), \tag{8}$$

[5] We note that the covering argument in the proof of Theorem 2 also employs δ-balls in similar fashion but only considers the probability that $d_H = 0$, rather than the concentration inequality.

then with probability exceeding $1-\eta$*, the mapping* A *is a B*ϵ*SE of order* K *for sparse vectors.*

By choosing $\Phi \sim \mathcal{N}^{M\times N}(0,1)$ with $M = O(K\log N)$, with high probability we ensure that the mapping A is a BϵSE. Additionally, from (8) we find that the error decreases as

$$\epsilon \;=\; O\Big((K/M)^{(1-\alpha)/2}\,(\log N)^{1/2}\Big),$$

for arbitrarily small $\alpha > 0$. Unfortunately, this decay is at a slower rate (roughly by a factor of $\sqrt{K/M}$) than the lower bound on the error given in Section II-A. This error rate results from an application of the Chernoff-Hoeffding inequality in the proof of Theorem 3. An open question is whether it is possible to obtain a tighter bound (with optimal error rate) for this robustness property.

As with Theorem 2, Gaussian matrices provide a universal mapping, i.e., the result remains valid for sparse signals in a basis $\Psi \in \mathbb{R}^{N\times N}$. Moreover, Theorem 3 can also be extended to rows of Φ that are drawn uniformly on the sphere, since the rows of Φ in Theorem 3 can be normalized without affecting the outcome of the proof. Note that by normalizing the Gaussian rows of Φ, is is as if they had been drawn from a uniform distribution of unit-norm signals.

We have now established a large class of robust BϵSEs: 1-bit quantized Gaussian projections. We now make use of this robustness by considering an example where the measurements are corrupted by Gaussian noise.

3.3 Noisy measurements and compressible signals

In practice, hardware systems may be inaccurate when taking measurements; this is often modeled by additive noise. The mapping A is robust to noise in an unusual way. After quantization, the measurements can only take the values -1 or 1. Thus, we can analyze the reconstruction performance from corrupted measurements by considering how many measurements flip their signs. For example, we analyze the specific case of Gaussian noise on the measurements prior to quantization, i.e.,

$$A_n(\boldsymbol{x}) := \operatorname{sign}(\Phi\boldsymbol{x} + \boldsymbol{n}), \tag{9}$$

where $\boldsymbol{n} \in \mathbb{R}^M$ has i.i.d. elements $n_i \sim \mathcal{N}(0,\sigma^2)$. In this case, we demonstrate, via the following lemma, a bound on the Hamming distance between the corrupted and ideal measurements with the BϵSE from Theorem 3 (see Appendix A).

Lemma 5. *Let* Φ *be a matrix generated as* $\Phi \sim \mathcal{N}^{M\times N}(0,1)$*, let the mapping* $A : \mathbb{R}^N \to \mathcal{B}^M$ *be defined as in (3), and let* $A_n : \mathbb{R}^N \to \mathcal{B}^M$ *be defined as in (9). Let* $\boldsymbol{n} \in \mathbb{R}^M$ *be a Gaussian random vector with* i.i.d. *components* $n_i \sim \mathcal{N}(0,\sigma^2)$*. Fix* $\gamma > 0$*. Then for any* $\boldsymbol{x} \in \mathbb{R}^N$*, we have*

$$\begin{aligned}
&\mathbb{E}\left(\, d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big)\,\right) \;\le\; e(\sigma, \|\boldsymbol{x}\|_2),\\
&\mathbb{P}\left(d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big) > \; e(\sigma, \|\boldsymbol{x}\|_2) + \gamma\,\right) \le\; e^{-2M\gamma^2},
\end{aligned}$$

where $e(\sigma, \|\boldsymbol{x}\|_2) = \frac{1}{2}\frac{\sigma}{\sqrt{\|\boldsymbol{x}\|_2^2+\sigma^2}} \le \frac{1}{2}\frac{\sigma}{\|\boldsymbol{x}\|_2}$.

If $\boldsymbol{x}_n^*$ is the estimate from a sparse consistent reconstruction algorithm $\Delta^{\text{1bit}}(A_n(\boldsymbol{x}), \Phi, K)$ from the measurements $A_n(\boldsymbol{x})$, then it immediately follows from Lemma 5 and Theorem 3 that

$$d_S(\boldsymbol{x}_n^*, \boldsymbol{x}) \;\le\; d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big) + \epsilon \;\le\; \tfrac{1}{2}\tfrac{\sigma}{\|\boldsymbol{x}\|_2} + \gamma + \epsilon, \tag{10}$$

with high probability (depending on M and γ). Given alternative noise distributions, e.g., Poisson noise, a similar analysis can be carried out to determine the likely number of sign flips and thus provide a bound on the error due to noise.

Another practical consideration is that real signals are not always strictly K-sparse. Indeed, it may be the case that signals are *compressible*; i.e., they can be closely approximated by a K-sparse signal. Lemma 5 can be extended to compressible signals. To do this, we consider the small coefficients, i.e., the "tail" of the residual of a best K-term approximation of $\boldsymbol{x}$, to be a source of Gaussian noise on the measurements and then apply Lemma 5. This is possible due to our particular Gaussian choice of Φ and the fact that for binary measurements, we are only concerned with the number of measurements that change sign.

Corollary 2. *Let Φ be a matrix generated as $\Phi \sim \mathcal{N}^{M\times N}(0,1)$ and let the mapping $A : \mathbb{R}^N \to \mathcal{B}^M$ be defined as in (3). Furthermore, let Φ have RIP constant δ_K. Let $\gamma > 0$. Then for any $\boldsymbol{x} \in \mathbb{S}^{N-1}$ we have*

$$\begin{aligned}
&\mathbb{E}\left(\, d_H\big(A(\boldsymbol{x}), A(\boldsymbol{x}_K)\big)\,\right) \;\leq\; \tfrac{\rho(\boldsymbol{x},K)}{2\|\boldsymbol{x}_K\|_2},\\
&\mathbb{P}\left(d_H\big(A(\boldsymbol{x}), A(\boldsymbol{x}_K)\big) > \;\tfrac{\rho(\boldsymbol{x},K)}{2\|\boldsymbol{x}_K\|_2} + \gamma\right) \leq\; e^{-2M\gamma^2},\\
&\rho(\boldsymbol{x},K) = \sqrt{1+\delta_K}\big(\|\boldsymbol{x}-\boldsymbol{x}_K\|_2 + \|\boldsymbol{x}-\boldsymbol{x}_K\|_1/\sqrt{K}\big),
\end{aligned}$$

where $\boldsymbol{x}_K$ is the best K-term approximation of $\boldsymbol{x}$.

The proof, which uses Lemma 6.1 of [13], is given in Appendix A. In similar fashion to (10), this result implies that with high probability (depending on M and γ), the angular reconstruction error of $\boldsymbol{x}^* = \Delta^{1\text{bit}}(A(\boldsymbol{x}), \Phi, K)$ for any signal $\boldsymbol{x}$ (sparse or compressible) is bounded as

$$d_S(\boldsymbol{x}^*, \boldsymbol{x}) \leq \frac{\sqrt{1+\delta_K}}{2\|\boldsymbol{x}_K\|_2}\left(\|\boldsymbol{x}-\boldsymbol{x}_K\|_2 + \frac{\|\boldsymbol{x}-\boldsymbol{x}_K\|_1}{\sqrt{K}}\right) + \gamma + \epsilon,$$

Much like conventional CS results, the reconstruction error on the order of the best K-term approximation error of the signal.

Thus far we have demonstrated a lower bound on the reconstruction error from 1-bit measurements (Theorem 2) and introduced a condition on the mapping A that enables stable reconstruction in noiseless, noisy, and compressible settings (Definition 1). We have furthermore demonstrated that a large class of random matrices—specifically matrices with coefficients draw from a Gaussian distribution and matrices with rows drawn uniformly from the unit sphere—provide good mappings (Theorem 3). We now provide a more practical contribution in the form of a new algorithm for reconstruction of sparse signals from 1-bit measurements.

4 BIHT: A Simple First-Order Reconstruction Algorithm

4.1 Problem formulation and algorithm definition

We now introduce a simple algorithm for the reconstruction of sparse signals from 1-bit compressive measurements. Our algorithm, *Binary Iterative Hard Thresholding* (BIHT), is a simple modification of IHT, the real-valued algorithm from which is takes its name [14]. The IHT algorithm has recently

been extended to handle measurement non-linearities [54]; however, these results do not apply to quantized measurements since quantization does not satisfy the requirements in [54].

We briefly recall that the IHT algorithm consists of two steps that can interpreted as follows. The first step can be thought of as a gradient descent to reduce the least squares objective $\|\boldsymbol{y} - \Phi\boldsymbol{x}\|_2^2/2$. Thus, at iteration l, IHT proceeds by setting $\boldsymbol{a}^{l+1} = \boldsymbol{x}^l + \Phi^T(\boldsymbol{y} - \Phi\boldsymbol{x})$. The second step imposes a sparse signal model by projecting $\boldsymbol{a}^{l+1}$ onto the "ℓ_0 ball", i.e., selecting the K largest in magnitude elements. Thus, IHT for CS can be thought of as trying to solve the problem

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \ \tfrac{1}{2}\|\boldsymbol{y} - \Phi\boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_0 = K. \tag{11}$$

The BIHT algorithm simply modifies the first step of IHT to instead minimize a consistency-enforcing objective. Specifically, given an initial estimate $\boldsymbol{x}^0 = \boldsymbol{0}$ and 1-bit measurements $\boldsymbol{y}_s$, at iteration l BIHT computes

$$\boldsymbol{a}^{l+1} = \boldsymbol{x}^l + \frac{\tau}{2}\Phi^T\big(\boldsymbol{y}_s - A(\boldsymbol{x}^l)\big), \tag{12}$$

$$\boldsymbol{x}^{l+1} = \eta_K(\boldsymbol{a}^{l+1}), \tag{13}$$

where A is defined as in (3), τ is a scalar that controls gradient descent step-size, and the function $\eta_K(\boldsymbol{v})$ computes the best K-term approximation of $\boldsymbol{v}$ by thresholding. Once the algorithm has terminated (either consistency is achieved or a maximum number of iterations have been reached), we then normalize the final estimate to project it onto the unit sphere. Section IV-B discusses several variations of this algorithm, each with different properties.

The key to understanding BIHT lies in the formulation of the objective. The following Lemma shows that the term $\Phi^T\big(\boldsymbol{y}_s - A(\boldsymbol{x}^l)\big)$ in (12) is in fact the negative subgradient of a convex objective $\mathcal{J}$. Let $[\cdot]_-$ denote the negative function, i.e., $([\boldsymbol{u}]_-)_i = [u_i]_-$ with $[u_i]_- = u_i$ if $u_i < 0$ and 0 else, and $\boldsymbol{u} \odot \boldsymbol{v}$ denote the Hadamard product, i.e., $(\boldsymbol{u} \odot \boldsymbol{v})_i = u_i v_i$ for two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$.

Lemma 6. *The quantity* $\frac{1}{2}\,\Phi^T\big(A(\boldsymbol{x}) - \boldsymbol{y}_s\big)$ *in (12) is a subgradient of the convex one-sided* ℓ_1*-norm*

$$\mathcal{J}(\boldsymbol{x}) = \|[\boldsymbol{y}_s \odot (\Phi\boldsymbol{x})]_-\|_1,$$

Thus, BIHT aims to decrease $\mathcal{J}$ at each step (12).

Proof. We first note that $\mathcal{J}$ is convex. We can write $\mathcal{J}(\boldsymbol{x}) = \sum_i \mathcal{J}_i(\boldsymbol{x})$ with each convex function $\mathcal{J}_i$ given by

$$\mathcal{J}_i(\boldsymbol{x}; \boldsymbol{y}_s, \Phi) = \begin{cases} |\langle \boldsymbol{\varphi}_i, \boldsymbol{x}\rangle|, & \text{if } A_i(\boldsymbol{x})\,(\boldsymbol{y}_s)_i < 0, \\ 0, & \text{else,} \end{cases}$$

where $\boldsymbol{\varphi}_i$ denotes a row of Φ and $A_i(\boldsymbol{x}) = \operatorname{sign}\langle \boldsymbol{\varphi}_i, \boldsymbol{x}\rangle$. Moreover, if $\langle \boldsymbol{\varphi}_i, \boldsymbol{x}\rangle \neq 0$, then the gradient of $\mathcal{J}_i$ is

$$\boldsymbol{\nabla}\mathcal{J}_i(\boldsymbol{x}; \boldsymbol{y}_s, \Phi) = \tfrac{1}{2}(A_i(\boldsymbol{x}) - (\boldsymbol{y}_s)_i)\,\boldsymbol{\varphi}_i = \begin{cases} A_i(\boldsymbol{x})\,\boldsymbol{\varphi}_i & \text{if } (\boldsymbol{y}_s)_i\,A_i(\boldsymbol{x}) < 0, \\ 0, & \text{else} \end{cases}$$

while if $\langle \boldsymbol{\varphi}_i, \boldsymbol{x}\rangle = 0$, then the gradient is replaced by the subdifferential set

$$\boldsymbol{\nabla}\mathcal{J}_i(\boldsymbol{x}; \boldsymbol{y}_s, \Phi) = \big\{\tfrac{\xi}{2}(A_i(\boldsymbol{x}) - (\boldsymbol{y}_s)_i)\,\boldsymbol{\varphi}_i : \xi \in [0,1]\big\} \ \ni \ \tfrac{1}{2}(A_i(\boldsymbol{x}) - (\boldsymbol{y}_s)_i)\,\boldsymbol{\varphi}_i.$$

Thus, by summing over i we conclude that $\frac{1}{2}\,\Phi^T\big(A(\boldsymbol{x}) - \boldsymbol{y}_s\big) \in \boldsymbol{\nabla} J(\boldsymbol{x}; \boldsymbol{y}_s, \Phi)$. □

Consequently, the BIHT algorithm can be thought of as trying to solve the problem:

$$\boldsymbol{x}^* \;=\; \underset{\boldsymbol{x}}{\operatorname{argmin}}\; \tau \,\|[\boldsymbol{y}_s \odot (\Phi \boldsymbol{x})]_-\|_1 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_0 = K,\, \|\boldsymbol{x}\|_2 = 1.$$

Observe that since $\boldsymbol{y}_s \odot (\Phi \boldsymbol{x})$ simply scales the elements of $\Phi \boldsymbol{x}$ by the signs $\boldsymbol{y}_s$, minimizing the one-sided ℓ_1 objective enforces a positivity requirement,

$$\boldsymbol{y}_s \odot (\Phi \boldsymbol{x}) \geq 0, \tag{14}$$

that, when satisfied, implies consistency.

Previously proposed 1-bit CS algorithms have used a one-sided ℓ_2-norm to impose consistency [28–31]. Specifically, they have applyied a constraint or objective that takes the form $\|[\boldsymbol{y}_s \odot (\Phi \boldsymbol{x})]_-\|_2^2/2$. Both the one-sided ℓ_1 and ℓ_2 functions imply a consistent solution when they evaluate to zero, and thus, both approaches are capable of enforcing consistency. However, the choice of the ℓ_1 vs. ℓ_2 penalty term makes a significant difference in performance depending on the noise conditions. We explore this difference in the experiments in Section V.

4.2 BIHT shifts

Several modifications can be made to the BIHT algorithm that may improve certain performance aspects, such as consistency, reconstruction error, or convergence speed. While a comprehensive comparison is beyond the scope of this paper, we believe that such variations exhibit interesting and useful properties that should be mentioned.

Projection onto sphere at each iteration. We can enforce that every intermediate solution have unit ℓ_2 norm. To do this, we modify the "impose signal model" step (13) by normalizing after choosing the best K-term approximation, i.e., we compute

$$\boldsymbol{x}^{l+1} = U\big(\eta_K(\boldsymbol{a}^{l+1})\big), \tag{15}$$

where $U(\boldsymbol{v}) = \boldsymbol{v}/\|\boldsymbol{v}\|_2$. While this step is necessary for previous algorithms such as [29–31], it is in general not necessary in the BIHT case.

If we choose to impose the projection, Φ must be appropriately normalized or, equivalently, the step size of the gradient descent must be carefully chosen. Otherwise, the algorithm will not converge. Empirically, we have found that for a Gaussian matrix, an appropriate scaling is $1/(\sqrt{M}\|\Phi\|_2)$, where the $1/\|\Phi\|_2$ controls the amplification of the estimate from Φ^T in the gradient descent step (12) and the $1/\sqrt{M}$ ensures that $\|\boldsymbol{y}_s - A(\boldsymbol{x}^l)\|_2 \leq 2$. Similar gradient step scaling requirements have been imposed in the conventional IHT algorithm and other sparse recovery algorithms as well (e.g., [9]).

Minimizing hinge loss. The one-sided ℓ_1-norm is related to the *hinge-loss* function in the machine learning literature, which is known for its robustness to outliers [55]. Binary classification algorithms seek to enforce the same consistency function as in (14) by minimizing a function $\sum[\kappa - \boldsymbol{y}_s \odot (\Phi \boldsymbol{x})]_+$, where $[\cdot]_+$ sets negative elements to zero. When $\kappa > 0$, the objective is both convex and has a non-trivial solution. Further connections and interpretations are discussed in Section V. Thus, rather than minimizing the one-sided ℓ_1 norm, we can instead minimize the hinge-loss. The hinge-loss can be interpreted as ensuring that the minimum value that an unquantized measurement $(\Phi \boldsymbol{x})_i$ can take is bounded away form zero, i.e., $|(\Phi \boldsymbol{x})_i| \geq \kappa$. This requirement is

similar to the sphere constraint in that it avoids a trivial solution; however, will perform differently than the sphere constraint. In this case, in the gradient descent step (12), we instead compute

$$\boldsymbol{a}^{l+1} = \boldsymbol{x}^l - \tau\Psi^T(\text{sign}(\Psi\boldsymbol{x}^l - \kappa) - 1)/2$$

where $\Psi = (\boldsymbol{y}_s \odot \Phi)$ scales the rows of Φ by the signs of $\boldsymbol{y}_s$. Again, the step size must be chosen appropriately, this time as $C_\kappa/\|\Phi\|_2$, where C_κ is a parameter that depends on κ.

Minimizing other one-sided objectives. In general, any function $\mathcal{R}(\boldsymbol{x}) = \sum \mathcal{R}_i(x_i)$, where $\mathcal{R}_i$ is continuous and has a negative gradient for $x_i \leq 0$ and is 0 for $x_i > 0$, can be used to enforce consistency. To employ such functions, we simply compute the gradient of $\mathcal{R}$ and apply it in (12).

As an example, the previously mentioned one-sided ℓ_2-norm has been used to enforce consistency in several algorithms. We can use it in BIHT by computing

$$\boldsymbol{a}^l = \boldsymbol{x}^l + \tau\Phi^T[\boldsymbol{y}_s \odot \Phi\boldsymbol{x}^l]_+$$

in (12). We compare and contrast the behavior of the one-sided ℓ_1 and ℓ_2 norms in Section V.

As another example, in similar fashion to the Huber norm [56], we can combine the ℓ_1 and ℓ_2 functions in a piecewise fashion. One potentially useful objective is $\sum \mathcal{R}_i(\boldsymbol{x})$, where $\mathcal{R}_i$ is defined as follows:

$$\mathcal{R}_i(\boldsymbol{x}) = \begin{cases} 0, & x_i \geq 0, \\ |x_i|, & -\frac{1}{2} \leq x_i < 0, \\ x_i^2 + \frac{1}{4}, & x_i < -\frac{1}{2}. \end{cases} \tag{16}$$

While similar, this is not exactly a one-sided Huber norm. In a one-sided Huber-norm, the square (ℓ_2) term would be applied to values near zero and the magnitude (ℓ_1) term would be applied to values significantly less than zero, the reverse of what we propose here.

This objective can provide different robustness properties or convergence rates than the previously mentioned objectives. Specifically, during each iteration it may allow us to take advantage of the shallow gradient of the one-sided ℓ_2 cost for large numbers of measurement sign discrepancies and the steeper gradient of the one-sided ℓ_1 cost when most measurements have the correct sign. This objective can be applied in BIHT as with the other objectives, by computing its gradient and plugging it into (12).

5 Experiments

In this section we explore the performance of the BIHT algorithm and compare it to the performance of previous algorithms for 1-bit CS. To make the comparison as straightforward as possible, we reproduced the experiments of [31] with the BIHT algorithm.

The experimental setup is as follows. For each data point, we draw a length-N, K-sparse signal with the non-zero entries drawn uniformly at random on the unit sphere, and we draw a new $M \times N$ matrix Φ with each entry $\phi_{ij} \sim \mathcal{N}(0, 1)$. We then compute the binary measurements $\boldsymbol{y}_s$ according to (3). Reconstruction of $\boldsymbol{x}^*$ is performed from $\boldsymbol{y}_s$ with three algorithms: *matching sign pursuit* (MSP) [30], *restricted-step shrinkage* (RSS) [31], and BIHT (this paper); the algorithms will be depicted by dashed, dotted, and triangle lines, respectively. Each reconstruction in this setup is repeated for 1000 trials and with a fixed $N = 1000$ and $K = 10$ unless otherwise noted. Furthermore, we perform the trials for M/N within the range $[0, 2]$. Note that when $M/N > 1$, we

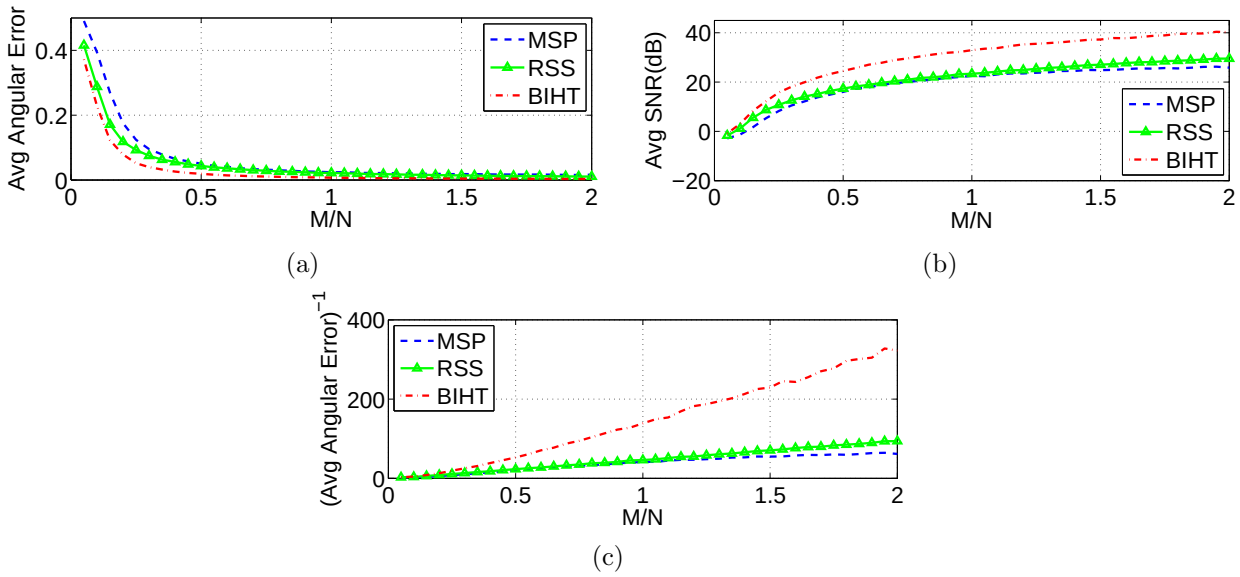


Figure 2: Average reconstruction angular error ϵ_{sim} vs. M/N, plotted three ways. (a) Angular error ϵ_{sim}, (b) SNR in decibels, and (c) Inverse angular error $\epsilon_{\text{sim}}^{-1}$. The plot demonstrates that BIHT yields a considerable improvement in reconstruction error, achieving an SNR as high as 40dB when $M/N = 2$. Furthermore, we see that the error behaves according $\epsilon_{\text{sim}} = O\left(1/M\right)$, implying that on average we achieve the optimal performance rate given in Theorem 1.

are acquiring more measurements than the ambient dimension of the signal. While the $M/N > 1$ regime is not interesting in conventional CS, it may be very practical in 1-bit systems that can acquire sign measurements at extremely high, super-Nyquist rates.

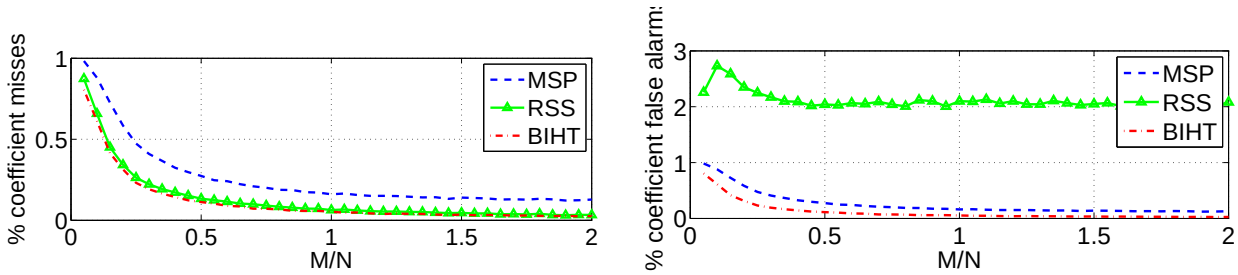
Average error. We begin by measuring the average reconstruction angular error ϵ_{sim} over the 1000 trials. The results of this are depicted in Figure 2. We display the results of this experiment three different ways: *(i)* the true angular error in Figure 2(a), which we denote as ϵ_{sim}, to demonstrate typical values achieved, *(ii)* the signal-to-noise ratio (SNR)[6] in Figure 2(b), to demonstrate that the performance of these techniques is practical (since the angular error is unintuitive to most observers), and *(iii)* the inverse of the angular error squared, i.e., $\epsilon_{\text{sim}}^{-1}$ in Figure 2(c), to compare with the performance predicted by Theorem 2.

We begin by comparing the performance of the algorithms. While the angular error of each algorithm appears to follow the same trend, BIHT obtains smaller error (or higher SNR) than the others, significantly so when M/N is greater than 0.35. The discrepancy in performance could be due to difference in the algorithms themselves, or perhaps, differences in their formulations for enforcing consistency. This is explored later in this section.

We now consider the actual performance trend. We see from Figure 2(c) that, above $M/N = 0.35$ each line appears fairly linear, albeit with a different slope, implying that with all other variables fixed, $\epsilon_{\text{sim}} = O\left(1/M\right)$. This is on the order of the optimal performance as given by the bound given in Theorem 1 and predicted by Theorem 2 for Gaussian matrices.

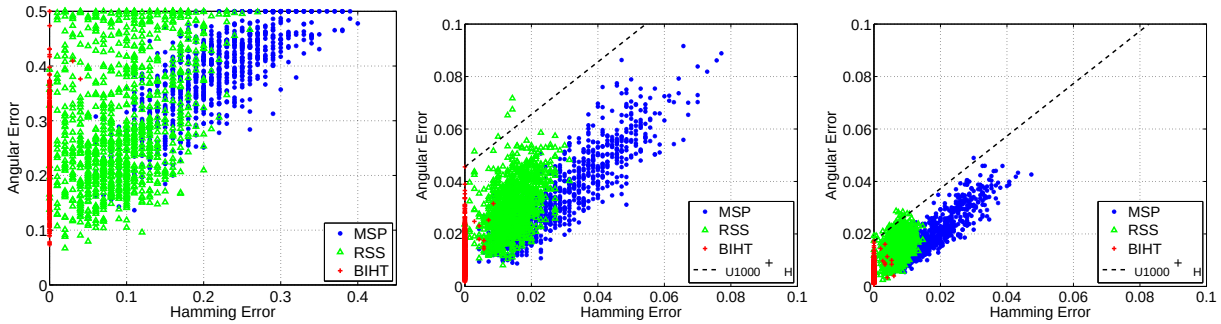
Misses and false alarms. We dig a little deeper into the source of errors by examining the

[6]In this paper we define the reconstruction SNR in decibels as $\text{SNR}(\boldsymbol{x}) := 10\log_{10}(\|\boldsymbol{x}\|_2^2/\|\boldsymbol{x} - \boldsymbol{x}^*\|_2^2)$. Note that this metric uses the standard euclidean error and not angular error.



(a) (b)

Figure 3: Reconstructed signal coefficient (a) misses, and (b) false-alarms. The MSP algorithm is most likely to miss a coefficient, while RSS and BIHT perform comparably. The RSS algorithm returns a large number of coefficients that are close to zero and thus performs poorly in the false-alarms metric. Both BIHT and MSP are restricted to have at most K false alarms by design.



(a) M/N = 0.1 (b) M/N = 0.7 (c) M/N =1.5

Figure 4: Reconstruction angular error ϵ_{sim} vs. measurement Hamming error ϵ_H. BIHT returns a consistent solution in most trials, even when the number of measurements is too low to permit a small angular error (see (a) M/N = 0.1). For larger M/N regimes, we see a linear relationship $\epsilon_{\text{sim}} \approx C + \epsilon_H$ between the average angular error ϵ_{sim} and the hamming error ϵ_H where C is constant (see (b) and (c)). The BϵSE formulation in Definition 1 predicts that the angular error is bounded by the hamming error ϵ_H in addition to an offset ϵ. The dashed line $\epsilon_{U1000} + \epsilon_H$ denotes the empirical upper bound for 1000 trials.

reconstruction "misses," i.e., those coefficients that were identified as zero that are non-zero in the true signal, as well as the "false-alarms", i.e., those coefficients that were identified as non-zero that are zero in the true signal. The results are depicted in Figure 3(a) and (b), respectively. In both cases, BIHT outperforms the other algorithms, although it is very close to the RSS algorithm in the number of misses. While both RSS and MSP have significantly more false-alarms than BIHT, by design, MSP can return at most K non-zero coefficients and thus cannot have more than K false alarms. Meanwhile, the RSS algorithm may have many coefficients that are significantly close to zero but are numerically counted as non-zeros.

Consistency. We also expose the relationship between the Hamming distance $d_H(A(\boldsymbol{x}), A(\boldsymbol{x}^*))$ between the measurements of the true and reconstructed signal and the angular error of the true and reconstructed signal. Figure 4 depicts the Hamming distance vs. angular error for three different values of M/N. The particularly striking result is that BIHT returns significantly more consistent reconstructions than the two other algorithms. This is clear from the fact that most of the red (plus) points lie on the y-axis while the majority of blue (dot) or green (triangle) points do not.

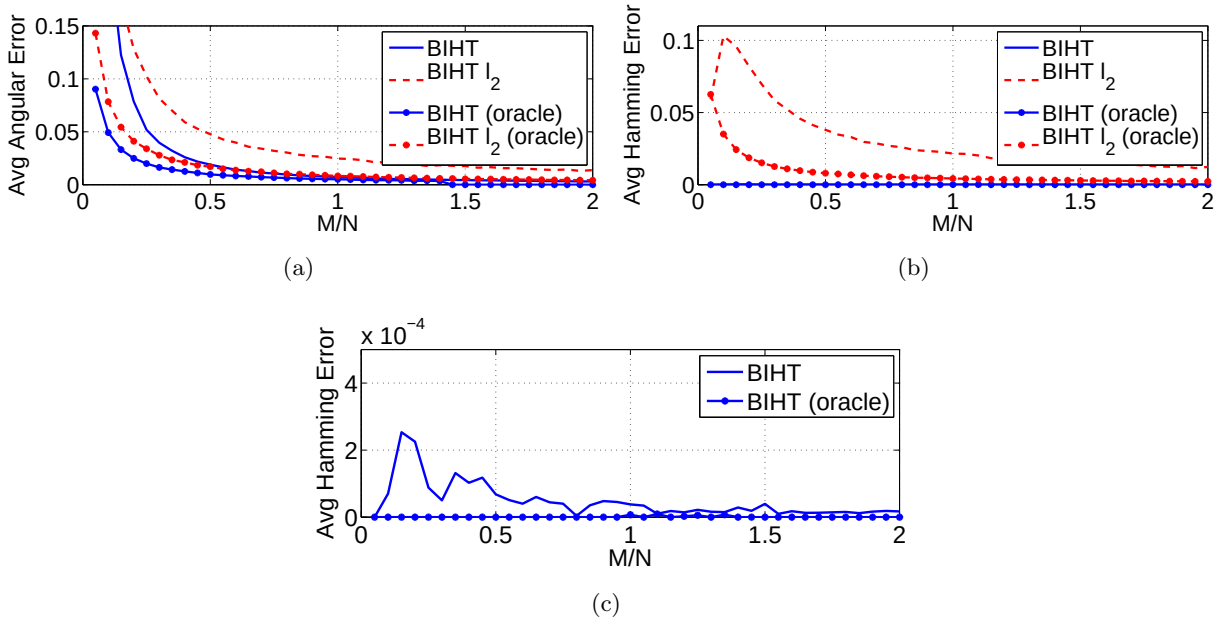


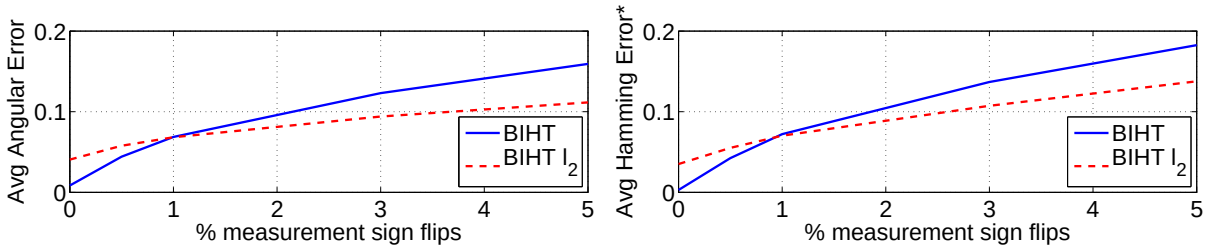
Figure 5: Enforcing consistency: One-sided ℓ_1 vs. one-sided ℓ_2 BIHT. When BIHT attempts to minimize a one-sided ℓ_2 instead of a one-sided ℓ_1 objective, the performance significantly decreases. We find this to be the case even when an oracle provides the true signal support a priori. Note: (c) is simply a zoomed version (b).

We find that, even in significantly "under-sampled" regimes like $M/N = 0.1$, where the BϵSE is unlikely to hold, BIHT is likely to return a consistent solution (albeit with high variance of angular errors). We also find that in "over-sampled" regimes such as $M/N = 1.7$, the range of angular errors on the y-axis is small.

We can infer an interesting performance trend from Figures 4(b) and (c), where the BϵSE property may hold. Since the RSS and MSP algorithms often do not return a consistent solution, we can visualize the relationship between angular error and hamming error. Specifically, on average the angular reconstruction error is a linear function of hamming error, $\epsilon_H = d_H(A(\boldsymbol{x}), A(\boldsymbol{x}^*))$, as similarly expressed by the reconstruction error bound provided by BϵSE. Furthermore, if we let ϵ_{1000} be the largest angular error (with consistent measurements) over 1000 trials, then we can suggest an empirical upper bound for BIHT of $\epsilon_{1000} + \epsilon_H$. This upper bound is denoted by the dashed line in Figures 4(b) and (c).

One-sided ℓ_1 vs. one-sided ℓ_2 objectives. As demonstrated in Figures 2 and 4, the BIHT algorithm achieves significantly improved performance over MSP and RSS in both angular error and Hamming error (consistency). A significant difference between these algorithms and BIHT is that MSP and RSS seek to impose consistency via a one-sided ℓ_2-norm, as described in Section IV-B. Minimizing either the one-sided ℓ_1 or one-sided ℓ_2 objectives will enforce consistency on the measurements of the solution; however, the behavior of these two terms appears to be significantly different, according to the previously discussed experiments.

To test the hypothesis that this term is the key differentiator between the algorithms, we implemented BIHT-ℓ_2, a one-sided ℓ_2 variation of the BIHT algorithm that enabled a fair comparison of the one-sided objectives (see Section IV-B for details). We compared both the angular error and



(a) (b)

Figure 6: Enforcing consistency with noise: One-sided ℓ_1 vs. one-sided ℓ_2 BIHT. When BIHT attempts to minimize a one-sided ℓ_2 instead of the one-sided ℓ_1 objective, the algorithm is more robust to flips of measurement signs. *Note that the Hamming error in (b) is measured with regards to the *noisy* measurements, e.g., a Hamming error of zero means that we reconstructed the signs of the noisy measurements exactly.

Hamming error performance of BIHT and BIHT-ℓ_2. Furthermore, we implemented *oracle assisted* variations of these algorithms where the true support of the signal is given a priori, i.e., η_K in (13) is replaced by an operator that always selects the true support, and thus the algorithm only needs to estimate the correct coefficient values. The oracle assisted case can be thought of as a "best performance" bound for these algorithms. Using these algorithms, we perform the same experiment detailed at the beginning of the section.

The results are depicted in Figure 5. The angular error behavior of BIHT-ℓ_2 is very similar to that of MSP and RSS and underperforms when compared to BIHT. We see the same situation with regards to Hamming error: BIHT finds consistent solutions for the majority of trials, but BIHT-ℓ_2 does not. Thus, the results of this simulation suggest that the one-sided term plays a significant role in the quality of the solution obtained.

One way to explain the performance discrepancy between the two objectives comes from observing the deep connection between our reconstruction problem and binary classification. As explained previously, in the classification context, the one-sided ℓ_1 objective is similar to the hinge-loss, and furthermore, the one-sided ℓ_2 objective is similar to the so-called *square-loss*. Previous results in machine learning have shown that for typical convex loss functions, the minimizer of the hinge loss has the tightest bound between expected risk and the Bayes optimal solution [57] and good error rates, especially when considering robustness to outliers [57, 58]. Thus, the hinge loss is often considered superior to the square loss for binary classification.[7] One might suspect that since the one-sided ℓ_1-objective is very similar to the hinge loss, it too should outperform other objectives in our context. Understanding why in our context, the geometry of the ℓ_1 and ℓ_2 objectives results in different performance is an interesting open problem.

We probed the one-sided ℓ_1/ℓ_2 objectives further by testing the two versions of BIHT on noisy measurements. We flipped a number of measurement signs at random in each trial. For this experiment, $N = M = 1000$ and $K = 10$ are fixed, and we performed 100 trials. We varied the number of sign flips between 0% and 5% of the measurements. The results of the experiment are depicted in Figure 6. We see that for both the angular error in Figure 6(a) and Hamming error

[7] Additional "well-behaved" loss functions (e.g., the Huber-ized hinge loss) have been proposed [59] and a host of classification algorithms related to this problem exist [58–62], both of which may prove useful in the 1-bit CS framework in the future.

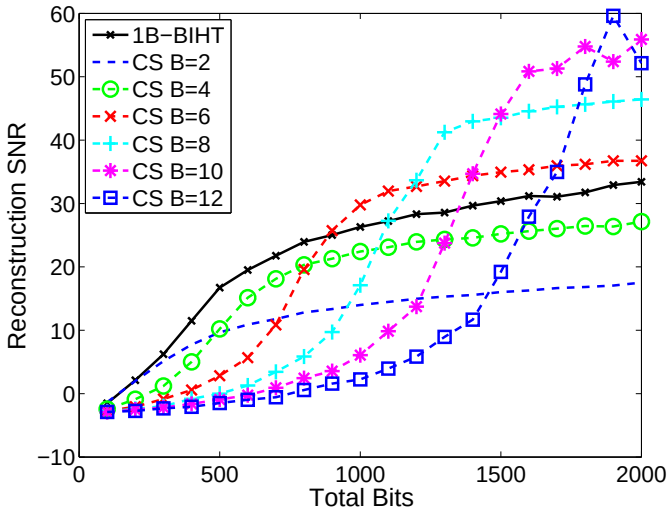


Figure 7: Comparison of BIHT to conventional CS multibit uniform scalar quantization (multibit reconstructions performed using BPDN [8]). BIHT is competitive with standard CS working with multibit measurements when the total number of bits is severely constrained. In particular, the BIHT algorithm performs strictly better than CS with 4 bits per measurement.

in Figure 6(b), that the one-sided ℓ_1 objective performs better when there are only a few errors and the one-sided ℓ_2 objective performs better when there are significantly more errors. This is expected since the ℓ_1 objective promotes sparse errors. This experiment implies that BIHT-ℓ_2 (and the other one-sided ℓ_2-based algorithms) may be more useful when the measurements contain significant noise that might cause a large number of sign flips, such as Gaussian noise.

Performance with a fixed bit-budget. In some applications we are interested in reducing the total number of bits acquired due to storage or communication costs. Thus, given a fixed total number of bits, an interesting question is how well 1-it CS performs in comparison to conventional CS quantization schemes and algorithms. For the sake of brevity, we give a simple comparison here between the 1-bit techniques and uniform quantization with *Basis Pursuit DeNoising* (BPDN) [8] reconstruction. While BPDN is not the optimal reconstruction technique for quantized measurements, it (and its variants such as the LASSO [59]) is considered a benchmark technique for reconstruction from measurements with noise and furthermore, is widely used in practice.

The experiment proceeds as follows. Given a total number of bits and a (uniform) quantization bit-depth B (i.e., number of bits per measurement), we choose the number of measurements as $M = \text{total bits}/B$, $N = 2000$, and the sparsity $K = 20$. The remainder of the experiment proceeds as described earlier (in terms of drawing matrices and signals). For bit depth greater than 1, we reconstruct using BPDN with an optimal choice of noise parameter and we scale the quantizer to such that signal can take full advantage of its dynamic range.

The results of this experiment are depicted in Figure 7. We see a common trend in each line: lackluster performance until "sufficient" measurements are acquired, then a slow but steady increase in performance as additional measurement are added, until a performance plateau is reached. Thus, since lower bit-depth implies that a larger number of measurements will be used, 1-bit CS reaches the performance plateau earlier than in the multi-bit case (indeed, the transition point is achieved at a higher number of total bits as the bit-depth is increased). This enables significantly improved

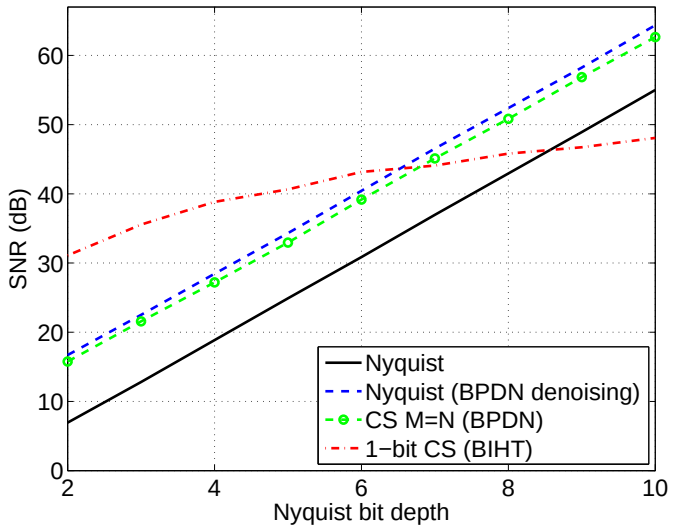


Figure 8: Comparison of uniformly quantized Nyquist-rate samples with linear reconstruction (solid) and BPDN denoising (dashed), CS with $M = N$ and BPDN reconstruction (dash-circle), and 1-bit quantized CS measurements with BIHT reconstruction (dash-dotted). Nyquist samples were quantized with bit-depth $\beta \in [2, 10]$ and 1-bit CS used $M = \beta N$ measurements; the same number of bits is used in each reconstruction. The Nyquist-rate lines have the classical 6.02dB/bit-depth slope, as expected. For a fixed number of bits, 1-bit CS does *not* follow this slope and outperforms conventional quantization when $\beta < 6$.

performance when the rate is severely constrained and higher bit-rates per measurements would significantly reduce the number of available measurements. For higher bit-rates, as expected from the analysis in [33], using fewer measurements with refined quantization achieves better performance.

It is also important to note that, regardless of trend, the BIHT algorithm performs strictly better than BPDN with 4 bits per measurement and uniform quantization for the parameters tested here. This gain is consistent with similar gains observed in [29, 30]. A more thorough comparison of additional CS quantization techniques with 1-bit CS is a subject for future study.

Comparison to quantized Nyquist samples. In our final experiment, we compare the performance of the 1-bit CS technique to the performance of a conventional uniform quantizer applied to uniform Nyquist-rate samples. Specifically, in each trial we draw a new Nyquist-sampled signal in the same way as in our previous experiments and with fixed $N = 2000$ and $K = 20$; however, now the signals are sparse in the discrete cosine transform (DCT) domain. We consider four reconstruction experiments. First, we quantize the Nyquist-rate signal with a bit-depth of β bits per sample (and optimal quantizer scale) and perform linear reconstruction (i.e., we just use the quantized samples as sample values). Second, we apply BPDN to the quantized Nyquist-rate samples with optimal choice of noise parameter, thus denoising the signal using a sparsity model. Third, we draw a new Gaussian matrix with $M = N$, quantize the measurements to β bits, again at optimal quantizer scale, and reconstruct using BPDN. Fourth, we draw a new Gaussian matrix with $M = \beta N$ and compute measurements, quantize to one bit per measurement by maintaining their sign, and perform reconstruction with BIHT. Note that the same total number of bits is used in each experiment.

Figure 8 depicts the average SNR obtained by performing 100 of the above trials. The linear, BPDN, Gaussian measurements with BPDN, and BIHT reconstructions are depicted by solid,

dashed, dash-circled, and dash-dotted lines, respectively. The linear reconstruction has a slope of 6.02dB/bit-depth, exhibiting a well-known trade-off for conventional uniform quantization. The BPDN reconstruction (without projections) follows the same trend, but obtains an SNR that is at least 10dB higher than the linear reconstruction. This is because BPDN imposes the sparse signal model to denoise the signal. We see about the same performance with the Gaussian projections at $M = N$, although it performs slightly worse than without projections since the Gaussian measurements require a slightly larger quantizer range. Similarly to the results in Fig. 7, in low Nyquist bit-depth regimes ($\beta < 6$), 1-bit CS achieves a significantly higher SNR than the other two techniques. When $6 < \beta < 8$, 1-bit CS is competitive with the BPDN scenario. This simulation demonstrates that for a fixed number of bits, 1-bit CS is competitive to conventional sampling with uniform quantization, especially in low bit-depth regimes.

6 Discussion

In this paper we have developed a rigorous mathematical foundation for 1-bit CS. Specifically, we have demonstrated a lower bound on reconstruction error as a function of the number of measurements and the sparsity of the signal. We have demonstrated that Gaussian random projections almost reach this lower bound (up to a log factor) in the noiseless case. This behavior is consistent with and extends existing results in the literature on multibit scalar quantization and 1-bit quantization of non-sparse signals.

We have also introduced reconstruction robustness guarantees through the binary ϵ-stable embedding (BϵSE) property. This property can be thought of as extending the RIP to 1-bit quantized measurements. To our knowledge, this is the first time such a property has been introduced in the context of quantization. To be able to use this property we established that a large class of random projections satisfy this property. This class is not as exhaustive as the class of matrices satisfying the RIP. Extending this property to a larger class of matrices is an interesting topic for future research.

Using the BϵSE, we have proven that 1-bit CS systems are robust to measurement noise added before quantization as well as to signals that are not exactly sparse but compressible.

We have introduced a new 1-bit CS algorithm, BIHT, that achieves better performance over previous algorithms in the noiseless case. This improvement is due to the enforcement of consistency using a one-sided linear objective, as opposed to a quadratic one. The linear objective is similar to the hinge loss from the machine learning literature.

We remind the reader that the central goal of this paper has been signal *acquisition* with quantization. As explained previously, one motivation for our work is the development of very high speed samplers. In this case, we are interested in building fast samplers by relaxing the requirements on the primary hardware burden, the quantizer. Such devices are susceptible to noise. Thus, while our noiseless results extend previous 1-bit quantization results (e.g., see [47] and [45]) to the sparse signal model setting and are of theoretical interest, a major contribution has been the further development of the robust guarantees, even if they produce error rates that seem suboptimal when compared to the noiseless case.

A number of interesting questions remain unanswered. As we discuss in Section III-B earlier, we have found that the BϵSE holds for Gaussian matrices with angular error roughly on the order of $O(\sqrt{K/M})$ worse than the optimal. One question is whether this gap can be closed with an alter-

native derivation, or whether it is a fundamental requirement for stability. Another useful pursuit would be to provide a more rigorous understanding of the discrepancy between the performance of the one-sided ℓ_1 and ℓ_2 objectives. Analysis of the performance behavior might lead to better one-sided functions.

Acknowledgments

Thanks to Rachel Ward for pointing us to the right reference with regards to the lower bound (18) and recommending several useful articles as well as Vivek Goyal for pointing us to additional prior work in this area. Thanks also to Zaiwen Wen and Wotao Yin for sharing some of the data from [31] for our algorithm comparisons, as well as engaging in numerous conversations on this topic. Finally, thanks to Nathan Srebro for his advice and discussion related to the one-sided penalty comparison and connections to binary classification.

A Lemma 1: Intersections of Orthants by Subspaces

In this section, we demonstrate that while there are 2^M available quantization points provided by 1-bit measurements, a K sparse signal will not use all of them. To understand how effectively the quantization bits are used, we first need to investigate how the K-dimensional subspaces projected from the N-dimensional K-sparse signal spaces intersect orthants in the M-dimensional measurement space.

An orthant in M dimensions is a set of points in $\mathbb{R}^M$ that all have the same sign pattern:

$$\mathcal{O}_{\boldsymbol{s}} = \{\boldsymbol{x} \mid \operatorname{sign} \boldsymbol{x} = \boldsymbol{s}\},$$

where $\boldsymbol{s}$ is a vector of ± 1. Each orthant has M boundaries of dimension $M-1$, defined as the subspace with a coordinate set to 0:

$$\mathfrak{B}_i = \{\boldsymbol{x} \mid (\boldsymbol{x})_i = 0\}.$$

We split each boundary into 2^{M-1} faces, defined as the set

$$\mathcal{F}_{i,\boldsymbol{s}} = \{\boldsymbol{x} \mid (\boldsymbol{x})_i = 0 \text{ and } \operatorname{sign}(\boldsymbol{x})_j = (\boldsymbol{s})_j \text{ for all } i \neq j\},$$

where $\boldsymbol{s}$ is the sign vector of a bordering orthant, and i is the boundary in which the face lies. Each face borders two orthants. Note that the faces are $M-1$ dimensional orthants in the $M-1$ dimensional boundary subspace. The geometry of the problem in $\mathbb{R}^3$ is summarized in Figure 9(a).

We use $I(M,K)$ to denote the maximum number of orthants in M dimensions intersected by a K-dimensional subspaces (with $I(M,1) = 2$). We upper bound $I(M,K)$ using an inductive argument that relies on the following two lemmas:

Lemma 7. *If a K-dimensional subspace $\mathcal{S} \subset \mathbb{R}^M$ is not the subset of a boundary $\mathfrak{B}_i$, then the subspace and boundary do intersect and their intersection is a $K-1$ dimensional subspace of $\mathfrak{B}_i$.*

Proof. We count the dimensions of the relevant spaces. If $\mathcal{S}$ is not a subset of $\mathfrak{B}_i$, then it equals the direct sum $\mathcal{S} = (\mathcal{S} \cap \mathfrak{B}_i) \oplus \mathcal{W}$, where $\mathcal{W} \subset \mathbb{R}^M$ is also not a subspace of $\mathfrak{B}_i$. Since $\dim \mathfrak{B}_i = M-1$, $\dim \mathcal{W} \leq 1$, and $\dim \mathcal{S} \cap \mathfrak{B}_i = K-1$ follows. □

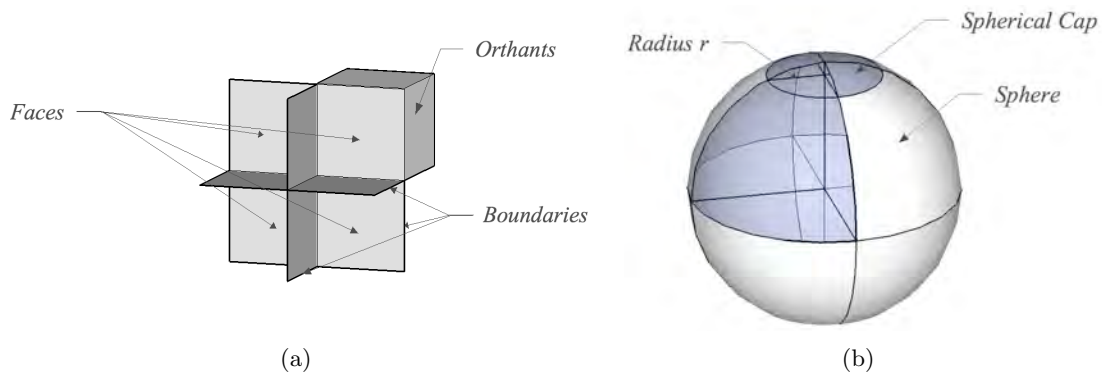


Figure 9: (a) The geometry of orthants in $\mathbb{R}^3$. (b) The geometry of spherical caps.

Lemma 8. *For $K > 1$, a K-dimensional subspace that intersects an orthant also non-trivially intersects at least K faces bordering that orthant.*

Proof. Consider a K-subspace $\mathcal{S}$, a point $\boldsymbol{p} \in \mathcal{S}$ interior to the orthant $\mathcal{O}_{\text{sign}\,\boldsymbol{p}}$, and a vector $\boldsymbol{x}_1 \in \mathcal{S}$ non-parallel to $\boldsymbol{p}$. The following iterative procedure can be used to prove the result:

1. Starting from 0, grow a until the set $\boldsymbol{p} \pm a\boldsymbol{x}_l$ intersects a boundary $\mathfrak{B}_i$, say at $a = a_l$. It is straightforward to show that as a grows, a boundary will be intersected. The point of intersection is in the face $\mathcal{F}_{i,\text{sign}\,\boldsymbol{p}}$. The set $\{\boldsymbol{p} \pm a\boldsymbol{x}_l | a \in (0, a_l)\}$ is in the orthant $\mathcal{O}_{\text{sign}\,\boldsymbol{p}}$.
2. Determine a vector $\boldsymbol{x}_{l+1} \in \mathcal{S}$ parallel to all the boundaries already intersected and not parallel to $\boldsymbol{p}$, set $l = l + 1$ and iterate from step 1.

A vector can always be found in step 2 for the first K iterations since $\mathcal{S}$ is K dimensional. The vector is parallel to all the boundaries intersected in the previous iterations and therefore $\boldsymbol{p} \pm a\boldsymbol{x}_l$ always intersects a boundary not intersected before. Therefore, at least K distinct faces are intersected. □

Lemmas 7 and 8 lead to the main result in this section. Lemma 1 in Section II-A follows trivially.

Lemma 9. *The number of orthants intersected by a K-dimensional subspace $\mathcal{S}$ in an M dimensional space $\mathcal{V}$ is upper bounded by*

$$I(M, K) \leq \binom{M}{K} 2^K.$$

Proof. The main intuition is that since the faces on each boundary are equivalent to orthants in the lower dimensional subspace of the boundary, the maximum number of faces intersected at each boundary is a problem of dimension $I(M-1, K-1)$.

If $\mathcal{S}$ is contained in one of the boundaries in $\mathcal{V}$, the number of orthants of $\mathcal{V}$ intersected is at most $I(M-1, K)$. Since $I(M, K)$ is non-decreasing in M and K, we can ignore this case in determining the upper bound.

If $\mathcal{S}$ is not contained in one of the boundaries then Lemma 7 shows that the intersection of $\mathcal{S}$ with any boundary $\mathfrak{B}_i$ is a $K-1$ dimensional subspace in $\mathfrak{B}_i$. To count the faces of $\mathfrak{B}_i$ intersected by $\mathcal{S}$ we use the observation in the definition of faces above, that each face is also an orthant of $\mathfrak{B}_i$. Therefore, the maximum number of faces of $\mathfrak{B}_i$ intersected is a recursion of the same problem in lower dimensions, i.e., is upper bounded by $I(M-1,K-1)$. Since there are M boundaries in $\mathcal{V}$, it follows that the number of faces in $\mathcal{V}$ intersected by $\mathcal{S}$ is upper bounded by $M \cdot I(M-1,K-1)$.

Using Lemma 8 we know that for an orthant to be intersected, at least K faces adjacent to it should be intersected. Since each face is adjacent to two orthants, the total number of orthants intersected cannot be greater than twice the number of faces intersected divided by K:

$$I(M,K) \leq \frac{2M \cdot I(M-1,K-1)}{K}. \tag{17}$$

The result follows by induction. □

A tighter achievable bound is also known [63, 64]:

$$\begin{aligned} I(M,K) &\leq 2\sum_{l=0}^{K-1}\binom{M-1}{l} \\ &\leq 2K\binom{M-1}{K-1} \text{ if } K \leq \frac{M-1}{2}. \end{aligned} \tag{18}$$

Although (18) is tighter and achieved with a subspace in a general configuration, it leads to expressions on the same asymptotical order of our main results. We use (17) for the remainder of this paper because of its simpler form.

B Theorem 1: Distributing Signals to Quantization Points

To prove Theorem 1 we consider how the available quantization points optimally cover the set of signals of interest. We consider unit norm signals that belong in a union of L subspaces, each of dimension K. Thus the set of interest is the union of L unit spheres of K dimensions.

First we need to understand how to measure the sets of signals of interest. We denote the unit sphere in K dimensions—which is the surface of the K-dimensional unit ball—using S^{K-1}, and the rotationally invariant area measure on the sphere using $\sigma(\cdot)$. Thus the area of the whole sphere is equal to $\sigma(S^{K-1})$. If subspaces intersect, the area of the sphere inside the intersection has measure zero. Therefore, the total surface area of all L spheres is LS^{K-1}.

The most efficient cover of this area is achieved if every point covers a spherical cap of radius r, denoted using $C(r)$. The geometry of the problem is demonstrated in Figure 9(b). From [65] the surface area of a spherical cap of radius r satisfies

$$\sigma(C(r)) \leq r^K \sigma(S^{K-1}),$$

For $L\binom{M}{K}2^K$ points to cover the area $L\sigma(S^{K-1})$ we require

$$\begin{aligned} L\binom{M}{K}2^K\sigma(C(r)) \geq L\sigma(S^{K-1}) &\Rightarrow \left(\frac{Me2r}{K}\right)^K \geq 1 \\ &\Rightarrow r \geq \frac{K}{2eM} = \Omega\left(K/M\right), \end{aligned}$$

using the bound $\binom{M}{K} \leq (eM/K)^K$. Incidentally, this proof gives an obvious solution to a Grassmanian covering problem of 1-dimensional subspaces in K dimensional spaces. Although Grassmanian packing problems have been examined in the literature (e.g., in the context of frame theory [66]), to our knowledge, the Grassmanian covering problem has not been posed or attempted.

C Theorem 2: Optimal Performance via Gaussian Projections

To prove Theorem 2, we follow the procedure given in [44, Theorem 3.3]. We begin by restricting our analysis to the support set $T \subset \{1, \cdots, N\}$ with $|T| \leq D \leq N$, and thus we consider vectors that lie on the (sub) sphere $\Sigma^*(T) = \{\boldsymbol{x} : \operatorname{supp} \boldsymbol{x} \subset T, \|\boldsymbol{x}\|_2 = 1\} \subset \mathbb{R}^N$. We remind the reader that $B_\delta(\boldsymbol{x}) := \{\boldsymbol{a} \in S^{N-1} : \|\boldsymbol{x} - \boldsymbol{a}\|_2 < \delta\}$ is the ball of unit norm vectors of Euclidean distance δ around $\boldsymbol{x}$, and we write $B^*_\delta(\boldsymbol{x}) = B_\delta(\boldsymbol{x}) \cap \Sigma^*(T)$ as in Section III-B.

Given a vector $\boldsymbol{\varphi} \sim \mathcal{N}^{N\times 1}(0,1)$ and two distinct points $\boldsymbol{p}$ and $\boldsymbol{q}$ in Q_δ, we have that

$$\mathbb{P}\big[\forall\, \boldsymbol{u} \in B^*_\delta(\boldsymbol{p}), \forall \boldsymbol{v} \in B^*_\delta(\boldsymbol{q}) : \operatorname{sign} \boldsymbol{\varphi}^T \boldsymbol{u} \neq \operatorname{sign} \boldsymbol{\varphi}^T \boldsymbol{v}\big] \;\geq\; d_S(\boldsymbol{p}, \boldsymbol{q}) \;-\; \sqrt{\tfrac{\pi}{2} D}\, \delta,$$

from Lemma 10 (given in Section A). When $\epsilon_o > 2\delta$, we have the relationship

$$\pi\, d_S(\boldsymbol{p}, \boldsymbol{q}) \geq 2 \sin(\frac{\pi}{2}\, d_S(\boldsymbol{p}, \boldsymbol{q})) = \|\boldsymbol{p} - \boldsymbol{q}\|_2 \geq \|\boldsymbol{u} - \boldsymbol{v}\|_2 - 2\delta > \epsilon_o - 2\delta,$$

and thus

$$\mathbb{P}\big[\forall\, \boldsymbol{u} \in B^*_\delta(\boldsymbol{p}), \forall \boldsymbol{v} \in B^*_\delta(\boldsymbol{q}) : \operatorname{sign} \boldsymbol{\varphi}^T \boldsymbol{u} \neq \operatorname{sign} \boldsymbol{\varphi}^T \boldsymbol{v} \mid \|\boldsymbol{u} - \boldsymbol{v}\|_2 > \epsilon_o\big] \;\geq\; \tfrac{\epsilon_o}{\pi} \;-\; (\tfrac{2}{\pi} + \sqrt{\tfrac{\pi}{2} D})\, \delta.$$

By setting $\delta = \pi\epsilon_o/(4 + \pi\sqrt{2\pi D})$ (and reversing the inequality), we obtain

$$\mathbb{P}\big[\exists\, \boldsymbol{u} \in B^*_\delta(\boldsymbol{p}), \exists \boldsymbol{v} \in B^*_\delta(\boldsymbol{q}) : \operatorname{sign}(\boldsymbol{\varphi}^T \boldsymbol{u}) = \operatorname{sign}(\boldsymbol{\varphi}^T \boldsymbol{v}) \mid \|\boldsymbol{u} - \boldsymbol{v}\|_2 > \epsilon_o\big] \;\leq\; 1 - \tfrac{\epsilon_o}{2}.$$

Thus, for M different random vectors $\boldsymbol{\varphi}_i$ arranged in $\Phi = (\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_M)^T \sim \mathcal{N}^{M\times N}(0,1)$, and for the associated mapping A defined in (3), we have that

$$\mathbb{P}\big[\exists\, \boldsymbol{u} \in B^*_\delta(\boldsymbol{p}), \exists \boldsymbol{v} \in B^*_\delta(\boldsymbol{q}) : A(\boldsymbol{u}) = A(\boldsymbol{v}) \mid \|\boldsymbol{u} - \boldsymbol{v}\|_2 > \epsilon_o\big] \leq (1 - \tfrac{\epsilon_o}{2})^M.$$

In words, we have found a bound on the probability that two vectors' measurements are consistent, even if their euclidean distance is greater than ϵ_o, but only for vectors in the restricted (sub) sphere $\Sigma^*(T)$. Now we seek to cover the rest of the space Σ^*_K (unit norm K-sparse signals).

Given a radius $\delta > 0$, the sphere $\Sigma^*(T)$ can be covered with a finite set $Q_\delta \subset \Sigma^*(T)$ of no more than $(3/\delta)^D$ points such that, for any $\boldsymbol{w} \in \Sigma^*(T)$, there exists a $\boldsymbol{q} \in Q_\delta$ with $\boldsymbol{w} \in B^*_\delta(\boldsymbol{q})$ [15]. Since there are no more than $\binom{|Q_\delta|}{2} \leq (|Q_\delta|)^2 \leq (3/\delta)^{2D}$ pairs of distinct points in Q_δ, we find

$$\mathbb{P}\big[\exists\, \boldsymbol{u}, \boldsymbol{v} \in \Sigma^*(T) : d_H(A(\boldsymbol{u}), A(\boldsymbol{v})) = 0 \mid \|\boldsymbol{u} - \boldsymbol{v}\|_2 > \epsilon_o\big] \;\leq\; \big(\tfrac{1}{\pi\epsilon_o}(12 + 3\pi\sqrt{2\pi D})\big)^{2D}\, (1 - \tfrac{\epsilon_o}{2})^M.$$

To obtain the final bound, we observe that any pair of unit K-sparse vectors $\boldsymbol{x}$ and $\boldsymbol{s}$ in Σ^*_K belongs to some $\Sigma^*(T)$ with $T = \operatorname{supp} \boldsymbol{x} \cup \operatorname{supp} \boldsymbol{s}$ and $|T| \leq 2K$. There are no more than $\binom{N}{2K} \leq (N/2K)^{2K}$ of such sets T, and thus setting $D = 2K$ above yields

$$\begin{aligned}
&\mathbb{P}\big[\exists\, \boldsymbol{u}, \boldsymbol{v} \in \Sigma^*_K : d_H(A(\boldsymbol{u}), A(\boldsymbol{v})) = 0 \mid \|\boldsymbol{u} - \boldsymbol{v}\|_2 > \epsilon_o\big] \\
&\leq\; (\tfrac{N}{2K})^{2K}\; (\tfrac{1}{\pi\epsilon_o}(12 + 6\pi\sqrt{\pi K}))^{4K}\; (1 - \tfrac{\epsilon_o}{2})^M \\
&\leq \exp\big[2K \log(\tfrac{N}{2K}) + 4K \log(\tfrac{1}{\pi\epsilon_o}(12 + 6\pi\sqrt{\pi K})) - M\tfrac{\epsilon_o}{2}\big],
\end{aligned}$$

where the second inequality follows from $1 - \frac{\epsilon_o}{2} \leq \exp \frac{\epsilon_o}{2}$. By upper bounding this probability by η and solving for M, we obtain

$$M \geq \tfrac{1}{\epsilon_o} \left(2K \, \log \tfrac{N}{2K} + 4K \log(\tfrac{1}{\pi \epsilon_o}(12 + 6\pi\sqrt{\pi K})) + \log \tfrac{1}{\eta} \right).$$

Since $K \geq 1$, we have that $\frac{1}{\pi}(12 + 6\pi\sqrt{\pi K}) < 12\sqrt{\pi K} < 16\sqrt{2K}$, and thus the previous relation is then satisfied when

$$\begin{aligned} M \;&\geq\; \tfrac{1}{\epsilon_o} \left(2K \, \log \tfrac{N}{2K} + 4K \log(\tfrac{1}{\epsilon_o}(16\sqrt{2K})) + \log \tfrac{1}{\eta} \right) \\ &=\; \tfrac{1}{\epsilon_o} \left(2K \, \log \tfrac{N}{2} + 4K \log(\tfrac{1}{\epsilon_o}(16\sqrt{2})) + \log \tfrac{1}{\eta} \right) \\ &=\; \tfrac{1}{\epsilon_o} \left(2K \, \log N + 4K \log(\tfrac{16}{\epsilon_o}) + \log \tfrac{1}{\eta} \right). \end{aligned}$$

D Lemma 4: Concentration of Measure for δ-Balls

Proving Lemma 4 amounts to showing that, for some fixed $\epsilon > 0$ and $0 \leq \delta \leq 1$, given a Gaussian matrix $\Phi \in \mathbb{R}^{M \times D}$, the mapping $A : \mathbb{R}^D \to \mathcal{B}^M$ defined as $A(\boldsymbol{u}) = \operatorname{sign}(\Phi \boldsymbol{u})$, and for some $\boldsymbol{x}, \boldsymbol{s} \in S^{D-1}$, we have

$$\mathbb{P}\Big(\Big| d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;-\; d_S(\boldsymbol{x}, \boldsymbol{s}) \Big| \;\leq\; \epsilon + \sqrt{\tfrac{\pi}{2} D}\, \delta \Big) \;\geq\; 1 - 2\, e^{-2\epsilon^2 M}, \quad \forall \boldsymbol{u} \in B^*_\delta(\boldsymbol{x}),\ \forall \boldsymbol{v} \in B^*_\delta(\boldsymbol{s}),$$

where the balls B_δ are also restricted to $\mathbb{R}^D$.

Given $\boldsymbol{u} \in B^*_\delta(\boldsymbol{x})$ and $\boldsymbol{v} \in B^*_\delta(\boldsymbol{s})$, the quantity $M d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big)$ is the sum $\sum_i A_i(\boldsymbol{u}) \oplus A_i(\boldsymbol{v})$, where $A_i(\boldsymbol{u})$ stands for the i^{th} component of $A(\boldsymbol{u})$. For one index $1 \leq i \leq M$

$$\begin{aligned} A_i(\boldsymbol{u}) \oplus A_i(\boldsymbol{v}) \;&\leq\; Z_i^+ := \max\big\{ A_i(\boldsymbol{p}) \oplus A_i(\boldsymbol{q}) :\; \boldsymbol{p} \in B^*_\delta(\boldsymbol{x}), \boldsymbol{q} \in B^*_\delta(\boldsymbol{s}) \big\}, \\ A_i(\boldsymbol{u}) \oplus A_i(\boldsymbol{v}) \;&\geq\; Z_i^- := \min\big\{ A_i(\boldsymbol{p}) \oplus A_i(\boldsymbol{q}) :\; \boldsymbol{p} \in B^*_\delta(\boldsymbol{x}), \boldsymbol{q} \in B^*_\delta(\boldsymbol{s}) \big\}, \end{aligned}$$

and therefore

$$Z^- \;:=\; \sum_{i=1}^M Z_i^- \;\leq\; M\, d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;\leq\; \sum_{i=1}^M Z_i^+ \;=:\; Z^+.$$

Of course, the occurrence of $Z_i^+ = 0$ ($Z_i^- = 1$) means that all vector pairs taken separately in $B^*_\delta(\boldsymbol{x})$ and $B^*_\delta(\boldsymbol{s})$ have consistent (or respectively, inconsistent) measurements on the i^{th} sensing component A_i. More precisely, since $\boldsymbol{\varphi}_i \sim \mathcal{N}^{N \times 1}(0, 1)$, $Z_i^\pm$ are binary random variables such that $\mathbb{P}[Z_i^+ = 1] = 1 - p_0$ and $\mathbb{P}[Z_i^- = 1] = p_1$ independently of i, where the probabilities p_0 and p_1 are defined by

$$\begin{aligned} p_0(d_S(\boldsymbol{x}, \boldsymbol{s}), \delta) \;&=\; \mathbb{P}[Z_i^+ = 0] \;=\; \mathbb{P}\big[\, \forall \boldsymbol{p} \in B^*_\delta(\boldsymbol{x}), \forall \boldsymbol{q} \in B^*_\delta(\boldsymbol{s}),\ A_i(\boldsymbol{u}) = A_i(\boldsymbol{v}) \,\big], \\ p_1(d_S(\boldsymbol{x}, \boldsymbol{s}), \delta) \;&=\; \mathbb{P}\big[\, \forall \boldsymbol{p} \in B^*_\delta(\boldsymbol{x}), \forall \boldsymbol{q} \in B^*_\delta(\boldsymbol{s}),\ A_i(\boldsymbol{u}) \neq A_i(\boldsymbol{v}) \,\big]. \end{aligned}$$

In summary, Z^+ and Z^- are binomially distributed with M trials and probability of success $1 - p_0$ and p_1, respectively. Furthermore, we have that $\mathbb{E}Z^+ = M\,(1 - p_0)$ and $\mathbb{E}Z^- = M\, p_1$, thus by the Chernoff-Hoeffding inequality,

$$\begin{aligned} \mathbb{P}\big[\, Z^+ \;>\; M\,(1 - p_0) + M\epsilon \,\big] \;&\leq\; e^{-2M\epsilon^2}, \\ \mathbb{P}\big[\, Z^- \;<\; M\, p_1 - M\epsilon \,\big] \;&\leq\; e^{-2M\epsilon^2}. \end{aligned}$$

This indicates that with a probability higher than $1 - 2e^{-2M\epsilon^2}$, we have

$$p_1 - \epsilon \;\leq\; d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;\leq\; (1 - p_0) + \epsilon.$$

The final result follows by lower bounding p_0 and p_1 as in Lemma 10.

Lemma 10. *Given $0 \leq \delta < 1$ and two unit vectors $\boldsymbol{x}, \boldsymbol{s} \in S^{D-1}$, we have*

$$p_0 = \mathbb{P}\big[\,\forall \boldsymbol{u} \in B_\delta(\boldsymbol{x}),\ \forall \boldsymbol{v} \in B_\delta(\boldsymbol{s}),\, \operatorname{sign}\langle \boldsymbol{\varphi}, \boldsymbol{u}\rangle = \operatorname{sign}\langle \boldsymbol{\varphi}, \boldsymbol{v}\rangle\,\big] \;\geq\; 1 - d_S(\boldsymbol{x}, \boldsymbol{s}) - \sqrt{\tfrac{\pi}{2}D}\,\delta, \quad (19)$$

$$p_1 = \mathbb{P}\big[\,\forall \boldsymbol{u} \in B_\delta(\boldsymbol{x}),\ \forall \boldsymbol{v} \in B_\delta(\boldsymbol{s}),\, \operatorname{sign}\langle \boldsymbol{\varphi}, \boldsymbol{u}\rangle \neq \operatorname{sign}\langle \boldsymbol{\varphi}, \boldsymbol{v}\rangle\,\big] \;\geq\; d_S(\boldsymbol{x}, \boldsymbol{s}) - \sqrt{\tfrac{\pi}{2}D}\,\delta. \quad (20)$$

Proof of Lemma 10. We begin by introducing some useful properties of Gaussian vector distribution. If $\boldsymbol{\varphi} \sim \mathcal{N}^{D\times 1}(0,1)$, the probability that $\boldsymbol{\varphi} \in \mathcal{A} \subset \mathbb{R}^D$ is simply the measure μ of $\mathcal{A}$ with respect to the standard Gaussian density $\gamma(\boldsymbol{\varphi}) = \frac{1}{(2\pi)^{D/2}}\, e^{-\|\boldsymbol{\varphi}\|^2/2}$, i.e.,

$$\mathbb{P}[\,\boldsymbol{\varphi} \in \mathcal{A}\,] = \mu(\mathcal{A}) = \int_{\mathcal{A}} \mathrm{d}^D\boldsymbol{\varphi}\ \gamma(\boldsymbol{\varphi}),$$

with $\mu(\mathbb{R}^D) = 1$. It may be easier to perform this integration over a hyper-spherical set of coordinates. Specifically, we let any vector $\boldsymbol{\varphi}$ be represented by the values $(r, \phi_1, \cdots, \phi_{D-1})$ where $r \in \mathbb{R}_+$ stands for the vector length, $\phi_1, \cdots, \phi_{D-2} \in [0, \pi]$ corresponds to the vector angles in each dimension, and $\phi_{D-1} \in [0, 2\pi]$ being the angle of $\boldsymbol{\varphi}$ in the "$\boldsymbol{xs}$" plane. This is possible since γ is rotionally invariant and thus we may assume the "$\boldsymbol{xs}$" plane is spanned by the canonical vectors $\boldsymbol{e}_D = \boldsymbol{x}$ and $\boldsymbol{e}_{D-1}$ in the canonical basis $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_D\}$ of $\mathbb{R}^D$, with $\boldsymbol{e}_1 = (\boldsymbol{x} \wedge \boldsymbol{s}) \,/\, \|\boldsymbol{x} \wedge \boldsymbol{s}\|_2$ and $\boldsymbol{e}_{D-1} = \boldsymbol{e}_D \wedge \boldsymbol{e}_1$.

The change of coordinates is then defined as $\varphi_1 = r\cos\phi_1$, $\varphi_2 = r\sin\phi_1\cos\phi_2$, ..., $\varphi_{D-1} = r\,\sin\phi_1 \cdots \sin\phi_{D-2}\cos\phi_{D-1}$, and $\varphi_D = r\,\sin\phi_1 \cdots \sin\phi_{D-2}\sin\phi_{D-1}$, while, conversely, $r = \|\boldsymbol{\varphi}\|_2$, $\tan\phi_1 = (\varphi_D^2 + \cdots + \varphi_2^2)^{1/2}/\varphi_1$, ..., $\tan\phi_{D-2} = (\varphi_D^2 + \varphi_{D-1}^2)^{1/2} \,/\, \varphi_{D-2}$, and $\tan\phi_{D-1} = \varphi_D \,/\, \varphi_{D-1}$.[8]

We now seek a lower bound on p_1. Computing this probability amounts to estimating

$$p_1 = \mathbb{P}[\,\forall \boldsymbol{u} \in B_\delta(\boldsymbol{x}),\ \forall \boldsymbol{v} \in B_\delta(\boldsymbol{s}),\ \langle \boldsymbol{\varphi}, \boldsymbol{u}\rangle\langle \boldsymbol{\varphi}, \boldsymbol{v}\rangle \leq 0\,] \;=\; \mu(\mathcal{W}_\delta),$$

where $\mathcal{W}_\delta = \{\boldsymbol{\varphi} : \langle \boldsymbol{\varphi}, \boldsymbol{u}\rangle\langle \boldsymbol{\varphi}, \boldsymbol{v}\rangle \leq 0,\ \forall \boldsymbol{u} \in B_\delta(\boldsymbol{x}),\ \forall \boldsymbol{v} \in B_\delta(\boldsymbol{s})\}$ is the set of all vectors $\boldsymbol{\varphi}$ such that its inner product with $\boldsymbol{u}$ and $\boldsymbol{v}$ result in different signs. Note that if $B_\delta(\boldsymbol{x}) \cap B_\delta(\boldsymbol{s})$ is not empty, then we have $p_1 = 0$ since for $\boldsymbol{w} \in B_\delta(\boldsymbol{x}) \cap B_\delta(\boldsymbol{s})$, we have $\langle \boldsymbol{\varphi}, \boldsymbol{w}\rangle^2$. This term cannot be negative and thus $\mathcal{W}_\delta = \{\boldsymbol{\varphi} : \langle \boldsymbol{\varphi}, \boldsymbol{w}\rangle = 0\}$, which has measure zero with respect to μ. In order to avoid this trouble, we must choose $d_S(\boldsymbol{x}, \boldsymbol{s}) \geq \frac{4}{\pi}\arcsin\delta/2$. Furthermore, since $\arcsin\lambda \leq \frac{\pi}{2}\lambda$ for any $0 \leq \lambda \leq 1$, this occurs if $d_S(\boldsymbol{x}, \boldsymbol{s}) \geq \delta$.

The remainder of the proof is devoted to finding an appropriate way to integrate the set $\mathcal{W}_\delta$. To this end, we begin by demonstrating that estimating p_1 can be simplified with the following equivalence (proved just after the completion of the proof of Lemma 10).

Lemma 11. *The set $\mathcal{W}_\delta \subset \mathbb{R}^D$ is equal to the set*

$$\mathcal{V}_\delta^- \;=\; \{\boldsymbol{\varphi} : \langle \boldsymbol{\varphi}, \boldsymbol{x}\rangle\langle \boldsymbol{\varphi}, \boldsymbol{s}\rangle \leq 0, \|\boldsymbol{x} - \mathcal{P}_{\Pi(\varphi)}\,\boldsymbol{x}\| \geq \delta, \|\boldsymbol{s} - \mathcal{P}_{\Pi(\varphi)}\,\boldsymbol{s}\| \geq \delta\},$$

where $\mathcal{P}_{\Pi(\boldsymbol{\varphi})}$ is the orthogonal projection on the plane $\Pi(\boldsymbol{\varphi}) = \{\boldsymbol{u} \in \mathbb{R}^D : \langle \boldsymbol{\varphi}, \boldsymbol{u}\rangle = 0\}$.

[8] This change of coordinates can be very convenient. For instance, the proof of Lemma 3 relies on the computation $\mathbb{P}[A_i(\boldsymbol{x}) \neq A_i(\boldsymbol{s})] = \mu(\mathcal{A} = \{\boldsymbol{\varphi} : \phi_{D-1} \in [0, \pi\, d_S(\boldsymbol{x}, \boldsymbol{s})] \cup [\pi, \pi + \pi\, d_S(\boldsymbol{x}, \boldsymbol{s})]\}) = d_S(\boldsymbol{x}, \boldsymbol{s})$, since for (almost) all $\boldsymbol{\varphi} \in \mathcal{A}$, $\boldsymbol{x}$ and $\boldsymbol{s}$ live in the two different subvolumes determined by the plane $\{\boldsymbol{u} : \langle \boldsymbol{\varphi}, \boldsymbol{u}\rangle = 0\}$ [51, 52].

Using the hyper spherical coordinate system developed earlier, membership in $\mathcal{V}_\delta^-$ can be expressed as

$$\boldsymbol{\varphi} = (r, \phi_1, \cdots, \phi_{D-1}) \in \mathcal{V}_\delta^- \;\Leftrightarrow\; \begin{cases} \tan\phi_{D-1} \;\in\; [0, \tan\theta], & \text{(R1)} \\ \sin\phi_1 \cdots \sin\phi_{D-2}\,|\sin\phi_{D-1}| \;\geq\; \delta, & \text{(R2)} \\ \sin\phi_1 \cdots \sin\phi_{D-2}\,|\sin(\phi_{D-1}-\theta)| \;\geq\; \delta. & \text{(R3)} \end{cases}$$

Indeed, requirement (R1) enforces $\langle\boldsymbol{\varphi}, \boldsymbol{x}\rangle\langle\boldsymbol{\varphi}, \boldsymbol{s}\rangle \leq 0$, while (R2) and (R3) are direct translations of the requirements that $\|\boldsymbol{x} - \mathcal{P}_{\Pi(\varphi)}\,\boldsymbol{x}\| = |\langle\widehat{\boldsymbol{\varphi}}, \boldsymbol{x} = \boldsymbol{e}_D\rangle| \geq \delta$ and $\|\boldsymbol{s} - \mathcal{P}_{\Pi(\varphi)}\,\boldsymbol{s}\| = |\langle\widehat{\boldsymbol{\varphi}}, \boldsymbol{y} = -\sin\theta\,\boldsymbol{e}_D + \cos\theta\,\boldsymbol{e}_{D-1}\rangle| \geq \delta$, with $\widehat{\boldsymbol{\varphi}} = \frac{1}{\|\boldsymbol{\varphi}\|}\boldsymbol{\varphi}$.

We are now ready to integrate to find p_1:

$$p_1 = \mu(\mathcal{V}_\delta^-) \;= \tfrac{1}{(2\pi)^{D/2}} \int_{\mathbb{R}_+} \mathrm{d}r\; r^{D-1} e^{-r^2/2} \left[\Big(\int_0^\pi \mathrm{d}\phi_1\; \sin^{D-2}\phi_1\Big)\cdots\Big(\int_0^\pi \mathrm{d}\phi_{D-2}\; \sin\phi_{D-2}\Big)\right]\cdots$$
$$\Big[\int_{[0,\theta]\,\cup\,[\pi,\pi+\theta]} \mathrm{d}\phi_{D-1}\; \chi_{g(\delta,\boldsymbol{\varphi})}(\phi_{D-1})\,\chi_{g(\delta,\boldsymbol{\varphi})}(\phi_{D-1}-\theta)\Big],$$

with $\chi_\lambda(\phi) = 1$ if $|\sin\phi| \geq \lambda$ and 0 else, for some $\lambda \in [0,1]$, and $g(\delta, \boldsymbol{\varphi}) = \delta/(\sin\phi_1 \cdots \sin\phi_{D-2})$.

However,

$$\int_{[0,\theta]\,\cup\,[\pi,\pi+\theta]} \mathrm{d}\phi\; \chi_\lambda(\phi)\,\chi_\lambda(\phi-\theta) \;=\; 2\theta - 4\arcsin\lambda \;\geq\; 2\theta - 2\pi\lambda,$$

since $\lambda \leq \arcsin\lambda \leq \frac{\pi}{2}\lambda$ for any $\lambda \in [0,1]$. Consequently,

$$\begin{aligned}\mu(\mathcal{V}_\delta^-) \geq\; & \tfrac{1}{(2\pi)^{D/2}} \int_{\mathbb{R}_+} \mathrm{d}r\; r^{D-1} e^{-r^2/2} \;\cdots \\ & \left[\Big(\int_0^\pi \mathrm{d}\phi_1\; \sin^{D-2}\phi_1\Big)\cdots\Big(\int_0^\pi \mathrm{d}\phi_{D-2}\; \sin\phi_{D-2}\Big)\right]\Big(2\theta - \tfrac{2\pi\delta}{(\sin\phi_1\cdots\sin\phi_{D-2})}\Big) \\ =\; & \frac{\theta}{\pi} \;-\; \frac{\pi\,\delta\; I_{D-3}\, I_{D-4}\cdots I_0}{I_{D-2}\, I_{D-3}\, I_{D-4}\cdots I_0} \;=\; \frac{\theta}{\pi} \;-\; \frac{\pi\,\delta}{I_{D-2}},\end{aligned}$$

with $I_n = \int_0^\pi \mathrm{d}\phi\; \sin^n\phi$ and knowing that $(2\pi)^{D/2} = 2(I_{D-2}\cdots I_0)\int_{\mathbb{R}_+} \mathrm{d}r\; r^{D-1}e^{-r^2/2}$.

Using the fact that $I_n = \sqrt{\pi}\,\Gamma(\frac{n+1}{2})/\Gamma(\frac{n}{2}+1) \geq \sqrt{\pi}/\sqrt{\frac{n}{2}+\frac{1}{4}}$, we obtain $I_{D-2} \geq \frac{\sqrt{\pi}}{\sqrt{\frac{D}{2}-\frac{3}{4}}} \geq \sqrt{\frac{2\pi}{D}}$, and thus

$$p_1 \;\geq\; d_S(\boldsymbol{x}, \boldsymbol{s}) \;-\; \sqrt{\tfrac{\pi}{2}D}\,\delta.$$

If we want a meaningful bound for $p_1 \geq 0$, then we must have $d_S(\boldsymbol{x}, \boldsymbol{s}) \geq \sqrt{\frac{\pi}{2}d}\,\delta \geq \delta$. Therefore, as soon as the lower bound is positive, the aforementioned condition $d_S(\boldsymbol{x}, \boldsymbol{s}) \geq \delta$ always holds.

The lower bound for p_0 is obtained similarly. It is straightforward to show that $p_0 = \mu(\mathcal{V}_\delta^+)$, with $\mathcal{V}_\delta^+ = \{\boldsymbol{\varphi} : \langle\boldsymbol{\varphi}, \boldsymbol{x}\rangle\langle\boldsymbol{\varphi}, \boldsymbol{s}\rangle > 0, \|\boldsymbol{x} - \mathcal{P}_{\Pi(\varphi)}\,\boldsymbol{x}\| \geq \delta, \|\boldsymbol{y} - \mathcal{P}_{\Pi(\varphi)}\,\boldsymbol{s}\| \geq \delta\}$. Lower bounding $\mu(\mathcal{V}_\delta^+)$ as for $\mu(\mathcal{V}_\delta^+)$, the only difference occurring with the integral on ϕ_{D-2} given by

$$\int_{[\theta,\pi]\,\cup\,[\pi+\theta,2\pi]} \mathrm{d}\phi_{D-1}\; \chi_{g(\delta,\boldsymbol{\varphi})}(\phi_{D-1})\,\chi_{g(\delta,\boldsymbol{\varphi})}(\phi_{D-1}-\theta) \;\cdots$$
$$= \;2\pi - 2\theta - 4\arcsin g(\delta,\boldsymbol{\varphi}) \;\geq\; 2(\pi-\theta) - 2\pi g(\delta,\boldsymbol{\varphi}).$$

Therefore, the lower bound of p_0 amounts to change $\theta \to \pi - \theta$ in the one of p_1, which provides the result. $\square$

Proof of Lemma 11. If $\delta = 0$, there is nothing to prove. Therefore $\delta > 0$ and if $\boldsymbol{\varphi}^*$ belongs to either $\mathcal{V}_\delta$ or $\mathcal{W}_\delta$, we must have $\langle \boldsymbol{\varphi}, \boldsymbol{x} \rangle \langle \boldsymbol{\varphi}, \boldsymbol{s} \rangle < 0$. It is also sufficient to work on the restriction of $\mathcal{V}_\delta$ and $\mathcal{W}_\delta$ to unit vectors.

(i) $\mathcal{V}_\delta \subset \mathcal{W}_\delta$: By contradiction, let us assume that $\boldsymbol{\varphi}^* \in \mathcal{V}_\delta$ but $\boldsymbol{\varphi}^* \notin \mathcal{W}_\delta$. Without any loss of generality, $\langle \boldsymbol{\varphi}^*, \boldsymbol{x} \rangle > 0$ and $\langle \boldsymbol{\varphi}^*, \boldsymbol{s} \rangle < 0$. Since $\boldsymbol{\varphi}^* \notin \mathcal{W}_\delta$, there exist two vectors $\boldsymbol{u}^* \in B_\delta(\boldsymbol{x})$ and $\boldsymbol{v}^* \in B_\delta(\boldsymbol{y})$ such that $\langle \boldsymbol{\varphi}^*, \boldsymbol{u}^* \rangle \langle \boldsymbol{\varphi}^*, \boldsymbol{v}^* \rangle > 0$. If $\langle \boldsymbol{\varphi}^*, \boldsymbol{u}^* \rangle > 0$ and $\langle \boldsymbol{\varphi}^*, \boldsymbol{v}^* \rangle > 0$, then, since $\langle \boldsymbol{\varphi}^*, \boldsymbol{s} \rangle < 0$ and by continuity of the inner product, there exist a $\lambda \in (0,1)$ such that $\langle \boldsymbol{\varphi}^*, \boldsymbol{s}(\lambda) \rangle = 0$ with $\boldsymbol{s}(\lambda) = \boldsymbol{y} + \lambda(\boldsymbol{v}^* - \boldsymbol{s})$. Therefore, $\boldsymbol{s}(\lambda) \in \Pi(\boldsymbol{\varphi})$ and, by definition of the orthogonal projection, $\|\boldsymbol{s} - \mathcal{P}_{\Pi\varphi}\, \boldsymbol{s}\| \leq \|\boldsymbol{s} - \boldsymbol{s}(\lambda)\| \leq \lambda\delta < \delta$ which is a contradiction. If $\langle \boldsymbol{\varphi}^*, \boldsymbol{u}^* \rangle < 0$ and $\langle \boldsymbol{\varphi}^*, \boldsymbol{v}^* \rangle < 0$, we apply the same reasoning on $\boldsymbol{x}$ and $\boldsymbol{u}^*$. Therefore, $\mathcal{V}_\delta \subset \mathcal{W}_\delta$.

(ii) $\mathcal{W}_\delta \subset \mathcal{V}_\delta$: If $\boldsymbol{\varphi}^* \in \mathcal{W}_\delta$ with $\boldsymbol{\varphi}^* \notin \mathcal{V}_\delta$, we have either $\|\boldsymbol{x} - \mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{x}\| < \delta$ or $\|\boldsymbol{s} - \mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{s}\| < \delta$. Let us say that $\|\boldsymbol{x} - \mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{x}\| < \delta$. Then, for $\boldsymbol{w} = \boldsymbol{x} + \delta\, (\mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{x} - \boldsymbol{x}) / \|\mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{x} - \boldsymbol{x}\| \in B^*_\delta(\boldsymbol{x})$, $\langle \boldsymbol{\varphi}^*, \boldsymbol{x} \rangle \langle \boldsymbol{\varphi}^*, \boldsymbol{w} \rangle = (\langle \boldsymbol{\varphi}^*, \boldsymbol{x} \rangle)^2 \big(1 - \delta / \|\mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{x} - \boldsymbol{x}\|\big) + \delta\, \langle \boldsymbol{\varphi}^*, \mathcal{P}_{\Pi(\varphi^*)}\, \boldsymbol{x} \rangle < 0$. However, $\boldsymbol{\varphi}^* \in \mathcal{W}_\delta$ and $\langle \boldsymbol{\varphi}^*, \boldsymbol{x} \rangle \langle \boldsymbol{\varphi}^*, \boldsymbol{s} \rangle < 0$, leading to $\langle \boldsymbol{\varphi}^*, \boldsymbol{w} \rangle \langle \boldsymbol{\varphi}^*, \boldsymbol{s} \rangle > 0$, which is a contradiction. $\square$

E Theorem 3: Gaussian Matrices Provide BϵSEs

The strategy for proving Theorem 3 will be to count the number of pairs of K-sparse signals that are Euclidean distance δ apart. We will then apply the concentration results of Lemma 4 to demonstrate that the angles between these pairs are approximately preserved. We specifically proceed by focusing on a single K-dimensional subspace (intersected with the unit sphere) and then by applying a union bound to account for all possible subspaces.

Let $T \subset \{1, \ldots, N\}$ be an index set of size $|T| = K$, $\Sigma^*(T) = \{\boldsymbol{w} \in \mathbb{R}^N : \operatorname{supp} \boldsymbol{w} \subset T, \quad \|\boldsymbol{w}\|_2 = 1\}$ be the sphere of unit vectors with support T. We first use again the fact that the sphere $\Sigma^*(T)$ can be δ-covered by a finite set of points $Q_{T,\delta}$. That is, for any $\boldsymbol{w} \in \Sigma^*(T)$, there exists a $\boldsymbol{q} \in Q_{T,\delta}$ such that $\boldsymbol{w} \in B^*_\delta(\boldsymbol{q}) = B_\delta(\boldsymbol{q}) \cap \Sigma^*_T = \{\boldsymbol{w}' \in \Sigma^*_T : \|\boldsymbol{w}' - \boldsymbol{q}\|_2 \leq \delta\}$ [15]. Note that the size of $Q_{T,\delta}$ is bounded by $|Q_{T,\delta}| \leq C_\delta = (3/\delta)^K$.

Let Φ_T be the matrix formed by the columns of Φ indexed by T and note that $\Phi_T \boldsymbol{w} = \Phi \boldsymbol{w}$. Since $\epsilon \geq 0$ is given, then for all pairs of points $\boldsymbol{x}, \boldsymbol{y} \in Q_{T,\delta}$, we have

$$\mathbb{P}\Big(\Big| d_H\big(A(\boldsymbol{p}), A(\boldsymbol{q})\big) \;-\; d_S(\boldsymbol{x}, \boldsymbol{y}) \Big| \;\leq\; \epsilon + \sqrt{\tfrac{\pi}{2} K}\, \delta \Big) \;\geq\; 1 \;-\; 2\, (\tfrac{3}{\delta})^{2K}\, e^{-2\epsilon^2 M}, \tag{21}$$

for all $\boldsymbol{p} \in B^*_\delta(\boldsymbol{x})$ and $\boldsymbol{q} \in B^*_\delta(\boldsymbol{y})$. This follows from Lemma 4 with $D = K$, since Φ_T is a Gaussian matrix and by invoking the union bound, since there are $\binom{C_\delta}{2} \leq C^2_\delta = (3/\delta)^{2K}$ such pairs $\boldsymbol{x}, \boldsymbol{y}$.

The bound (21) can be extended to all possible index sets T of size K via the union bound. Specifically, for all $T \subset \{1, \cdots, N\}$ and all pairs of points $\boldsymbol{x}, \boldsymbol{y} \in Q_{T,\delta}$, we have

$$\mathbb{P}\Big(\Big| d_H\big(A(\boldsymbol{p}), A(\boldsymbol{q})\big) \;-\; d_S(\boldsymbol{x}, \boldsymbol{y}) \Big| \;\leq\; \epsilon + \sqrt{\tfrac{\pi}{2} K}\, \delta \Big) \;\geq\; 1 \;-\; 2\, (\tfrac{eN}{K})^K\, (\tfrac{3}{\delta})^{2K}\, e^{-2\epsilon^2 M} \tag{22}$$

for all $\boldsymbol{p} \in B^*_\delta(\boldsymbol{x})$ and $\boldsymbol{q} \in B^*_\delta(\boldsymbol{y})$, since there are no more than $\binom{N}{K} \leq (eN/K)^K$ possible T.

To summarize, for any points on the sphere $\boldsymbol{u}, \boldsymbol{v} \in S^{N-1}$ with $|\operatorname{supp} \boldsymbol{u} \cup \operatorname{supp} \boldsymbol{v}| \leq K$, there exists an index set T of size K such that $\boldsymbol{u}, \boldsymbol{v} \in \Sigma^*(T)$ and from (22) there exists two points $\boldsymbol{x}, \boldsymbol{y} \in Q_{T,\delta}$ such that $\boldsymbol{u} \in B^*_\delta(\boldsymbol{x})$ and $\boldsymbol{v} \in B^*_\delta(\boldsymbol{y})$ with a probability exceeding $1 - 2\,(\frac{eN}{K})^K\,(\frac{3}{\delta})^{2K}\,e^{-2\epsilon^2 M}$. Furthermore, when this occurs we have

$$\Big|\, d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;-\; d_S(\boldsymbol{x}, \boldsymbol{y})\,\Big| \;\leq\; \epsilon + \sqrt{\tfrac{\pi}{2}K}\,\delta. \tag{23}$$

To obtain our final bound, consider that $\boldsymbol{u} \in B^*_\delta(\boldsymbol{x})$ implies that $\pi\, d_S(\boldsymbol{u}, \boldsymbol{x}) \leq 2 \arcsin \delta/2 \leq \pi\delta/2$, and $d_S(\boldsymbol{v}, \boldsymbol{y})$ can be similarly bounded. Thus, $d_S(\boldsymbol{u}, \boldsymbol{v}) \geq d_S(\boldsymbol{x}, \boldsymbol{y}) - \delta$ and $d_S(\boldsymbol{u}, \boldsymbol{v}) \leq d_S(\boldsymbol{x}, \boldsymbol{y}) + \delta$, and (23) becomes

$$\Big|\, d_H\big(A(\boldsymbol{u}), A(\boldsymbol{v})\big) \;-\; d_S(\boldsymbol{u}, \boldsymbol{v})\,\Big| \;\leq\; \epsilon + (1 + \sqrt{\tfrac{\pi}{2}K})\,\delta. \tag{24}$$

By bounding the probability of failure as $2\,(\frac{eN}{K})^K\,(\frac{3}{\delta})^{2K}\,e^{-2\epsilon^2 M} \leq \eta$, where $0 < \eta < 1$, and setting $\epsilon = (1 + \sqrt{\frac{\pi}{2}K})\,\delta$, solving for M, we obtain

$$M \;\geq\; \tfrac{4}{\epsilon^2}\big(K\,\log(\tfrac{9eN}{K}) + 2K\,\log(\tfrac{2(1+\sqrt{2\pi K})}{\epsilon}) + \log(\tfrac{2}{\eta})\big).$$

Since $K \geq 1$, we have that $2(1 + \sqrt{2\pi K}) < 4\sqrt{2\pi K}$, and thus the previous relation is satisfied if

$$\begin{aligned} M \;&\geq\; \tfrac{4}{\epsilon^2}\big(K\,\log(\tfrac{9eN}{K}) + 2K\,\log(\tfrac{4\sqrt{2\pi K}}{\epsilon}) + \log(\tfrac{2}{\eta})\big), \\ &=\; \tfrac{4}{\epsilon^2}\big(K\,\log(9eN) + 2K\,\log(\tfrac{4\sqrt{2\pi}}{\epsilon}) + \log(\tfrac{2}{\eta})\big), \\ &=\; \tfrac{4}{\epsilon^2}\big(K\,\log(N) + 2K\,\log(\tfrac{12\sqrt{2\pi e}}{\epsilon}) + \log(\tfrac{2}{\eta})\big), \end{aligned}$$

which can be further simplified to $M \;\geq\; \tfrac{4}{\epsilon^2}\big(K\,\log(N) + 2K\,\log(\tfrac{50}{\epsilon}) + \log(\tfrac{2}{\eta})\big)$.

F Lemma 5: Stability with Measurement Noise

In Lemma 5, since $\Phi \sim \mathcal{N}^{M\times N}(0,1)$, each $y_i = (\Phi\boldsymbol{x})_i$ follows a Gaussian distribution $\mathcal{N}(0, \|\boldsymbol{x}\|_2^2)$, and furthermore, since we have independent additive noise, $z_i = y_i + n_i = (\Phi\boldsymbol{x})_i + n_i$ follows the Gaussian distriubtion $\mathcal{N}(0, \|\boldsymbol{x}\|_2^2 + \sigma^2)$.

We begin by bounding the probability that any noisy measurement z_i has a different sign than the original corresponding measurement y_i, i.e., we bound $p_0 = \mathbb{P}(z_i y_i < 0)$. This quantity is interesting since $M\, d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big)$ follows a Binomial distribution with M trials and probability of success p_0 and thus we also have $\mathbb{E}\big(d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big)\big) = p_0$.

To solve for the bound, we compute

$$p_0 = \int_{\mathbb{R}} \mathrm{d}u\; \mathbb{P}(z_i y_i < 0 \,|\, y_i = u\,)\, f_{y_i}(u) \;=\; \int_{\mathbb{R}} \mathrm{d}u\; \mathbb{P}(u^2 + u n_i < 0)\, g(u; \|\boldsymbol{x}\|_2),$$

with the pdf $f_{y_i}(t) = g(t;\sigma') = \frac{1}{\sqrt{2\pi}t}\exp(-t^2/2\sigma'^2)$. This leads to

$$\begin{aligned} p_0 &= \int_0^\infty \mathrm{d}u\ \mathbb{P}(n_i < -u)\, g(u;\|\boldsymbol{x}\|_2) \ + \ \int_{-\infty}^0 \mathrm{d}u\ \mathbb{P}(n_i > -u)\, g(u;\|\boldsymbol{x}\|_2) \\ &= \int_0^\infty \mathrm{d}u\ 2\,Q(u/\sigma)\, g(u;\|\boldsymbol{x}\|_2) \ \leq \ \int_0^\infty \mathrm{d}u\ e^{-\frac{u^2}{2\sigma^2}}\, g(u;\|\boldsymbol{x}\|_2) \\ &= \ \tfrac{1}{\sqrt{2\pi}\|\boldsymbol{x}\|_2}\int_0^\infty \mathrm{d}u\ e^{-\frac{1}{2}(\frac{(\|\boldsymbol{x}\|_2^2+\sigma^2)\,u^2}{\sigma^2\|\boldsymbol{x}\|_2^2})} \ = \ \tfrac{1}{2}\frac{\sigma}{\sqrt{\|\boldsymbol{x}\|_2^2+\sigma^2}}, \end{aligned}$$

where $Q(u) = \int_u^\infty \mathrm{d}t\, g(t;1)$ denotes the tail integral of the standard Gaussian distribution which is bounded by $Q(t) \leq \frac{1}{2}e^{-t^2/2}$ for $t \geq 0$ (see for instance [67, Eq. (13.48)]).

Thus, we have $p_0 \leq e(\sigma, \|\boldsymbol{x}\|_2) = \frac{1}{2}\frac{\sigma}{\sqrt{\|\boldsymbol{x}\|_2^2+\sigma^2}}$ and, by applying the Chernoff-Hoeffding inequality to the distribution of $d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big)$,

$$\begin{aligned} \mathbb{P}\big[M\, d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big) > \ M\, e(\sigma,\|\boldsymbol{x}\|_2) + M\epsilon\big] & \\ \leq \ \mathbb{P}\big[M\, d_H\big(A_{\boldsymbol{n}}(\boldsymbol{x}), A(\boldsymbol{x})\big) > \ M\, p_0 + M\epsilon\big] & \\ \leq \ e^{-2M\epsilon^2}, & \end{aligned}$$

which proves the lemma.

G Corollary 2: Stability with Compressible Signals

The proof of Corollary 2 is as follows. Since $\boldsymbol{x} = \boldsymbol{x}_K + (\boldsymbol{x} - \boldsymbol{x}_K)$ then $\Phi\boldsymbol{x} = \Phi\boldsymbol{x}_K + \boldsymbol{n}$ where $\boldsymbol{n} = \Phi(\boldsymbol{x} - \boldsymbol{x}_K)$ is a random Gaussian vector. Thus $A(\boldsymbol{x}) = A_{\boldsymbol{n}}(\boldsymbol{x}_K)$ where A_n is defined as in Lemma 5. The vector $\boldsymbol{n}$ is also independent of $\Phi\boldsymbol{x}_K$ since the supports of $\boldsymbol{x}_K$ and $(\boldsymbol{x} - \boldsymbol{x}_K)$ are disjoint. Moreover, applying Lemma 6.1 of [13], we have that

$$\|\boldsymbol{n}\|_2 \ \leq \ \sqrt{M}\rho(\boldsymbol{x}, K),$$

if $\frac{1}{\sqrt{M}}\Phi$ is a RIP(K, δ_K) matrix. Finally, the variance σ^2 of each i.i.d. component n_i of $\boldsymbol{n}$ can be bounded by $\sigma^2 = \mathbb{E}n_i^2 = \mathbb{E}\|\boldsymbol{n}\|_2^2/M \leq \rho(\boldsymbol{x}, K)$, thus the result follows from Lemma 5 with the bound $e(\sigma, \|\boldsymbol{x}_K\|_2) \leq \frac{1}{2}\sigma/\|\boldsymbol{x}_K\|_2 \leq \frac{\rho(\boldsymbol{x},K)}{2\|\boldsymbol{x}_K\|_2}$.

References

[1] E. Candès, "Compressive sampling," in *Proc. Int. Congress Math.*, Madrid, Spain, Aug. 2006.

[2] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 6, no. 4, pp. 1289–1306, 2006.

[3] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse, bandlimited signals," *to appear in IEEE Trans. Inform. Theory*, 2009.

[4] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 83–91, 2008.

[5] B. K. Natarajan, "Sparse Approximate Solutions to Linear Systems," *SIAM J. on Comp.*, vol. 24, pp. 227, 1995.

[6] S. S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. on Sci. Comp.*, vol. 20, no. 1, pp. 33–61, 1998.

[7] E. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus de l'Académie des Sciences, Série I*, vol. 346, no. 9-10, pp. 589–592, 2008.

[8] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[9] E. T. Hale, W. Yin, and Y. Zhang, "A fixed-point continuation method for ℓ_1 regularized minimization with applications to compressed sensing," in *Preprint*, 2007.

[10] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. of Sel. Top. in Sig. Proc.*, Sept. 2007.

[11] 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, 2007.

[12] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comput. Math.*, vol. 9, no. 3, pp. 317–334, 2009.

[13] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.

[14] T. Blumensath and M. Davies, "Iterative hard thresholding for compressive sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[15] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Const. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.

[16] M. Davenport, *Random observations on random observations: Sparse signal acquisition and processing*, Ph.d. thesis, Rice University, Aug. 2010.

[17] J. Romberg, "Compressive sensing by random convolution," *SIAM J. Imaging Sciences*, 2009.

[18] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

[19] L. Jacques, P. Vandergheynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici, "Cmos compressed imaging by random convolution," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, 2009, pp. 1113–1116, doi:10.1109/ICASSP.2009.4959783.

[20] L. Jacques, D. Hammond, and M. Fadili, "Dequantizing compressed sensing: When oversampling and non-gaussian contraints combine," *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 559–571, 2011.

[21] J. Laska, P. Boufounos, M. Davenport, and R. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *to appear in App. Comp. and Harm. Anal. (ACHA)*, 2011.

[22] W. Dai, H. Pham, and O. Milenkovic, "Distortion-rate functions for quantized compressive sensing," *Preprint*, 2009.

[23] A. Zymnis, S. Boyd, and E. Candès, "Compressed sensing with quantized measurements," *IEEE Signal Processing Letters*, vol. 17, no. 2, Feb. 2010.

[24] J. Sun and V. Goyal, "Quantization for compressed sensing reconstruction," in *Proc. Sampling Theory and Applications (SampTA)*, Marseille, France, May 2009.

[25] G. Gray and G. Zeoli, "Quantization and saturation noise due to analog-to-digital conversion," *IEEE Trans. Aerospace and Elec. Systems*, vol. 7, no. 1, pp. 222–223, 1971.

[26] R. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Selected Areas in Comm.*, vol. 17, no. 4, pp. 539–550, 1999.

[27] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, "Analog-to-digital converters," *IEEE Signal Processing Mag.*, Nov. 2005.

[28] P. Boufounos, "Reconstruction of sparse signals from distorted randomized measurements," in *Proc. Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, Mar. 2010.

[29] P. Boufounos and R. Baraniuk, "1-bit compressive sensing," in *Proc. Conf. Inform. Science and Systems (CISS)*, Princeton, NJ, Mar. 2008.

[30] P. Boufounos, "Greedy sparse signal reconstruction from sign measurements," in *Proc. Asilomar Conf. on Signals Systems and Comput.*, Asilomar, California, Nov. 2009.

[31] J. Laska, Z. Wen, W. Yin, and R. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *Preprint*, 2010.

[32] P. Boufounos, *Quantization and erasures in frame representations*, Ph.D. thesis, MIT EECS, Cambridge, MA, Jan. 2006.

[33] P. Boufounos and R. Baraniuk, "Quantization of sparse representations," in *Proc. data compression conference (DCC)*, Mar. 2007, p. 378.

[34] J. Z. Sun and V. K. Goyal, "Optimal quantization of random measurements in compressed sensing," in *International Symposium on Information Theory (ISIT)*, June 2009.

[35] N. Thao and M. Vetterli, "Reduction of the mse in R-times oversampled A/D conversion $O(1/R)$ to $O(1/R^2)$," *IEEE Trans. Signal Processing*, vol. 42, no. 1, pp. 200–203, 2994.

[36] V. K. Goyal, M. Vetterli, and N. Thao, "Quantized overcomplete expansions in $\mathbb{R}^n$: Analysis, synthesis, and algorithms," *IEEE Trans. Inform. Theory*, vol. 44, no. 1, pp. 16–31, 1998.

[37] P. M. Aziz, H. V. Sorensen, and J. Van Der Spiegel, "An overview of Sigma-Delta converters," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 61–84, Jan. 1996.

[38] N. T. Thao, "Vector quantization analysis of $\Sigma\Delta$ modulation," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 808–817, Apr. 1996.

[39] J. C. Candy and G. C. Temes, Eds., *Oversampling Delta-Sigma Converters*, IEEE Press, 1992.

[40] J. J. Benedetto, A. M. Powell, and O. Yilmaz, "Sigma-delta quantization and finite frames," *IEEE Trans. Info. Theory*, vol. 52, no. 5, pp. 1990–2005, May 2006.

[41] P. Boufounos and A. V. Oppenheim, "Quantization noise shaping on arbitrary frame expansions," *EURASIP Journal on Applied Signal Processing, Special issue on Frames and Overcomplete Representations in Signal Processing, Communications, and Information Theory*, vol. 2006, pp. Article ID 53807, 12 pages, DOI:10.1155/ASP/2006/53807, 2006.

[42] P. Boufounos and R. Baraniuk, "Sigma Delta Quantization for Compressive Sensing," in *Proc. SPIE Waveletts XII. Proc. of SPIE Vol. 6701, 670104*, San Diego, CA, Aug. 2007.

[43] S. Gunturk, A. Powell, R. Saab, and O. Ylmaz, "Sobolev Duals for Random Frames and Sigma-Delta Quantization of Compressed Sensing Measurements," *preprint*, Feb. 2010.

[44] P. Boufounos, "Universal rate-efficient scalar quantization," *Preprint*, 2010, Submitted.

[45] H. Q. Nguyen, V.K. Goyal, and L.R. Varshney, "Frame Permutation Quantization," *Applied and Computational Harmonic Analysis (ACHA)*, Nov. 2010.

[46] P. Indyk and R. Motwani, "Approximate nearest neighbors: towards removing the curse of dimensionality," in *Proceedings of the thirtieth annual ACM symposium on Theory of computing*. ACM, 1998, pp. 604–613.

[47] A. Andoni, M. Datar, N. Immorlica, P. Indyk, and V. Mirrokni, "Locality-sensitive hashing scheme based on p-stable distributions," *Nearest neighbor methods in learning and vision: Theory and practice (book)*, 2006.

[48] A. Andoni and P. Indyk, "Near-optimal hashing algorithms for approximate nearest nieghbor in high dimensions," *Commun. ACM*, vol. 51, no. 1, pp. 117–122, 2008.

[49] M. Raginsky and S. Lazebnik, "Locality-sensitive binary codes from shift-invariant kernels," *The Neural Information Processing Systems*, vol. 22, 2009.

[50] M. Bridson, "Geometric and combinatorial group theory," in *The Princeton companion to mathematics*, T. Gowers, J. Barrow-Green, and I. Leader, Eds., chapter IV.10, pp. 431–447. Princeton Univ. Press, 2008.

[51] M. Goemans and D. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journ. ACM*, vol. 42, no. 6, pp. 1145, 1995.

[52] S. Shariati, L. Jacques, F. X. Standaert, B. Macq, M. A. Salhi, and P. Antoine, "Randomly driven fuzzy key extraction of unclonable images," in *IEEE Int. conf. on image proc. (ICIP)*, 2010.

[53] A. Gupta, B. Recht, and R. Nowak, "Sample complexity for 1-bit compressed sensing and sparse classification," in *Proc. Intl. Symp. on Information Theory (ISIT)*, 2010.

[54] T. Blumensath, "Compressed sensing with nonlinear observations," *Preprint*, 2010.

[55] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics, 2nd edition, Feb. 2009.

[56] P. J. Huber, "Robust regression: Asymptotics, conjectures, and Monte Carlo," *Statistics*, vol. 1, pp. 799–821, 1973.

[57] L. Rosasco, E. De Vito, A. Caponnetto, M. Piana, and A. Verri, "Are loss functions all the same?," *Neural Comp.*, vol. 16, no. 5, pp. 1063–1076, Mar. 2004.

[58] N. Srebro, K. Sridharan, and A. Tewari, "Smoothness, low-noise, and fast rates," *Preprint*, 2010.

[59] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[60] O. Bartlett, Y. Freund, W. S. Lee, and R. E. Schapire, "Boosting the margin: a new explanation for the effectiveness of voting methods," *Annals of Statistics*, vol. 26, no. 5, pp. 1651–1686, 1998.

[61] G. Ratsch and M. Warmuth, "Efficient margin maximizing with boosting," *The J. of Machine Learning Research*, vol. 6, Dec. 2005.

[62] Avrim Blum, *On-Line Algorithms in Machine Learning*, vol. 1442/1998 of *Lecture Notes in Computer Science*, Springer, 1998.

[63] T. M. Cover, "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition," *IEEE Trans. on Electronic Comp.*, vol. 3, pp. 326–334, Jun. 1965.

[64] L. Flatto, "A new proof of the transposition theorem," *Proc. American Mathematical Society*, vol. 24, no. 1, pp. 29–31, jan 1970.

[65] K. Ball, "An elementary introduction to modern convex geometry," in *Flavors of Geometry*, Silvio Levy, Ed., vol. 31 of *MSRI Publications*, pp. 1–58. Cambridge University Press, Cambridge, UK, 1997.

[66] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," *Applied and Computational Harmonic Analysis*, vol. 14, no. 3, pp. 257–275, 2003.

[67] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions, Volume 1*, Wiley, 1994.