UC Irvine
UC Irvine Previously Published Works

Title
One-Bit Compressed Sensing by Linear Programming

Permalink
https://escholarship.org/uc/item/0mf8d97t

Journal
Communications on Pure and Applied Mathematics, 66(8)

ISSN
0010-3640

Authors
Plan, Yaniv
Vershynin, Roman

Publication Date
2013-08-01

DOI
10.1002/cpa.21442

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, availalbe at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

arXiv:1109.4299v5 [cs.IT] 16 Mar 2012

ONE-BIT COMPRESSED SENSING BY LINEAR PROGRAMMING

YANIV PLAN AND ROMAN VERSHYNIN

Abstract. We give the first computationally tractable and almost optimal solution to the problem of one-bit compressed sensing, showing how to accurately recover an s-sparse vector $\boldsymbol{x} \in \mathbb{R}^n$ from the signs of $O(s \log^2(n/s))$ random linear measurements of $\boldsymbol{x}$. The recovery is achieved by a simple linear program. This result extends to approximately sparse vectors $\boldsymbol{x}$. Our result is universal in the sense that with high probability, one measurement scheme will successfully recover all sparse vectors simultaneously. The argument is based on solving an equivalent geometric problem on random hyperplane tessellations.

1. Introduction

Compressed sensing is a modern paradigm of data acquisition, which is having an impact on several disciplines, see [21]. The scientist has access to a measurement vector $\boldsymbol{v} \in \mathbb{R}^m$ obtained as

$$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}, \tag{1.1}$$

where $\boldsymbol{A}$ is a given $m \times n$ measurement matrix and $\boldsymbol{x} \in \mathbb{R}^n$ is an unknown signal that one needs to recover from $\boldsymbol{v}$. One would like to take $m \ll n$, rendering $\boldsymbol{A}$ non-invertible; the key ingredient to successful recovery of $\boldsymbol{x}$ is take into account its assumed structure – sparsity. Thus one assumes that $\boldsymbol{x}$ has at most s nonzero entries, although the support pattern is unknown. The strongest known results are for random measurement matrices $\boldsymbol{A}$. In particular, if $\boldsymbol{A}$ has Gaussian i.i.d. entries, then we may take $m = O(s \log(n/s))$ and still recover $\boldsymbol{x}$ exactly with high probability [10, 7]; see [26] for an overview. Furthermore, this recovery may be achieved in polynomial time by solving the convex minimization program

$$\min \|\boldsymbol{x}'\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x}' = \boldsymbol{v}. \tag{1.2}$$

Stability results are also available when noise is added to the problem [9, 8, 3, 27].

However, while the focus of compressed sensing is signal recovery with minimal information, the classical set-up (1.1), (1.2) assumes *infinite bit precision* of the measurements. This disaccord raises an important question: how many bits per measurement (i.e. per coordinate of $\boldsymbol{v}$) are sufficient for tractable and accurate sparse recovery? This paper shows that *one bit per measurement* is enough.

There are many applications where such severe quantization may be inherent or preferred — analog-to-digital conversion [20, 18], binomial regression in statistical modeling and threshold group testing [12], to name a few.

1.1. Main results.

This paper demonstrates that a simple modification of the convex program (1.2) is able to accurately estimate $\boldsymbol{x}$ from extremely quantized measurement vector

$$\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x}).$$

Date: September 19, 2011.

2000 *Mathematics Subject Classification.* 94A12; 60D05; 90C25.

Y.P. is supported by an NSF Postdoctoral Research Fellowship under award No. 1103909. R.V. is supported by NSF grants DMS 0918623 and 1001829.

Here $\boldsymbol{y}$ is the vector of signs of the coordinates of $\boldsymbol{A}\boldsymbol{x}$.[1]

Note that $\boldsymbol{y}$ contains no information about the magnitude of $\boldsymbol{x}$, and thus we can only hope to recover the normalized vector $\boldsymbol{x}/\|\boldsymbol{x}\|_2$. This problem was introduced and first studied by Boufounos and Baraniuk [6] under the name of *one-bit compressed sensing*; some related work is summarized in Section 1.2.

We shall show that the signal can be accurately recovered by solving the following convex minimization program

$$\min \|\boldsymbol{x}'\|_1 \quad \text{subject to} \quad \operatorname{sign}(\boldsymbol{A}\boldsymbol{x}') \equiv \boldsymbol{y} \quad \text{and} \quad \|\boldsymbol{A}\boldsymbol{x}'\|_1 = m. \tag{1.3}$$

The first constraint, $\operatorname{sign}(\boldsymbol{A}\boldsymbol{x}') \equiv \boldsymbol{y}$, keeps the solution consistent with the measurements. It is defined by the relation $\langle \boldsymbol{a}_i, \boldsymbol{x}' \rangle \cdot \boldsymbol{y}_i \geq 0$ for $i = 1, 2, \ldots, m$, where $\boldsymbol{a}_i$ is the i-th row of $\boldsymbol{A}$. The second constraint, $\|\boldsymbol{A}\boldsymbol{x}'\|_1 = m$, serves to prevent the program from returning a zero solution. Moreover, this constraint is linear as it can be represented as one linear equation $\sum_{i=1}^m y_i \langle \boldsymbol{a}_i, \boldsymbol{x}' \rangle = m$ where y_i denote the coordinates of $\boldsymbol{y}$. Therefore (1.3) is indeed a convex minimization program; furthermore one can easily represent it as a linear program, see (5.3) below. Note also that the number m in (1.3) is chosen for convenience of the analysis; it can be replaced by any other fixed positive number.

Theorem 1.1 (Recovery from one-bit measurements)**.** *Let $n, m, s > 0$, and let $\boldsymbol{A}$ be an $m \times n$ random matrix with independent standard normal entries. Set*

$$\delta = C \left(\frac{s}{m} \log(2n/s) \log(2n/m + 2m/n) \right)^{1/5}. \tag{1.4}$$

Then, with probability at least $1 - C\exp(-c\delta m)$, the following holds uniformly for all signals $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \leq \sqrt{s}$. Let $\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x})$. Then the solution $\hat{\boldsymbol{x}}$ of the convex minimization program (1.3) *satisfies*

$$\left\| \frac{\hat{\boldsymbol{x}}}{\|\hat{\boldsymbol{x}}\|_2} - \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} \right\|_2 \leq \delta.$$

Here and thereafter C and c denote positive absolute constants; other standard notation is explained in Section 1.3.

Remark 1 (Effective sparsity). The Cauchy-Schwarz inequality implies that $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \leq \sqrt{\|\boldsymbol{x}\|_0}$ where $\|\boldsymbol{x}\|_0 = |\operatorname{supp}(\boldsymbol{x})|$ is the number of nonzero elements of $\boldsymbol{x}$. Therefore one can view the parameter $(\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2)^2$ as a measure of *effective sparsity* of the signal $\boldsymbol{x}$. The effective sparsity is thus a real valued and robust extension of the sparsity parameter $\|\boldsymbol{x}\|_0$, which allows one to handle approximately sparse vectors.

Let us then state the partial case of Theorem 1.1 for sparse signals:

Corollary 1.2 (Sparse recovery from one-bit measurements)**.** *Let $n, m, s > 0$, and set δ as in* (1.4)*. Then, with probability at least $1 - C\exp(-c\delta m)$, the following holds uniformly for all signals $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\|x\|_0 \leq s$. Let $\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x})$. Then the solution $\hat{\boldsymbol{x}}$ of the convex minimization program* (1.3) *satisfies*

$$\left\| \frac{\hat{\boldsymbol{x}}}{\|\hat{\boldsymbol{x}}\|_2} - \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} \right\|_2 \leq \delta.$$

[1] To be precise, for a scalar $z \neq 0$ we define $\operatorname{sign}(z) = z/|z|$, and $\operatorname{sign}(0) = 0$. We allow the sign function to act on a vector by acting individually on each element.

Remark 2 (Number of measurements). The conclusion of Corollary 1.2 can be stated in the following useful way. With high probability, *an arbitrarily accurate estimation of every s-sparse vector $\boldsymbol{x}$ can be achieved from*

$$m = O(s \log^2(n/s))$$

one-bit random measurements. The implicit factor in the $O(\cdot)$ notation depends only on the desired accuracy level δ; more precisely $m \sim \delta^{-5} s \log^2(n/s)$ up to an absolute constant factor. The same holds if $\boldsymbol{x}$ is only effectively s-sparse as in Theorem 1.1. The central point here is that the number of measurements is almost linear in the sparsity s, which can be much smaller than the ambient dimension n.

Remark 3 (Non-gaussian measurements). Most results in compressed sensing, and in random matrix theory in general, are valid not only for Gaussian random matrices but also for general random matrix ensembles. In one-bit compressed sensing, since the measurements $\operatorname{sign}(\boldsymbol{A}\boldsymbol{x})$ do not depend on the scaling of the rows of $\boldsymbol{A}$, it is clear that our results will not change if the rows of $\boldsymbol{A}$ are sampled independently from *any rotationally invariant distribution* in $\mathbb{R}^n$ (for example, the uniform distribution on the unit Euclidean sphere S^{n-1}).

However, in contrast to the widespread universality phenomenon, one-bit compressed sensing cannot be generalized to some of the simplest *discrete distributions*, such as Bernoulli. Indeed, suppose the entries of $\boldsymbol{A}$ are independent ± 1 valued symmetric random variables. Then for the vectors $\boldsymbol{x} = (1, 0, 0, \ldots, 0)$ and $\boldsymbol{x}' = (1, \frac{1}{2}, 0, \ldots, 0)$ one can easily check that $\operatorname{sign}(\boldsymbol{A}\boldsymbol{x}) = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x}')$ for any number of measurements m. So one-bit measurements can not distinguish between two fixed distinct signals $\boldsymbol{x}$ and $\boldsymbol{x}'$ no matter how many measurements are taken.

Remark 4 (Optimality). For a fixed level of accuracy, our estimate on the number of measurements $m = O(s \log^2(n/s))$ matches the best known number of measurements in the classical (not quantized) compressed sensing problem up to the exponent 2 of the logarithm, and up to an absolute constant factor. However, we believe that the exponent 2 can be reduced to 1. We also believe that the error δ in Theorem 1.1 may decrease more quickly as $s/m \to 0$. In particular, Jacques et al. [18] demonstrate that $\boldsymbol{x}$ is exactly sparse and is estimated using an ℓ_0-minimization-based approach, the error is upper bounded as $\delta = O((s/m)^{1-o(1)} \log n)$; they also demonstrate a lower error bound $\delta = \Omega(s/m)$ regardless of what algorithm is used. In fact, such a result is not possible when $\boldsymbol{x}$ is only known to be effectively sparse (i.e., $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \leq \sqrt{s}$). Instead, the best possible bound is of the form $\delta = O(\sqrt{(s/m)\log(n/s)})$ (this can be checked via entropy arguments). We believe this is achievable (and is optimal) for the convex program (1.3).

1.2. **Prior work.** While there have been several numerical results for quantized compressed sensing [6, 4, 5, 20, 28], as well as guarantees on the convergence of many of the algorithms used for these numerical results, theoretical accuracy guarantees have been much less developed. One may endeavor to circumvent this problem by considering quantization errors as a source of noise, thereby reducing the quantized compressed sensing problem to the noisy classical compressed sensing problem. Further, in some cases the theory and algorithms of noisy compressed sensing may be adapted to this problem as in [28, 11, 17, 25]; the method of quantization may be specialized in order to minimize the recovery error. As noted in [19] if the range of the signal is unspecified, then such a noise source is unbounded, and so the classical theory does not apply. However, in the setup of our paper we may assume without loss of generality that $\|\boldsymbol{x}\|_2 = 1$, and thus it is possible that the methods of Candes and Tao [8] can be adapted to derive a version of Corollary 1.2 for a fixed sparse signal $\boldsymbol{x}$. Nevertheless, we do not see any way to deduce by these methods a uniform result over all sparse signals $\boldsymbol{x}$.

In a complementary line of research Ardestanizadeh et al. [2] consider compressed sensing with a finite number of bits per measurement. However, the number of bits per measurement there is not

one (or constant); this number depends on the sparsity level s and the dynamic range of the signal $\boldsymbol{x}$. Similarly, in the work of Gunturk et al. [14, 15] on sigma-delta quantization, the number of bits per measurement depends on the dynamic range of $\boldsymbol{x}$. On the other hand, by considering sigma-delta quantization and multiple bits, the Gunturk et al. are able to provide excellent guarantees on the speed of decay of the error δ as s/m decreases.

The framework of one-bit compressed sensing was introduced by Boufounos and Baraniuk in [6]. Jacques et al. [18] show that $O(s \log n)$ one-bit measurements are sufficient to recover an s-sparse vector with arbitrary precision; their results are also robust to bit flips. In particular, their results require the estimate $\hat{\boldsymbol{x}}$ to be as sparse as $\boldsymbol{x}$, have unit norm, and be consistent with the data. The difficulty is that the first two of these constraints are non-convex, and thus the only known program which is known to return such an estimate is ℓ_0 minimization with the unit norm constraint—this is generally considered to be intractable. Gupta et al. [16] demonstrate that one may tractably recover the support of $\boldsymbol{x}$ from $O(s \log n)$ measurements. They give two measurement schemes. One is non-adaptive, but the number of measurements has a quadratic dependence on the dynamic range of the signal. The other has no such dependence but is adaptive. Our results settle several of these issues: (a) we make no assumption about the dynamic range of the signal, (b) the one-bit measurements are non-adaptive, and (c) the signal is recovered by a tractable algorithm (linear programming).

1.3. **Notation and organization of the paper.** Throughout the paper, C, c, C_1, etc. denote absolute constants whose values may change from line to line. For integer n, we denote $[n] = \{1, \dots, n\}$. Vectors are written in bold italics, e.g., $\boldsymbol{x}$, and their coordinates are written in plain text so that the i-th component of $\boldsymbol{x}$ is x_i. For a subset $T \subset [n]$, $\boldsymbol{x}_T$ is the vector $\boldsymbol{x}$ restricted to the elements indexed by T. The ℓ_1 and ℓ_2 norms of a vector $\boldsymbol{x} \in \mathbb{R}^n$ are defined as $\|\boldsymbol{x}\|_1 = \sum_{i=1}^n |x_i|$ and $\|\boldsymbol{x}\|_2 = (\sum_{i=1}^n x_i^2)^{1/2}$ respectively. The number of non-zero coordinates of $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|_0 = |\operatorname{supp}(\boldsymbol{x})|$. The unit balls with respect to ℓ_1 and ℓ_2 norms are denoted by $B_1^n = \{x \in \mathbb{R}^n : \|x\|_1 \le 1\}$ and $B_2^n = \{x \in \mathbb{R}^n : \|x\|_2 \le 1\}$ respectively. The unit Euclidean sphere is denoted $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$.

The rest of the paper is devoted to proving Theorem 1.1. In Section 2 we reduce this task to the following two ingredients: (a) Theorem 2.3 which states states that *a solution to* (1.3) *is effectively sparse*, and (b) Theorem 2.2 which analyzes a simpler but non-convex version of (1.3) where the constraint $\|\boldsymbol{A}\boldsymbol{x}'\|_1 = m$ is replaced by $\|\boldsymbol{x}'\|_2 = 1$. The latter result can be interpreted in a geometric way in terms of *random hyperplane tessellations* of a subset K of the Euclidean sphere, specifically for the set of effectively sparse signals $K = S^{n-1} \cap \sqrt{s} B_1^n$. In Section 3 we estimate the metric entropy of K, and we use this in Section 4 to prove our main geometric result of independent interest: $m = O(s \log(n/s))$ *random hyperplanes are enough to cut K into small pieces*, yielding that all cells of the resulting tessellation have arbitrarily small diameter. This will complete part (b) above. For part (a), we prove Theorem 2.3 on the effective sparsity of solutions in Section 5. The proof is based on counting all possible solutions of (1.3), which are the vertices of the feasible polytope. This will allow us to use standard concentration inequalities from the Appendix and to conclude the argument by a union bound.

Acknowledgement. The authors are grateful to Sinan Güntürk for pointing out an inaccuracy in the statement of Lemma 3.4 in an earlier version of this paper.

2. Strategy of the proof

Our proof of Theorem 1.1 has two main ingredients which we explain in this section. Throughout the paper, $\boldsymbol{a}_i$ will denote the rows of $\boldsymbol{A}$, which are i.i.d. standard normal vectors in $\mathbb{R}^n$.

Let us revisit the second constraint $\|\boldsymbol{A}\boldsymbol{x}'\|_1 = m$ in the convex minimization program (1.3). Consider a fixed signal $\boldsymbol{x}'$ for the moment. Taking the expectation with respect to the random matrix $\boldsymbol{A}$, we see that

$$\mathbb{E}\,\|\boldsymbol{A}\boldsymbol{x}'\|_1 = \sum_{i=1}^{m} \mathbb{E}\,|\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle| = cm\|\boldsymbol{x}'\|_2$$

where $c = \sqrt{2/\pi}$. Here we used that the first absolute moment of the standard normal random variable equals c. So *in expectation*, the constraint $\|\boldsymbol{A}\boldsymbol{x}'\|_1 = m$ is equivalent to $\|\boldsymbol{x}'\|_2 = 1$ up to constant factor c.

This observation suggests that we may first try to analyze the simpler minimization program

$$\min \|\boldsymbol{x}'\|_1 \quad \text{subject to} \quad \operatorname{sign}(\boldsymbol{A}\boldsymbol{x}') = \boldsymbol{y} \quad \text{and} \quad \|\boldsymbol{x}'\|_2 = 1. \tag{2.1}$$

This optimization program was first proposed in [6]. Unfortunately, it is non-convex due to the constraint $\|\boldsymbol{x}'\|_2 = 1$, and therefore seems to be computationally intractable. On the other hand, we find that the non-convex program (2.1) is more amenable to theoretical analysis than the convex program (1.3).

The first ingredient of our theory will be to demonstrate that the non-convex optimization program (2.1) leads to accurate recovery of an effectively sparse signal $\boldsymbol{x}$. One can reformulate this as a geometric problem about *random hyperplane tessellations*. We will discuss tessellations in Section 4; the main result of that section is Theorem 4.2 which immediately implies the following result:

Theorem 2.1. *Let $n, m, s > 0$, and set*

$$\delta = C\left(\frac{s}{m}\log(2n/s)\right)^{1/5}. \tag{2.2}$$

Then, with probability at least $1 - C\exp(-c\delta m)$, the following holds uniformly for all $\boldsymbol{x}, \hat{\boldsymbol{x}} \in \mathbb{R}^n$ that satisfy $\|\boldsymbol{x}\|_2 = \|\hat{\boldsymbol{x}}\|_2 = 1$, $\|\boldsymbol{x}\|_1 \le \sqrt{s}$, $\|\hat{\boldsymbol{x}}\|_1 \le \sqrt{s}$:

$$\operatorname{sign}(\boldsymbol{A}\hat{\boldsymbol{x}}) = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x}) \quad \textit{implies} \quad \|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2 \le \delta.$$

Theorem 2.1 yields a version of our main Theorem 1.1 for the non-convex program (2.1):

Theorem 2.2 (Non-convex recovery)**.** *Let $n, m, s > 0$, and set δ as in* (2.1)*. Then, with probability at least $1 - C\exp(-c\delta m)$, the following holds uniformly for all signals $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \le \sqrt{s}$. Let $\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x})$. Then the solution $\hat{\boldsymbol{x}}$ of the non-convex minimization program* (2.1) *satisfies*

$$\left\| \hat{\boldsymbol{x}} - \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} \right\|_2 \le \delta.$$

Proof. We can assume without loss of generality that $\|\boldsymbol{x}\|_2 = 1$ and thus $\|\boldsymbol{x}\|_1 \le \sqrt{s}$. Since $\boldsymbol{x}$ is feasible for the program (2.1), we also have $\|\hat{\boldsymbol{x}}\|_1 \le \|\boldsymbol{x}\|_1 \le \sqrt{s}$, and thus $\hat{\boldsymbol{x}} \in S^{n-1}$. Therefore Theorem 2.1 applies to $\boldsymbol{x}, \hat{\boldsymbol{x}}$, and it yields that $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2 \le \delta$ as required. □

Remark 5 (Prior work). A version of Theorem 2.1 was recently proved in [18] for exactly sparse signals $\boldsymbol{x}, \hat{\boldsymbol{x}}$, i.e. such that $\|\boldsymbol{x}\|_2 = \|\hat{\boldsymbol{x}}\|_2 = 1$, $\|\boldsymbol{x}\|_0 \le s$, $\|\hat{\boldsymbol{x}}\|_0 \le s$. This latter result holds with $\delta = C(s/m)^{1-o(1)}\log(2n)$. However, from the proof of Theorem 2.2 given above one sees that the result of [18] would not be sufficient to deduce our main results, even Corollary 1.2 for exactly sparse vectors. The reason is that our goal is to solve a tractable program that involves the ℓ_1 norm, and thus we cannot directly assume that our estimate will be in the low-dimensional set of exactly sparse vectors. Our proof of Theorem 2.1 has to overcome some additional difficulties compared to [18] caused by the absence of any control of the supports of the signals $\boldsymbol{x}, \hat{\boldsymbol{x}}$. In particular, the

metric entropy of the set of unit-normed, sparse vectors only grows logarithmically with the inverse of the covering accuracy. This allows the consideration of a very fine cover in the proofs in [18]. In contrast, the metric entropy of the set of vectors satisfying $\|\boldsymbol{x}\|_2 \le 1$ and $\|\boldsymbol{x}\|_1 \le \sqrt{s}$ is much larger at fine scales, thus necessitating a different strategy of proof.

Theorem 1.1 would follow if we could demonstrate that the convex program (1.3) and the non-convex program (2.1) were equivalent. Rather than doing this explicitly, we shall prove that the solution $\hat{\boldsymbol{x}}$ of the convex program (1.3) essentially preserves the effective sparsity of a signal $\boldsymbol{x}$, and we finish off by applying Theorem 2.1.

Theorem 2.3 (Preserving effective sparsity)**.** *Let $n, s > 0$ and suppose that $m \ge Cs\log(n/s)$. Then, with probability at least $1 - C\exp(-cm)$, the following holds uniformly for all signals $\boldsymbol{x}$ satisfying $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \le \sqrt{s}$. Let $\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x})$. Then the solution $\hat{\boldsymbol{x}}$ of the convex minimization program* (1.3) *satisfies*

$$\frac{\|\hat{\boldsymbol{x}}\|_1}{\|\hat{\boldsymbol{x}}\|_2} \le \frac{\|\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_2} \cdot C\sqrt{\log(2n/m + 2m/n)}.$$

This result is the second main ingredient of our argument, and it will be proved in Section 5. Now we are ready to deduce Theorem 1.1.

Proof of Theorem 1.1. Consider a signal $\boldsymbol{x}$ as in Theorem 1.1, so $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \le \sqrt{s}$. In view of the application of Theorem 2.3, we may assume without loss of generality that $m \ge Cs\log(n/s)$. Indeed, otherwise we have $\delta \ge 2$ and the conclusion of Theorem 1.1 is trivial. So Theorem 2.3 applies and gives

$$\frac{\|\hat{\boldsymbol{x}}\|_1}{\|\hat{\boldsymbol{x}}\|_2} \le C\sqrt{s\log(2n/m + 2m/n)} =: \sqrt{s_0}.$$

Also, as we noted above, $\|\boldsymbol{x}\|_1 / \|\boldsymbol{x}\|_2 \le \sqrt{s} \le \sqrt{s_0}$. So Theorem 2.1 applies for the normalized vectors $\boldsymbol{x}/\|\boldsymbol{x}\|_2$, $\hat{\boldsymbol{x}}/\|\hat{\boldsymbol{x}}\|_2$ and for s_0. Note that $\operatorname{sign}(\boldsymbol{A}\hat{\boldsymbol{x}}) = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{y}$ because $\hat{\boldsymbol{x}}$ is a feasible vector for the program (1.3). Therefore Theorem 2.1 yields

$$\left\| \frac{\hat{\boldsymbol{x}}}{\|\hat{\boldsymbol{x}}\|_2} - \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} \right\|_2 \le \delta$$

where

$$\delta = C\left(\frac{s_0}{m}\log(2n/s)\right)^{1/5} = C'\left(\frac{s}{m}\log(2n/s)\log(2n/m + 2m/n)\right)^{1/5}.$$

This completes the proof. □

For the rest of the paper, our task will be to prove the two ingredients above – Theorem 2.1, which we shall relate to a more general hyperplane tessellation problem, and Theorem 2.3 on the effective sparsity of the solution.

3. Geometry of signal sets

Our arguments are based on the geometry of the set of effectively s-sparse signals

$$K_{n,s} := \big\{\boldsymbol{x} \in \mathbb{R}^n : \ \|\boldsymbol{x}\|_2 \le 1, \ \|\boldsymbol{x}\|_1 \le \sqrt{s}\big\} = B_2^n \cap \sqrt{s}B_1^n$$

and the set of s-sparse signals

$$S_{n,s} := \big\{\boldsymbol{x} \in \mathbb{R}^n : \ \|\boldsymbol{x}\|_2 \le 1, \ \|\boldsymbol{x}\|_0 \le s\big\}.$$

While the set $S_{n,s}$ is not convex, $K_{n,s}$ is, and moreover it is a convexification of $S_{n,s}$ in the following sense. Below, for a set K, we define $\operatorname{conv}(K)$ to be its convex hull.

Lemma 3.1 (Convexification)**.** *One has* $\operatorname{conv}(S_{n,s}) \subset K_{n,s} \subset 2\operatorname{conv}(S_{n,s})$.

Proof. The first containment follows by Cauchy-Schwartz inequality, which implies for each $\boldsymbol{x} \in S_{n,s}$ that $\|\boldsymbol{x}\|_1 \leq \sqrt{s}$. The second containment is proved using a common technique in the compressed sensing literature. Let $\boldsymbol{x} \in K_{n,s}$. Partition the support of $\boldsymbol{x}$ into disjoint subsets $T_1, T_2, \ldots$ so that T_1 indexes the largest s elements of $\boldsymbol{x}$ (in magnitude), T_2 indexes the next s largest elements, and so on. Since all $\boldsymbol{x}_{T_i} \in S_{n,s}$, in order to complete the proof it suffices to show that

$$\sum_{i \geq 1} \|\boldsymbol{x}_{T_i}\|_2 \leq 2.$$

To prove this, first note that $\|\boldsymbol{x}_{T_1}\|_2 \leq \|\boldsymbol{x}\|_2 \leq 1$. Second, note that for $i \geq 2$, each element of $\boldsymbol{x}_{T_i}$ is bounded in magnitude by $\left\|\boldsymbol{x}_{T_{i-1}}\right\|_1 / s$, and thus $\|\boldsymbol{x}_{T_i}\|_2 \leq \left\|\boldsymbol{x}_{T_{i-1}}\right\|_1 / \sqrt{s}$. Combining these two facts we obtain

$$\sum_{i \geq 1} \|\boldsymbol{x}_{T_i}\|_2 \leq 1 + \sum_{i \geq 2} \|\boldsymbol{x}_{T_i}\|_2 \leq 1 + \sum_{i \geq 2} \|\boldsymbol{x}_{T_i}\|_1 / \sqrt{s} \leq 1 + \|\boldsymbol{x}\|_1 / \sqrt{s} \leq 2, \tag{3.1}$$

where in the last inequality we used that $\|\boldsymbol{x}\|_1 \leq \sqrt{s}$ for $\boldsymbol{x} \in K_{n,s}$. The proof is complete. □

Our arguments will rely on entropy bounds for the set $K_{n,s}$. Consider a more general situation, where K is a bounded subset of $\mathbb{R}^n$ and $\varepsilon > 0$ is a fixed number. A subset $\mathcal{N} \subseteq K$ is called an *ε-net* of K if for every $\boldsymbol{x} \in K$ one can find $\boldsymbol{y} \in \mathcal{N}$ so that $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon$. The minimal cardinality of an ε-net of K is called the *covering number* and denoted $N(K, \varepsilon)$. The number $\log N(K, \varepsilon)$ is called the *metric entropy* of K. The covering numbers are (almost) increasing by inclusion:

$$K' \subseteq K \quad \text{implies} \quad N(K', 2\varepsilon) \leq N(K, \varepsilon). \tag{3.2}$$

Specializing to our sets of signals $K_{n,s}$ and $S_{n,s}$, we come across a useful example of an ε-net:

Lemma 3.2 (Sparse net)**.** *Let $s \leq t$. Then $S_{n,t} \cap K_{n,s}$ is an $\sqrt{s/t}$-net of $K_{n,s}$.*

Proof. Let $\boldsymbol{x} \in K_{n,s}$, and let $T \subseteq [n]$ denote the set of the indices of the t largest coefficients of $\boldsymbol{x}$ (in magnitude). Using the decomposition $\boldsymbol{x} = \boldsymbol{x}_T + \boldsymbol{x}_{T^c}$ and noting that $\boldsymbol{x}_T \in S_{n,t} \cap K_{n,s}$, we see that it suffices to check that $\|\boldsymbol{x}_{T^c}\|_2 \leq \sqrt{s/t}$. This will follow from the same steps as in (3.1). In particular, we have

$$\|\boldsymbol{x}_{T^c}\|_2 \leq \|\boldsymbol{x}\|_1 / \sqrt{|T|} \leq \sqrt{s/t}$$

as required. □

Next we pass to quantitative entropy estimates. The entropy of the Euclidean ball can be estimated using a standard volume comparison argument, as follows (see [24, Lemma 4.16]):

$$N(B_2^n, \varepsilon) \leq (3/\varepsilon)^n, \quad \varepsilon \in (0, 1). \tag{3.3}$$

From this we deduce a known bound on the entropy of $S_{n,s}$:

Lemma 3.3 (Entropy of $S_{n,s}$)**.** *For $\varepsilon \in (0, 1)$ and $s \leq n$, we have*

$$\log N(S_{n,s}, \varepsilon) \leq s \log\Big(\frac{9n}{\varepsilon s}\Big).$$

Proof. We represent $S_{n,s}$ as the union of the unit Euclidean balls $B_2^n \cap \mathbb{R}^I$ in all s-dimensional coordinate subspaces, $I \subset [n]$, $|I| = s$. Each ball $B_2^n \cap \mathbb{R}^I$ has an ε-net for of cardinality at most $(3/\varepsilon)^s$, according to (3.3). The union of these nets forms an ε-net of $S_{n,s}$, and since the number of possible I is $\binom{n}{\lfloor s \rfloor}$, the resulting net has cardinality at most $\binom{n}{\lfloor s \rfloor}(3/\varepsilon)^{\lfloor s \rfloor} \leq (3en/\varepsilon s)^s$. Taking the logarithm completes the proof. □

As a consequence, we obtain an entropy bound for $K_{n,s}$:

Lemma 3.4 (Entropy of $K_{n,s}$)**.** *For $\varepsilon \in (0,1)$, we have*

$$\log N(K_{n,s},\varepsilon) \le \begin{cases} n\log\left(\frac{6}{\varepsilon}\right) & \text{if } 0<\varepsilon<2\sqrt{\frac{s}{n}} \\ \frac{4s}{\varepsilon^2}\log\left(\frac{9\varepsilon n}{s}\right) & \text{if } 2\sqrt{\frac{s}{n}} \le \varepsilon \le 1 \end{cases}$$
$$\le \frac{Cs}{\varepsilon^2}\log\Big(\frac{2n}{s}\Big).$$

Proof. First note that $K_{n,s} \subset B_2^n$. Then the monotonicity property (3.2) followed by the volumetric estimate (3.3) yield the first desired bound $N(K_{n,s},\varepsilon) \le N(B_2^n,\varepsilon/2) \le n\log(6/\varepsilon)$ for all $\varepsilon \in (0,1)$.

Next, suppose that $2\sqrt{\frac{s}{n}} < \varepsilon < 1$. Then set $t := 4s/\varepsilon^2 \le n$. Lemma 3.2 states that $S_{n,t} \cap K_{n,s}$ is an $(\varepsilon/2)$-net of $K_{n,s}$. Furthermore, to find an $(\varepsilon/2)$-net of $S_{n,t}$, we use Lemma 3.3 for $\varepsilon/4$ and for t. Taking into account the monotonicity property (3.2), we see that there exists an $(\varepsilon/2)$-net $\mathcal{N}$ of $S_{n,t} \cap K_{n,s}$ and such that

$$\log|\mathcal{N}| \le t\log\Big(\frac{36n}{\varepsilon t}\Big) = \frac{4s}{\varepsilon^2}\log\Big(\frac{9\varepsilon n}{s}\Big).$$

It follows that $\mathcal{N}$ is an ε-net of $K_{n,s}$, and its cardinality is as required. □

4. Random hyperplane tessellations

In this section we prove a generalization of Theorem 2.1. We consider a set $K \subseteq \mathbb{R}^n$ and a collection of m random hyperplanes in $\mathbb{R}^n$, chosen independently and uniformly from the Haar measure. The resulting partition of K by this collection of hyperplanes is called a *random tessellation* of K. The cells of the tessellation are formed by intersection of K and the m random half-spaces with particular orientations. The main interest in the theory of random tessellations is the typical shape of the cells.

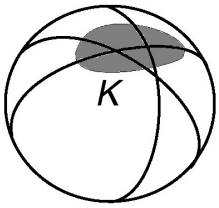

Figure 1. Hyperplane tessellation of a subset K of a sphere

We shall study the situation where K is a subset of the sphere S^{n-1}, see Figure 1. The particular example of $K = S^{n-1}$ is a natural model of random hyperplane tessellation in the sperical space S^{n-1}. The more classical and well studied model of random hyperplane tessellation is in Euclidean space $\mathbb{R}^n$, where the hyperlanes are allowed to be affine, see [23] for the history of this field. The random hyperplane tessellations of the sphere is studied in particular in [22].

Here we focus on the following question. *How many random hyperplanes ensure that all the cells of the tessellation of K have small diameter* (such as $1/2$)? For the purposes of this paper, we shall address this problem for a specific set, namely for

$$K = S^{n-1} \cap \sqrt{s}B_1^n = S^{n-1} \cap K_{n,s}.$$

We shall prove that $m = O(s\log(n/s))$ hyperplanes suffice with high probability. Our argument can be extended to more general sets K, but we defer generalizations to a later paper.

Theorem 4.1 (Random hyperplane tessellations)**.** *Let $s \le n$ and m be positive integers. Consider the tessellation of the set $K = S^{n-1} \cap \sqrt{s} B_1^n$ by m random hyperplanes in $\mathbb{R}^n$ chosen independently and uniformly from the Haar measure. Let $\delta \in (0,1)$, and assume that*

$$m \ge C\delta^{-5} s \log(2n/s).$$

Then, with probability at least $1 - 2\exp(-\delta m)$, all cells of the tessellation of K have diameter at most δ.

It is convenient to represent the random hyperplanes in Theorem 4.1 as $(\boldsymbol{a}_i)^\perp$, $i = 1, \dots, m$, where $\boldsymbol{a}_i$ are i.i.d. standard normal vectors in $\mathbb{R}^n$. The claim that all cells of the tessellation of K have diameter at most δ can be restated in the following way. Every pair of points $\boldsymbol{x}, \boldsymbol{y} \in K$ satisfying $\|\boldsymbol{x} - \boldsymbol{y}\|_2 > \delta$ is separated by at least one of the hyperplanes, so there exists $i \in [m]$ such that

$$\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle > 0, \quad \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle < 0.$$

Theorem 4.1 is then a direct consequence of the following slightly stronger result.

Theorem 4.2 (Separation by a set of hyperplanes)**.** *Let $s \le n$ and m be positive integers. Consider the set $K = S^{n-1} \cap \sqrt{s} B_1^n$ and independent random vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m \sim \mathcal{N}(0, \mathrm{Id})$ in $\mathbb{R}^n$. Let $\delta \in (0,1)$, and assume that*

$$m \ge C\delta^{-5} s \log(2n/s).$$

Then, with probability at least $1 - 2\exp(-\delta m)$, the following holds. For every pair of points $\boldsymbol{x}, \boldsymbol{y} \in K$ satisfying $\|\boldsymbol{x} - \boldsymbol{y}\|_2 > \delta$, there is a set of at least $c\delta m$ of the indices $i \in [m]$ that satisfy

$$\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle > c\delta, \quad \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle < -c\delta.$$

We will prove Theorem 4.2 by the following covering argument, which will allow us to uniformly handle all pairs $\boldsymbol{x}, \boldsymbol{y} \in K$ satisfying $\|\boldsymbol{x} - \boldsymbol{y}\|_2 > \delta$. We choose an ε-net $\mathcal{N}_\varepsilon$ of K as in Lemma 3.4. We decompose the vector $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}'$ where $\boldsymbol{x}_0 \in \mathcal{N}_\varepsilon$ is a "center" and $\boldsymbol{x}' \in \varepsilon B_2^n \cap K$ is a "tail", and we do similarly for $\boldsymbol{y}$. An elementary probabilistic argument and a union bound will allow us to nicely separate each pair of centers $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{N}_\varepsilon$ satisfying $\|\boldsymbol{x}_0 - \boldsymbol{y}_0\|_2 > \delta$ by $\Omega(m)$ hyperplanes. (Specifically, it will follow that $\langle \boldsymbol{a}_i, \boldsymbol{x}_0 \rangle > c\delta$, $\langle \boldsymbol{a}_i, \boldsymbol{y}_0 \rangle < -c\delta$ for at least $c\delta m$ of the indices $i \in [m]$.)

Furthermore, the tails $\boldsymbol{x}', \boldsymbol{y}' \in \varepsilon B_2^n \cap \sqrt{s} B_1^n$ can be uniformly controlled using Lemma 5.4, which implies that all tails are in a good position with respect to $m - o(m)$ hyperplanes. (Specifically, for small ε one can deduce that $|\langle \boldsymbol{a}_i, \boldsymbol{x}' \rangle| < c\delta/2$, $|\langle \boldsymbol{a}_i, \boldsymbol{y}' \rangle| < c\delta/2$ for at least $m - c\delta m/2$ of the indices $i \in [m]$.) Putting the centers and the tails together, we shall conclude that $\boldsymbol{x}$ and $\boldsymbol{y}$ are separated at least $\Omega(m) + m - o(m) > \Omega(m)$ hyperplanes, as required.

Now we present the full proof of Theorem 4.2.

4.1. **Step 1: Decomposition into centers and tails.** Let $\varepsilon \in (0,1)$ be a number to be determined later. Let $\mathcal{N}_\varepsilon$ be an ε-net of K. Since $K \subseteq K_{n,s}$, Lemma 3.4 along with monotonicity property of entropy (3.2) guarantee that $\mathcal{N}_\varepsilon$ can be chosen so that

$$\log |\mathcal{N}_\varepsilon| \le \frac{Cs}{\varepsilon^2} \log\Big(\frac{2n}{s}\Big). \tag{4.1}$$

Lemma 4.3 (Decomposition into centers and tails)**.** *Let $t = 4s/\varepsilon^2$. Then every vector $\boldsymbol{x} \in K$ can be represented as*

$$\boldsymbol{x} = \boldsymbol{x}_0 + \varepsilon \boldsymbol{x}' \tag{4.2}$$

where $\boldsymbol{x}_0 \in \mathcal{N}_\varepsilon$, $\boldsymbol{x}' \in K_{n,t}$.

Proof. Since $\mathcal{N}_\varepsilon$ is an ε-net of K, representation (4.2) holds for some $\boldsymbol{x}' \in B_2^n$. Since $K_{n,t} = B_2^n \cap \sqrt{t}B_1^n$, it remains to check that $\boldsymbol{x}' \in \sqrt{t}B_1^n$. Note that $\boldsymbol{x} \in K \subset \sqrt{s}B_1^n$ and $\boldsymbol{x}_0 \in \mathcal{N}_\varepsilon \subset K \subset \sqrt{s}B_1^n$. By the triangle inequality this implies that $\varepsilon \boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{x}_0 \in 2\sqrt{s}B_1^n$. Thus $\boldsymbol{x}' \in (2\sqrt{s}/\varepsilon)B_1^n = \sqrt{t}B_1^n$, as claimed. □

4.2. **Step 2: Separation of the centers.** Our next task is to separate the centers $\boldsymbol{x}_0$, $\boldsymbol{y}_0$ of each pair of points $\boldsymbol{x}, \boldsymbol{y} \in K$ that are far apart by $\Omega(m)$ hyperplanes. For a fixed pair of points and for one hyperplane, it is easy to estimate the probability of a nice separation.

Lemma 4.4 (Separation by one hyperplane)**.** *Let $\boldsymbol{x}, \boldsymbol{y} \in S^{n-1}$ and assume that $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \geq \delta$ for some $\delta > 0$. Let $\boldsymbol{a} \sim \mathcal{N}(0, \mathrm{Id})$. Then for $\delta_0 = \delta/12$ we have*

$$\mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle > \delta_0,\ \langle \boldsymbol{a}, \boldsymbol{y}\rangle < -\delta_0\right\} \geq \delta_0.$$

Proof. Note that

$$\begin{aligned}\mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle > \delta_0,\ \langle \boldsymbol{a}, \boldsymbol{y}\rangle < -\delta_0\right\} &= \mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle > 0 \text{ and } \langle \boldsymbol{a}, \boldsymbol{y}\rangle < 0 \text{ and } \langle \boldsymbol{a}, \boldsymbol{x}\rangle \notin (0, \delta_0] \text{ and } \langle \boldsymbol{a}, \boldsymbol{y}\rangle \notin [-\delta_0, 0)\right\} \\ &\geq 1 - \mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \leq 0 \text{ or } \langle \boldsymbol{a}, \boldsymbol{y}\rangle \geq 0\right\} - \mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \in (0, \delta_0]\right\} - \mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{y}\rangle \in [-\delta_0, 0)\right\}.\end{aligned}$$

The inequality above follows by the union bound. Now, since $\langle \boldsymbol{a}, \boldsymbol{x}\rangle \sim \mathcal{N}(0, 1)$ we have

$$\mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \in (0, \delta_0]\right\} \leq \frac{\delta_0}{\sqrt{2\pi}} \quad \text{and} \quad \mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{y}\rangle \in [-\delta_0, 0)\right\} \leq \frac{\delta_0}{\sqrt{2\pi}}.$$

Also, denoting the geodesic distance in S^{n-1} by $d(\cdot, \cdot)$ it is not hard to show that

$$\mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \leq 0 \text{ or } \langle \boldsymbol{a}, \boldsymbol{y}\rangle \geq 0\right\} = 1 - \frac{d(\boldsymbol{x}, \boldsymbol{y})}{2\pi} \leq 1 - \frac{\|\boldsymbol{x} - \boldsymbol{y}\|_2}{2\pi} \leq 1 - \frac{\delta}{2\pi}$$

(see [13, Lemma 3.2]). Thus

$$\mathbb{P}\left\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle > \delta_0,\ \langle \boldsymbol{a}, \boldsymbol{y}\rangle < -\delta_0\right\} \geq \frac{\delta}{2\pi} - \frac{2\delta_0}{\sqrt{2\pi}} \geq \delta_0$$

as claimed. □

Now we will pay attention to the number of hyperplanes that nicely separate a given pair of points.

Definition 4.5 (Separating set)**.** *Let $\delta_0 \in (0, 1)$. The separating index set of a pair of points $\boldsymbol{x}, \boldsymbol{y} \in S^{n-1}$ is defined as*

$$I_{\delta_0}(\boldsymbol{x}, \boldsymbol{y}) := \left\{i \in [m] :\ \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle > \delta_0,\ \langle \boldsymbol{a}_i, \boldsymbol{y}\rangle < -\delta_0\right\}.$$

The cardinality $|I_{\delta_0}(\boldsymbol{x}, \boldsymbol{y})|$ is a binomial random variable, which is the sum of m indicator functions of the independent events $\{\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle > \delta_0,\ \langle \boldsymbol{a}_i, \boldsymbol{y}\rangle < -\delta_0\}$. The probability of each such event can be estimated using Lemma 4.4. Indeed, suppose $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \geq \delta$ for some $\delta > 0$, and let $\delta_0 = \delta/12$. Then the probability of each of the events above is at least δ_0. Then $|I_{\delta_0}(\boldsymbol{x}, \boldsymbol{y})| \sim \text{Binomial}(m, p)$ with $p > \delta_0$. A standard deviation inequality (e.g. [1, Theorem A.1.13]) yields

$$\mathbb{P}\left\{|I_{\delta_0}(\boldsymbol{x}, \boldsymbol{y})| < \delta_0 m/2\right\} \leq e^{-\delta_0 m/8}. \tag{4.3}$$

Now we take a union bound over pairs of centers in the net $\mathcal{N}_\varepsilon$ that was chosen in the beginning of Section 4.1.

Lemma 4.6 (Separation of the centers)**.** *Let $\varepsilon, \delta \in (0, 1)$, and let $\mathcal{N}_\varepsilon$ be an ε-net of K whose cardinality satisfies* (4.1)*. Assume that*

$$m \geq \frac{C_1 s}{\varepsilon^2 \delta} \log\left(\frac{2n}{s}\right). \tag{4.4}$$

Then, with probability at least $1 - \exp(-\delta m/100)$*, the following event holds:*

(4.5) *For every* $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{N}_\varepsilon$ *such that* $\|\boldsymbol{x}_0 - \boldsymbol{y}_0\|_2 > \delta$*, one has* $|I_{\delta/12}(\boldsymbol{x}_0, \boldsymbol{y}_0)| \geq \delta m/24$.

Proof. For a fixed pair $\boldsymbol{x}_0, \boldsymbol{y}_0$ as above, we can rewrite (4.3) as

$$\mathbb{P}\left\{|I_{\delta/12}(\boldsymbol{x}_0, \boldsymbol{y}_0)| < \delta m/24\right\} \leq e^{-\delta m/96}.$$

A union bound over all pairs $\boldsymbol{x}_0, \boldsymbol{y}_0$ implies that the event in (4.5) fails with probability at most

$$|\mathcal{N}_\varepsilon|^2 \cdot e^{-\delta m/96}.$$

By (4.1) and (4.4), this quantity is further bounded by

$$\exp\Big[\frac{Cs}{\varepsilon^2}\log\Big(\frac{2n}{s}\Big) - \frac{\delta m}{96}\Big] \leq \exp(-\delta m/100)$$

provided the absolute constant C_1 is chosen sufficiently large. The proof is complete. □

4.3. **Step 3: Control of the tails.** Now we provide a uniform control of the tails $\boldsymbol{x}' \in K_{n,t}$ that arise from the decomposition given in Lemma 4.3. The next result is a direct consequence of Lemma 5.4.

Lemma 4.7 (Control of the tails)**.** *Let* $t \geq 1$ *and let* $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m \sim \mathcal{N}(0, \mathrm{Id})$ *be independent random vectors in* $\mathbb{R}^n$*. Assume that*

(4.6) $$m \geq Ct\log(2n/t).$$

Then, with probability at least $1 - 2\exp(-cm)$*, the following event holds:*

$$\sup_{\boldsymbol{x}' \in K_{n,t}} \frac{1}{m}\sum_{i=1}^m \big|\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle\big| \leq 1.$$

4.4. **Step 4: Putting the centers and tails together.** Let $\varepsilon = c_0\delta^2$ for a sufficiently small absolute constant $c_0 > 0$. To control the tails, we choose an ε-net $\mathcal{N}_\varepsilon$ of K as in Lemma 4.6, and we shall apply this lemma with $\delta/2$ instead of δ. Note that requirement (4.4) becomes

$$m \geq C_2\delta^{-5}s\log\Big(\frac{2n}{s}\Big),$$

and it is satisfied by the assumption of Theorem 4.2, for a sufficiently large absolute constant C. So Lemma 4.6 yields that with probability at least $1 - \exp(-\delta m/200)$, the following separation of centers holds:

(4.7) For every $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{N}_\varepsilon$ such that $\|\boldsymbol{x}_0 - \boldsymbol{y}_0\|_2 > \delta/2$, one has $|I_{\delta/24}(\boldsymbol{x}_0, \boldsymbol{y}_0)| \geq \delta m/48$.

To control the tails, we choose $t = 4s/\varepsilon^2 \sim s/\delta^4$ as in Decomposition Lemma 4.3, and we shall apply Lemma 4.7. Note that requirement (4.6) becomes

$$m \geq C_3\delta^{-4}s\log\Big(\frac{C_3 n}{s}\Big),$$

and it is satisfied by the assumption of Theorem 4.2, for a sufficiently large absolute constant C. So Lemma 4.7 yields that with probability at least $1 - 2\exp(-cm)$, the following control of tails holds:

(4.8) $$\text{For every } \boldsymbol{x}' \in K_{n,t}, \text{ one has } \frac{1}{m}\sum_{i=1}^m \big|\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle\big| \leq 1.$$

Now we combine the centers and the tails. With probability at least $1-2\exp(-c\delta m)$, both events (4.7) and (4.8) hold. Suppose both these events indeed hold, and consider a pair of vectors $\boldsymbol{x}, \boldsymbol{y} \in K$ as in the assumption, so $\|\boldsymbol{x}-\boldsymbol{y}\|_2 > \delta$. We decompose these vectors according to Lemma 4.3:

$$\boldsymbol{x} = \boldsymbol{x}_0 + \varepsilon x', \quad \boldsymbol{y} = \boldsymbol{y}_0 + \varepsilon \boldsymbol{y}' \tag{4.9}$$

where $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{N}_\varepsilon$ and $\boldsymbol{x}', \boldsymbol{y}' \in K_{n,t}$. By the triangle inequality and the choice of ε, the centers are far apart:

$$\|\boldsymbol{x}_0 - \boldsymbol{y}_0\|_2 \geq \|\boldsymbol{x}-\boldsymbol{y}\|_2 - 2\varepsilon > \delta - 2\varepsilon = \delta - 2c_0\delta^2 \geq \delta/2.$$

Then event (4.7) implies that the separating set

$$I_0 := I_{\delta/24}(\boldsymbol{x}, \boldsymbol{y}) \text{ satisfies } |I_0| \geq \delta m/48. \tag{4.10}$$

Furthermore, using (4.8) for the tails $\boldsymbol{x}'$ and $\boldsymbol{y}'$ we see that

$$\frac{1}{m}\sum_{i=1}^m \big|\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle\big| + \frac{1}{m}\sum_{i=1}^m \big|\langle \boldsymbol{a}_i, \boldsymbol{y}'\rangle\big| \leq 2.$$

By Markov's inequality, the set

$$I' := \left\{ i \in [m] : \ \big|\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle\big| + \big|\langle \boldsymbol{a}_i, \boldsymbol{y}'\rangle\big| \leq \frac{192}{\delta} \right\} \text{ satisfies } |(I')^c| \leq \frac{\delta m}{96}.$$

We claim that

$$I := I_0 \cap I'$$

is a set of indices i that satisfies the conclusion of Theorem 4.2. Indeed, the number of indices in I is as required since

$$|I| \geq |I_0| - |(I')^c| \geq \frac{\delta m}{48} - \frac{\delta m}{96} = \frac{\delta m}{96}.$$

Further, let us fix $i \in I$. Using decomposition (4.9) we can write

$$\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = \langle \boldsymbol{a}_i, \boldsymbol{x}_0\rangle + \varepsilon \langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle.$$

Since $i \in I \subseteq I_0 = I_{\delta/24}(\boldsymbol{x}, \boldsymbol{y})$, we have $\langle \boldsymbol{a}_i, \boldsymbol{x}_0\rangle > \delta/24$, while from $i \in I \subseteq I'$ we obtain $\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle \geq -192/\delta$. Thus

$$\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle > \frac{\delta}{24} - \frac{192\varepsilon}{\delta} \geq \frac{\delta}{30},$$

where the last estimate follows by the choice of $\varepsilon = c_0\delta^2$ for a sufficiently small absolute constant $c_0 > 0$. In a similar way one can show that

$$\langle \boldsymbol{a}_i, \boldsymbol{y}\rangle < -\frac{\delta}{24} + \frac{192\varepsilon}{\delta} \leq -\frac{\delta}{30}.$$

This completes the proof of Theorem 4.2. □

5. Effective sparsity of solutions

In this section we prove Theorem 2.3 about the effective sparsity of the solution of the convex optimization problem (1.3). Our proof consists of two steps – a lower bound for $\|\hat{\boldsymbol{x}}\|_2$ proved in Lemma 5.1 below, and an upper bound on $\|\boldsymbol{x}\|_2$ which we can deduce from Lemma 5.4 in the Appendix.

Lemma 5.1 (Euclidean norm of solutions)**.** *Let $n, m > 0$. Then, with probability at least $1 - C\exp(-cm\log(2n/m + 2m/n))$, the following holds uniformly for all signals $\boldsymbol{x} \in \mathbb{R}^n$. Let $\boldsymbol{y} = \operatorname{sign}(\boldsymbol{A}\boldsymbol{x})$. Then the solution $\hat{\boldsymbol{x}}$ of the convex minimization program* (1.3) *satisfies*

$$\|\hat{\boldsymbol{x}}\|_2 \geq c/\sqrt{\log(2n/m + 2m/n)}.$$

Remark 6. Note that the sparsity of the signal $\boldsymbol{x}$ plays no role in Lemma 5.1; the result holds uniformly for all signals $\boldsymbol{x}$.

Let us assume that Lemma 5.1 is true for a moment, and show how together with Lemma 5.4 it implies Theorem 2.3.

Proof of Theorem 2.3. With probability at least $1-C\exp(-cm)$, the conclusions of both Lemma 5.1 and Lemma 5.4 with $t = 1/4$ hold. Assume this event occurs. Consider a signal $\boldsymbol{x}$ as in Theorem 2.3 and the corresponding solution $\hat{\boldsymbol{x}}$ of (1.3). By Lemma 5.1, the latter satisfies

$$\|\hat{\boldsymbol{x}}\|_2 \geq c/\sqrt{\log(2n/m + 2m/n)}. \tag{5.1}$$

Next, consider

$$\lambda = \frac{1}{m}\sum_{i=1}^{m} |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| = \frac{1}{m}\|\boldsymbol{A}\boldsymbol{x}\|_1.$$

Since by the assumption on $\boldsymbol{x}$ we have $\boldsymbol{x}/\|\boldsymbol{x}\|_2 \in K_{n,s} \cap S^{n-1}$, Lemma 5.4 with $t = 1/4$ implies that

$$\lambda \geq \frac{\|\boldsymbol{x}\|_2}{2}. \tag{5.2}$$

By definition of λ, the vector $\lambda^{-1}\boldsymbol{x}$ is feasible for the program (1.3), so the solution $\hat{\boldsymbol{x}}$ of this program satisfies

$$\|\hat{\boldsymbol{x}}\|_1 \leq \|\lambda^{-1}\boldsymbol{x}\|_1 = \lambda^{-1}\|\boldsymbol{x}\|_1.$$

Putting this together with (5.2) and (5.1), we conclude that

$$\frac{\|\hat{\boldsymbol{x}}\|_1}{\|\hat{\boldsymbol{x}}\|_2} \leq \frac{\|\boldsymbol{x}\|_1}{\lambda\,\|\hat{\boldsymbol{x}}\|_2} \leq \frac{2\,\|\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_2\,\|\hat{\boldsymbol{x}}\|_2} \leq \frac{\|\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_2} \cdot C\sqrt{\log(2n/m + 2m/n)}.$$

This completes the proof of Theorem 2.3. □

In the rest of this section we prove Lemma 5.1. The argument based on the observation that the set of possible solutions $\hat{\boldsymbol{x}}$ of the convex program (1.3) for all $\boldsymbol{x}$ and corresponding $\boldsymbol{y}$ is finite, and its cardinality can be bounded by $\exp(Cm\log(2n/m + 2m/n))$. For each fixed solution $\hat{\boldsymbol{x}}$, a lower bound on $\|\hat{\boldsymbol{x}}\|_2$ will be deduced from Gaussian concentration inequalities, and the argument will be finished by taking a union bound over $\hat{\boldsymbol{x}}$.

It may be convenient to recast the convex minimization program (1.3) as a *linear program* by introducing the dummy variables $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$:

$$\min \sum_{i=1}^{n} u_i \quad \text{such that:} \quad \begin{cases} -u_i \leq x'_i \leq u_i, & i = 1, 2, \ldots, n; \\ y_i\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle \geq 0, & i = 1, 2, \ldots, m; \\ \frac{1}{m}\sum_{i=1}^{m} y_i\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle \geq 1. \end{cases} \tag{5.3}$$

The feasible set of the linear program (5.3) is a polytope in $\mathbb{R}^{2n}$, and the linear program attains a solution on a vertex of this polytope. Further, since the random Gaussian vectors $\boldsymbol{a}_i$ are in general position, one can check that the solution of the linear program is unique with probability 1. Thus, by characterizing these vertices and pointing out the relationship between u_i and $\hat{x}_i$, we may reduce the space of possible solutions $\hat{\boldsymbol{x}}$. This is the content of our next lemma. Given subsets $T \subset \{1, 2, \ldots, n\}$, $\Omega \subset \{1, 2, \ldots, m\}$, we define $\boldsymbol{A}_T^{\Omega}$ to be the submatrix of $\boldsymbol{A}$ with columns indexed by T and rows indexed by Ω.

Lemma 5.2 (Vertices of the feasible polytope)**.** *With probability* 1*, the linear program* (5.3) *attains a solution* $(\hat{\boldsymbol{x}}, \boldsymbol{u})$ *at a point which satisfies the following for some* $T \subset \{1, 2, \ldots, n\}$ *and* $\Omega \subset \{1, 2, \ldots, m\}$*:*

(1) $u_i = |\hat{x}_i|$*;*
(2) $\operatorname{supp}(\hat{\boldsymbol{x}}) = T$*;*
(3) $|T| = |\Omega| + 1$*;*
(4) $\boldsymbol{A}_T^{\Omega}\hat{\boldsymbol{x}}_T = 0$*;*
(5) $\frac{1}{m}\sum_{i=1}^{m} |\langle \boldsymbol{a}_i, \hat{\boldsymbol{x}}\rangle| = 1$.

Proof. Part (1) follows since we are minimizing $\sum u_i$. Part (5) follows since

$$\frac{1}{m}\sum_{i=1}^{m} y_i\langle \boldsymbol{a}_i, \hat{\boldsymbol{x}}\rangle = \frac{1}{m}\sum_{i=1}^{m} |\langle \boldsymbol{a}_i, \hat{\boldsymbol{x}}\rangle|$$

combined with the fact that we are implicitly minimizing $\|\boldsymbol{x}'\|_1$. Parts (2)–(4) will follow from the fact that (5.3) achieves its minimum at a vertex. The vertices are precisely the feasible points for which some d of the inequality constraints achieve equality, provided $\hat{\boldsymbol{x}}$ is the unique solution to those d equalities. Since $(\hat{\boldsymbol{x}}, \boldsymbol{u}) \in \mathbb{R}^{2n}$ at least $2n$ of the constraints must be equalities. We now count equalities based on T and Ω.

We first consider the constraints $-u_i \le x_i' \le u_i, i = 1, 2, \ldots, n$. If $\hat{x}_i = 0$ we have two equalities, $-u_i = \hat{x}_i$ and $u_i = \hat{x}_i$, otherwise, we have one. This gives $n + |T^c|$ equalities. Part (5) gives one more equality. This leaves us with at least $2n - n - |T^c| - 1 = |T| - 1$ equalities that must be satisfied out of the equations $y_i\langle \boldsymbol{a}_i, \hat{\boldsymbol{x}}\rangle \ge 0$. Thus, we may take $|\Omega| = |T| - 1$. □

Proof of Lemma 5.1. We may disregard the dummy variables (u_i) and consider that the solution $\hat{\boldsymbol{x}} = \boldsymbol{x}'$ must satisfy conditions (2)–(5) above for some T and Ω. We will show that with high probability, any such vector $\boldsymbol{x}' \in \mathbb{R}^n$ is lower bounded in the Euclidean norm.

Let us first fix sets T and Ω, and consider a vector $\boldsymbol{x}'$ satisfying (2)–(5). We represent it as

$$\boldsymbol{x}' = \mu\bar{\boldsymbol{x}} \quad \text{for some } \mu > 0 \text{ and } \|\bar{\boldsymbol{x}}\|_2 = 1.$$

Our goal is to lower bound μ. By condition (4) above, we have $\boldsymbol{A}_T^{\Omega}\bar{\boldsymbol{x}}_T = 0$ which, with probability 1, completely determines the vector $\bar{\boldsymbol{x}}$ up to multiplication by ± 1 (since $|T| = |\Omega| + 1$ and $\bar{\boldsymbol{x}}_{T^c} = 0$). Moreover, since $\operatorname{supp}(\bar{\boldsymbol{x}}) = \operatorname{supp}(\boldsymbol{x}') = T$, we have $0 = \boldsymbol{A}_T^{\Omega}\bar{\boldsymbol{x}}_T = \boldsymbol{A}^{\Omega}\bar{\boldsymbol{x}}$, so $\langle \boldsymbol{a}_i, \boldsymbol{x}'\rangle = 0$ for $i \in \Omega$. Using this with together with condition (5), we obtain

$$1 = \mu\frac{1}{m}\sum_{i=1}^{m} |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}}\rangle| = \mu\frac{1}{m}\sum_{i\notin\Omega} |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}}\rangle|$$

and thus

$$\left\|\boldsymbol{x}'\right\|_2 = \mu = \Big(\frac{1}{m}\sum_{i\notin\Omega} |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}}\rangle|\Big)^{-1}. \tag{5.4}$$

We proceed to upper bound $\frac{1}{m}\sum_{i\notin\Omega} |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}}\rangle|$.

Since the random vector $\bar{\boldsymbol{x}}$ depends entirely on $\boldsymbol{A}_T^{\Omega}$, it is independent of $\boldsymbol{a}_i$ for $i \notin \Omega$. Thus, by the rotational invariance of the Gaussian distribution, for any fixed vector $\boldsymbol{z}$ with unit norm, we have the following distributional estimates:[2]

$$\frac{1}{m}\sum_{i\notin\Omega} |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}}\rangle| \overset{\text{dist}}{=} \frac{1}{m}\sum_{i\notin\Omega} |\langle \boldsymbol{a}_i, \boldsymbol{z}\rangle| \overset{\text{dist}}{\le} \frac{1}{m}\sum_{i=1}^{m} |\langle \boldsymbol{a}_i, \boldsymbol{z}\rangle|\,.$$

[2]For random variables X, Y, the distributional inequality $X \overset{\text{dist}}{\le} Y$ means that $\mathbb{P}\{X > t\} \le \mathbb{P}\{Y > t\}$ for all $t \in \mathbb{R}$.

The last term is a sum of independent sub-Gaussian random variables, and it can be bounded using standard concentration inequalities. Specifically, applying Lemma 5.3 from the Appendix, we obtain

$$\mathbb{P}\left\{\frac{1}{m}\sum_{i=1}^{m}|\langle \boldsymbol{a}_i, \boldsymbol{z}\rangle| > t\right\} \le C\exp(-cmt^2) \quad \text{for } t \ge 2.$$

Using (5.4), this is equivalent to

$$\mathbb{P}\left\{\left\|\boldsymbol{x}'\right\|_2 < 1/t\right\} \le C\exp(-cmt^2) \quad \text{for } t \ge 2.$$

It is left to upper bound the number of vectors satisfying conditions (2)–(5) and to use the union bound. Since $|T| = |\Omega| + 1$, the total number of possible choices for T and Ω (and hence of $\boldsymbol{x}'$) is

$$\sum_{i=0}^{\min(m,n-1)} \binom{n}{i+1}\binom{m}{i} \le \exp(Cm\log(2n/m + 2m/n)).$$

Thus, by picking $t = C_0\sqrt{\log(2n/m + 2m/n)}$ with a sufficiently large absolute constant C_0, we find that all $\boldsymbol{x}'$ uniformly satisfy the required estimate $\|\boldsymbol{x}'\|_2 \ge c/\sqrt{\log(2n/m + 2m/n)}$ with probability at least $1 - \exp(Cm\log(2n/m + 2m/n)) \cdot C\exp(-cmt^2) = 1 - C\exp(-cm\log(2n/m + 2m/n))$. Lemma 5.1 is proved. □

Appendix. Uniform concentration inequality.

In this section we prove concentration inequalities for

$$\|\boldsymbol{A}\boldsymbol{x}\|_1 = \sum_{i=1}^{m}|\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle|.$$

In the situation where the vector $\boldsymbol{x}$ is fixed, we have a sum of independent random variables, which can be controlled by standard concentration inequalities:

Lemma 5.3 (Concentration)**.** *Let $n, m \in \mathbb{N}$ and $\boldsymbol{x} \in \mathbb{R}^n$. Then, for every $t > 0$ one has*

$$\mathbb{P}\left\{\left|\frac{1}{m}\sum_{i=1}^{m}|\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| - \sqrt{\frac{2}{\pi}}\|\boldsymbol{x}\|_2\right| > t\|\boldsymbol{x}\|_2\right\} \le C\exp(-cmt^2).$$

Proof. Without loss of generality we can assume that $\|\boldsymbol{x}\|_2 = 1$. Then $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle$ are independent standard normal random variables, so $\mathbb{E}\,|\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| = \sqrt{2/\pi}$. Therefore $X_i := |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| - \sqrt{2/\pi}$ are independent and identically distributed centered random variables. Moreover, X_i are sub-gaussian random variable with $\|X_i\|_{\psi_2} \le C$, see [26, Remark 18]. An application of Hoeffding-type inequality (see [26, Proposition 10]) yields

$$\mathbb{P}\left\{\left|\frac{1}{m}\sum_{i=1}^{m}X_i\right| > t\right\} \le C\exp(-cmt^2).$$

This completes the proof. □

We will now prove a stronger version of Lemma 5.3 that is uniform over all effectively sparse signals $\boldsymbol{x}$.

Lemma 5.4 (Uniform concentration)**.** *Let $n \in \mathbb{N}$, $t \in [0, \sqrt{2/\pi}]$, and suppose that $m \ge Ct^{-4}s\log(2n/s)$. Then*

$$\mathbb{P}\left\{\sup_{\boldsymbol{x}\in K_{n,s}\cap S^{n-1}}\left|\frac{1}{m}\sum_{i=1}^{m}|\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| - \sqrt{\frac{2}{\pi}}\right| > t\right\} \le C\exp(-cmt^2).$$

Proof. This is a standard covering argument, although the approximation step requires a little extra care. Let $\mathcal{M}$ be a $t/4$-net of $K_{n,s} \cap S^{n-1}$. Since $K_{n,s} \cap S^{n-1} \subseteq K_{n,s}$, we can arrange by Lemma 3.4 that

$$|\mathcal{M}| \le \exp(Ct^{-2} s \log(2n/s)).$$

By definition, for any $\boldsymbol{x} \in K_{n,s} \cap S^{n-1}$ one can find $\bar{\boldsymbol{x}} \in \mathcal{M}$ such that $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_2 \le t/4$. So the triangle inequality yields

$$\left| \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle| - \sqrt{\frac{2}{\pi}} \right| \le \left| \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}} \rangle| - \sqrt{\frac{2}{\pi}} \right| + \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle| .$$

Note that $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_1 \le \|\boldsymbol{x}\|_1 + \|\bar{\boldsymbol{x}}\|_1 \le 2\sqrt{s}$. Together with $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_2 \le t/4$ this means that

$$\boldsymbol{x} - \bar{\boldsymbol{x}} \in \frac{t}{4} \cdot K_{n,64s/t^2}.$$

Consequently,

(5.5)

$$\begin{aligned} \sup_{\boldsymbol{x} \in K_{n,s} \cap S^{n-1}} \left| \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle| - \sqrt{\frac{2}{\pi}} \right| &\le \sup_{\bar{\boldsymbol{x}} \in \mathcal{M}} \left| \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \bar{\boldsymbol{x}} \rangle| - \sqrt{\frac{2}{\pi}} \right| + \frac{t}{4} \cdot \sup_{\boldsymbol{w} \in K_{n,64s/t^2}} \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \boldsymbol{w} \rangle| \\ &=: R_1 + \frac{t}{4} \cdot R_2. \end{aligned}$$

We bound the terms R_1 and R_2 separately. For simplicity of notation, we assume that $64s/t^2$ is an integer, as the non-integer case will have no significant effect on the result.

A bound on R_1 follows from the concentration estimate in Lemma 5.3 and a union bound:

$$\mathbb{P}\{R_1 > t/4\} \le |\mathcal{M}| \cdot C \exp(-cmt^2) \le C \exp(Ct^{-2} s \log(2n/s) - cmt^2) \le C \exp(-cmt^2) \tag{5.6}$$

provided that $m \ge Ct^{-4} s \log(2n/s)$.

Next, due to Lemma 3.1 and Jensen's inequality, we have

$$R_2 \le 2 \sup_{\boldsymbol{w} \in S_{n,64s/t^2}} \frac{1}{m} \sum_{i=1}^m |\langle \boldsymbol{a}_i, \boldsymbol{w} \rangle| \le 2 \sup_{\boldsymbol{w} \in S_{n,64s/t^2}} \Big(\frac{1}{m} \sum_{i=1}^m \langle \boldsymbol{a}_i, \boldsymbol{w} \rangle^2 \Big)^{1/2} =: 2R_2'.$$

The quantity R_2' has been well studied in compressed sensing; it is bounded by the restricted isometry constant of the matrix $\frac{1}{\sqrt{m}} A$ at sparsity level $64s/t^2$. Probabilistic bounds for the restricted isometry constants of Gaussian matrices are well known, and have been derived in the earliest compressed sensing works [10]. We use the bound in [26, Theorem 65] that gives

$$\mathbb{P}\left\{R_2' > 1.5\right\} \le 2 \exp(-cm) \tag{5.7}$$

provided that $m \ge Ct^{-2} s \log(n/s)$. Thus

$$\mathbb{P}\{R_2 > 3\} \le 2 \exp(-cm).$$

Combining this and (5.6) we conclude that

$$\mathbb{P}\left\{ R_1 + \frac{t}{4} \cdot R_2 > t \right\} \le C' \exp(-cmt^2)$$

where we used the assumption that $t \le \sqrt{2/\pi}$. This and (5.5) complete the proof. □

References

[1] Alon, N., and Spencer, J. H. *The probabilistic method*, second ed. Wiley-Interscience (John Wiley and Sons), New York, 2000.

[2] Ardestanizadeh, E., Cheraghchi, M., and Shokrollahi, A. Bit precision analysis for compressed sensing. In *International Symposium on Information Theory (ISIT)* (2009), IEEE.

[3] Bickel, P. J., Ritov, Y., and Tsybakov, A. B. Simultaneous analysis of Lasso and Dantzig selector. *The Annals of Statistics 37*, 4 (Aug. 2009), 1705–1732.

[4] Boufounos, P. T. Greedy sparse signal reconstruction from sign measurements. In *Asilomar Conference on Signals, Systems and Computers* (November 2009).

[5] Boufounos, P. T. Reconstruction of sparse signals from distorted randomized measurements. In *International Conference on Acoustics, Speech and Signal Processing* (2010).

[6] Boufounos, P. T., and Baraniuk, R. G. 1-Bit compressive sensing. In *42nd Annual Conference on Information Sciences and Systems (CISS)* (Mar. 2008).

[7] Candes, E., Rudelson, M., Vershynin, R., and Tao, T. Error correction via linear programming. *46th Annual Symposium on Foundations of Computer Science* (2005), 668–681.

[8] Candes, E., and Tao, T. The Dantzig selector: Statistical estimation when p is much larger than n. *The Annals of Statistics 35*, 6 (Dec. 2007), 2313–2351.

[9] Candès, E. J., Romberg, J. K., and Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics 59*, 8 (Aug. 2006), 1207–1223.

[10] Candes, E. J., and Tao, T. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory 52*, 12 (Dec. 2006), 5406–5425.

[11] Dai, W., Pham, H. V., and Milenkovic, O. A comparative study of quantized compressive sensing schemes. In *International Symposium on Information Theory* (June 2009).

[12] Damaschke, P. Threshold group testing. *Electronic Notes in Discrete Mathematics 21* (2005), 265–271.

[13] Goemans, M. X., and Williamson, D. P. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM 42*, 6 (Nov. 1995), 1115–1145.

[14] Gunturk, C., Lammers, M., Powell, A., Saab, R., and Ylmaz, O. Sigma delta quantization for compressed sensing. In *44th Annual Conference on Information Sciences and Systems (CISS)* (2010), IEEE.

[15] Gunturk, C., Lammers, M., Powell, A., Saab, R., and Ylmaz, O. Sobolev duals for random frames and sigma-delta quantization of compressed sensing measurements. Preprint. Available at `http://arxiv.org/abs/1002.0182`.

[16] Gupta, A., Nowak, R., and Recht, B. Sample complexity for 1-bit compressed sensing and sparse classification. In *International Symposium on Information Theory (ISIT)* (2010), IEEE.

[17] Jacques, L., Hammond, D., and Fadili, J. Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine. *IEEE Transactions on Information Theory 57*, 1 (2011), 559–571.

[18] Jacques, L., Laska, J. N., Boufounos, P. T., and Baraniuk, R. G. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. Preprint. Available at `http://arxiv.org/abs/1104.3160`.

[19] Laska, J. N., Boufounos, P. T., Davenport, M. A., and Baraniuk, R. G. Democracy in action: Quantization, saturation, and compressive sensing. *Applied and Computational Harmonic Analysis* (Feb. 2011).

[20] Laska, J. N., Wen, Z., Yin, W., and Baraniuk, R. G. Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements. Preprint. Available at `http://dsp.rice.edu/sites/dsp.rice.edu/files/publications/journal-article/2010/trustbutverify.pdf`.

[21] Mackenzie, D. Compressed sensing makes every pixel count. *What's Happening in the Mathematical Sciences*, 7 (2009), 114–127.

[22] Miles, R. Random points, sets and tessellations on the surface of a sphere. *Sankhyā: The Indian Journal of Statistics, Series A 33* (1971), 145–174.

[23] Møller, J., and Stoyan, D. Stochastic geometry and random tessellations. In *Tessellations in the sciences: virtues, techniques and applications of geometric tilings*, R. van de Weijgaert, G. Vegter, V. Icke, and J. Ritzerveld, Eds. Springer-Verlag. To appear.

[24] Pisier, G. *The volume of convex bodies and Banach space geometry*. Cambridge University Press, 1989.

[25] Sun, J., and Goyal, V. Optimal quantization of random measurements in compressed sensing. In *IEEE International Symposium on Information Theory (ISIT)* (2009), IEEE.

[26] Vershynin, R. Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing: Theory and Applications*, Y. Eldar and G. Kutyniok, Eds. Cambridge University Press. To Appear (2010).

[27] Wojtaszczyk, P. Stability and instance optimality for Gaussian measurements in compressed sensing. *Foundations of Computational Mathematics 10*, 1 (Apr. 2009), 1–13.

[28] Zymnis, A., Boyd, S., and Candes, E. Compressed sensing with quantized measurements. *IEEE Signal Processing Letters 17*, 2 (Feb. 2010), 149–152.

Department of Mathematics, University of Michigan, 530 Church St., Ann Arbor, MI 48109, U.S.A.
E-mail address: {yplan,romanv}@umich.edu