Improved Bounds for Universal One-Bit Compressive Sensing

Jayadev Acharya* Cornell University
Arnab Bhattacharyya† Indian Institute of Science
Pritish Kamath‡ MIT

May 3, 2017

Abstract

Unlike compressive sensing where the measurement outputs are assumed to be real-valued and have infinite precision, in *one-bit compressive sensing*, measurements are quantized to one bit, their signs. In this work, we show how to recover the support of sparse high-dimensional vectors in the one-bit compressive sensing framework with an asymptotically near-optimal number of measurements. We also improve the bounds on the number of measurements for approximately recovering vectors from one-bit compressive sensing measurements. Our results are universal, namely the same measurement scheme works simultaneously for all sparse vectors.

Our proof of optimality for support recovery is obtained by showing an equivalence between the task of support recovery using 1-bit compressive sensing and a well-studied combinatorial object known as Union Free Families.

1 Introduction

The problem of recovering a sparse signal from a small number of measurements is a fundamental one in machine learning, statistics, and signal processing. When the measurements are linear, the process is called *compressive sensing*. Remarkable results from the last decade [Don06, CRT06] have shown that it is possible to efficiently reconstruct sparse signals using only $\Theta(k \log(n/k))$ linear measurements. Here, n is the ambient dimension of the input signal and k is its sparsity. A particularly striking result in compressive sensing is that with high probability, a Gaussian matrix with $\Theta(k \log(n/k))$ rows can be used as the sensing matrix for all sparse inputs simultaneously and is in that sense universal.

A criticism of compressive sensing is that it assumes infinite-precision real-valued measurements. Quantization of measurement outputs to very low bit-rates cannot be

*JA is supported by a start-up grant from Cornell University. Email: acharya@cornell.edu.

†AB is supported in part by a DST Ramanujan Fellowship. Part of the research was conducted while visiting the Simons Institute for Theory of Computing. Email: arnabb@csa.iisc.ernet.in.

‡PK is supported in part by NSF grants CCF-1420956, CCF-1420692, CCF-1218547, CCF-1650733. Email: pritish@mit.edu.

modeled simply as additive noise with bounded norm. To address this issue, Boufounos and Baraniuk [BB08] introduced the notion of 1-*bit compressive sensing* where each measurement is quantized to a single bit, namely its sign. This quantization can be cheaply implemented in hardware and is robust to certain non-linear distortions [Bou10]. One-bit compressive sensing is an active area of research (*e.g.*, [GNR10, YYO12, PV13, JLBB13, GNJN13, ZYJ14, KSW16, DSXZ16, LGX16, SS16]).

Formally, in 1-bit compressive sensing, given a sensing matrix A, measurements of a k-sparse[1] signal $\mathbf{x} \in \mathbb{R}^n$ are obtained by:

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

so that $\mathbf{y}$ is the vector of signs[2] of the coordinates of $A\mathbf{x}$. We consider noiseless measurements. Note that all information about the magnitude of $\mathbf{x}$ is lost by the sign operator, and we can only hope to reconstruct the normalized vector $\mathbf{x}/\left\|\mathbf{x}\right\|_2$ from $\mathbf{y}$.

In this work, we primarily consider the problem of *support recovery* of sparse vectors using 1-bit compressive sensing measurements. We focus on *universal* sensing matrices. This is commonly referred to as *for all*, or as *uniform bounds*. Universal sensing matrices have guarantees of the form, "with high probability, for all signals, the algorithm succeeds", which is in contrast to the general randomized setting where guarantees are slightly weaker, "for each signal, with high probability, the algorithm succeeds" [GSTV07]. Our objective is to minimize the total number of measurements needed (i.e., the number of rows in the sensing matrix) and the running time of the recovery algorithm. Formally:

Definition 1 (Support Recovery with 1-bit Compressed Sensing)**.** *A matrix $A \in \mathbb{R}^{m\times n}$ is a* 1-bit compressive sensing matrix *for support recovery of k-sparse vectors if there exists a recovery algorithm such that, for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\left\|\mathbf{x}\right\|_0 \leq k$, the algorithm on input $A\mathbf{x}$ returns* $\operatorname{supp}(\mathbf{x})$.

We will also consider the problem of approximate vector recovery using 1-bit compressive sensing, again focusing on the universality. Formally:

Definition 2 (Approximate Vector Recovery with 1-bit Compressed Sensing)**.** *A matrix $A \in \mathbb{R}^{m\times n}$ is a* 1-bit compressive sensing matrix *for ε-approximate vector recovery of k-sparse vectors if there exists a recovery algorithm such that, for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\left\|\mathbf{x}\right\|_0 \leq k$, the algorithm on input $A\mathbf{x}$ returns $\hat{\mathbf{x}}$ such that $\left|\frac{\mathbf{x}}{\left\|\mathbf{x}\right\|_2} - \frac{\hat{\mathbf{x}}}{\left\|\hat{\mathbf{x}}\right\|_2}\right| < \varepsilon$.*

[1] A vector is k-sparse if it has at most k nonzero components.

[2] To be precise, let $\operatorname{sign}(x) = x/|x|$ for nonzero x and $\operatorname{sign}(0) = 0$. Note that this seems to be returning more than 1 bit. But observe that if we instead define $\overline{\operatorname{sign}}(x) = \begin{cases} 1 & x > 0 \\ -1 & x \leq 0 \end{cases}$, then a measurement of $\operatorname{sign}(\langle \mathbf{a}, \mathbf{x}\rangle)$, can be simulated with two $\overline{\operatorname{sign}}$ measurements, namely using $\overline{\operatorname{sign}}(\langle \mathbf{a}, \mathbf{x}\rangle)$ and $\overline{\operatorname{sign}}(\langle -\mathbf{a}, \mathbf{x}\rangle)$.

Problem	Upper Bound	Lower Bound	Citation
Support Recovery for k sparse vectors in $\mathbb{R}^n$	$O(k^3 \log n)$	–	[GNJN13]
	–	$\Omega(k \log \frac{n}{k})$	folklore
	$O(k^2 \log n)$	$\Omega(k^2 \log n / \log k)$	This work
Approximate Recovery for k sparse vectors in $\mathbb{R}^n$	$\widetilde{O}\left(\frac{k}{\varepsilon} \log \frac{n}{k}\right)$	–	[JLBB13, GNJN13]
	$\widetilde{O}\left(k^3 \log \frac{n}{k} + \frac{k}{\varepsilon}\right)$	–	[GNJN13]
	–	$\Omega\left(k \log \frac{n}{k} + \frac{k}{\varepsilon} - k^{1.5}\right)$	[JLBB13]
	$\widetilde{O}\left(k^2 \log \frac{n}{k} + \frac{k}{\varepsilon}\right)$	$\Omega\left(k \log \frac{n}{k} + \frac{k}{\varepsilon}\right)$	This work

Table 1: Summary of results on universal 1-bit compressive sensing

1.1 Our Results

Our main contribution is to show nearly tight upper and lower bounds on the number of measurements needed for support recovery of k-sparse signals using 1-bit compressive sensing. We also provide some improvements on the bounds in approximate vector recovery. See Table 1 for a summary of our results.[3]

1.1.1 Support Recovery

Previously, Gopi et al [GNJN13] have shown a universal support recovery algorithm using $O(k^3 \log n)$ 1-bit measurements with $O(nk \log n)$ running time. If universality is not a constraint, then [HB11, GNR10] show that $O(k \log n)$ measurements suffice.

Our main contribution is showing that $\Theta(k^2 \log n)$ is a nearly tight bound for the number of 1-bit measurements needed for universal support recovery. Like in [GNJN13], our arguments exploit the structure of *Union Free Set Families* [EFF82]. While [GNJN13] uses Union Free Families to recover non-negative sparse vectors, we observe that a strengthened version of these set families can in fact be used to recover all sparse vectors. Moreover, we prove that any 1-bit compressive sensing matrix for support recovery can be converted into a Union Free Family, thus deepening the connection between the two notions. Formally, we obtain the following upper and lower bounds:

Theorem 3. (Upper bound for Support Recovery) *There exists a 1-bit compressive sensing matrix $A \in \mathbb{R}^{m \times n}$ for support recovery of k-sparse signals that uses $m = O(k^2 \log n)$ measurements. Moreover, the recovery algorithm runs in time $O(nk \log n)$.*

Theorem 4. (Lower bound for Support Recovery) *Let $A \in \mathbb{R}^{m \times n}$ be such that the map $\psi_A : \mathbb{R}^n \to \{0,1\}^m$, given by $\psi_A(\mathbf{x}) \stackrel{\text{def}}{=} \operatorname{sign}(A\mathbf{x})$ satisfies $\psi_A(\mathbf{x}_1) \neq \psi_A(\mathbf{x}_2)$ whenever $\|\mathbf{x}_1\|_0, \|\mathbf{x}_2\|_0 \leq k$ and $\operatorname{supp}(\mathbf{x}_1) \neq \operatorname{supp}(\mathbf{x}_2)$. Then, $m = \Omega(k^2 \log n / \log k)$.*

[3] the notation $\widetilde{O}(f(n))$ stands for $O(f(n) \cdot \operatorname{poly}\log(f(n)))$ for any $\operatorname{poly}\log(f(n))$.

Comparison to Group Testing. We remark that quantitatively similar results were known previously in the context of Group Testing [Dor43], which in the language of 1-bit compressive sensing, corresponds to the setting where the k-sparse signals have entries in $\{0,1\}$, and the measurements are restricted to be non-negative. Indeed, these results are obtained by showing a tight connection between Group Testing and Union-Free Families (also known as k-disjunct families). Group Testing has been an active research topic with a vast literature (See for eg, [BBTK96, DH00, ND00, AS12, CJBJ13, CJSA14, Maz16] and references therein).

Our contributions are as follows: (i) In Theorem 3, we use a strengthened notion of Union-Free Families, to obtain a better upper bound for support recovery of arbitrary k-sparse signals in $\mathbb{R}^n$; while surprisingly still using measurements vectors with entries in $\{0,1\}$. (ii) In Theorem 4, the lower bound we obtain is incomparable to the lower bound in the Group Testing problem. Our lower bound is stronger in the sense that it applies even when the measurements are arbitrary real vectors (instead of just non-negative), whereas it is weaker in the sense that the lower bound applies to measurements that can recovery the support for all k-sparse signals in $\mathbb{R}^n$ (instead of only 0-1 signals).

1.1.2 Approximate Vector Recovery

A number of papers have obtained bounds for approximate vector recovery [PV13, JLBB13, GNJN13]. The current universal 1-bit compressive sensing algorithms require

$$\min\left\{\widetilde{O}\left(\frac{k}{\varepsilon}\log\frac{n}{k}\right), \widetilde{O}\left(k^3\log\frac{n}{k}+\frac{k}{\varepsilon}\right)\right\}$$

measurements. [JLBB13] also proved a lower bound of $\Omega(k\log\frac{n}{k}+\frac{k}{\varepsilon}-k^{3/2})$ measurements.[4] As a function of ε, the second half of the bound is helpful only when $\varepsilon < 1/\sqrt{k}$.

As a corollary to Theorem 3, we can improve the upper bound term of $\widetilde{O}(k^3\log\frac{n}{k}+\frac{k}{\varepsilon})$ to $\widetilde{O}(k^2\log n+\frac{k}{\varepsilon})$. Moreover, in Section 4 we also improve the lower bound to $\Omega(k\log\frac{n}{k}+\frac{k}{\varepsilon})$, which holds for all $\varepsilon > 0$.

Corollary 5. (Improved Upper Bound for Approximate Recovery) *There exists a 1-bit compressive sensing matrix $A \in \mathbb{R}^{m\times n}$ for ε-approximate recovery of k-sparse signals that uses $m = \widetilde{O}\left(k^2\log n+\frac{k}{\varepsilon}\right)$ measurements.*

Theorem 6. (Improved Lower Bound for Approximate Recovery) *The number of measurements for ε-approximate recovery using 1-bit compressive sensing is at least $\Omega\left(k\log\frac{n}{k}+\frac{k}{\varepsilon}\right)$.*

2 Upper Bound for Support Recovery

In this section we prove Theorem 3. Gopi et al [GNJN13] present two techniques to obtain 1-bit compressive sensing matrices for support recovery of k-sparse signals. The first

[4] Strictly speaking, the lower bound of $\Omega(k\log\frac{n}{k})$ is folklore (we provide a proof in Section 4 for completeness), and [JLBB13] showed a lower bound of $\Omega(\frac{k}{\varepsilon}-k^{1.5})$.

technique is based on Union-Free-Families (UFF) to solve support recovery using only $O(k^2 \log n)$ measurements. However, this technique works only when the signals are non-negative. In order to handle all real-valued signals, they propose a technique based on expanders that uses $O(k^3 \log n)$ measurements. This expander based technique can be interpreted as implicitly constructing a generalization of UFFs called *Robust-UFF* (Definition 8). This construction is able to handle all real signals, albeit with an additional multiplicative factor of k in the number of measurements. Our upper bound uses Robust-UFFs constructed directly using the probabilistic method instead of going via expanders, thereby leading to a 1-bit compressive sensing matrix for support recovery using only $O(k^2 \log n)$ measurements.

Definition 7 (Union Free Family)**.** *A family of sets $\mathcal{F} = \{B_1, B_2, \ldots, B_n\}$, where each $B_i \subseteq [m]$[5] is an (n, m, k)*-UFF *if the following holds: for all distinct $j_0, j_1, \ldots, j_k \in [n]$, it is the case that $B_{j_0} \not\subseteq (B_{j_1} \cup B_{j_2} \cup \cdots \cup B_{j_k})$.*

Definition 8 (Robust Union Free Family)**.** *A family of sets $\mathcal{F} = \{B_1, B_2, \ldots, B_n\}$, where each $B_i \subseteq [m]$ is an (n, m, d, k, α)*-Robust-UFF *if the following holds: for all distinct $j_0, j_1, \ldots, j_k \in [n]$, it is the case that $|B_{j_0} \cap (B_{j_1} \cup B_{j_2} \cup \cdots \cup B_{j_k})| < \alpha |B_{j_0}|$ and $|B_j| = d$ for every $j \in [n]$.*

An easy application of the probabilistic method shows the existence of Robust-UFFs with certain desirable parameters, as done in [dW12].

Lemma 9 (Existence of Robust-UFF [dW12])**.** *There exists an (n, m, d, k, α)-Robust-UFF $\mathcal{F}$ with parameters satisfying $m = O\left(\frac{k^2 \log n}{\alpha^2}\right)$ and $d = O\left(\frac{k \log n}{\alpha}\right)$.*

Remark 10. *Union Free Families (UFF) are a special case of Robust-UFF when $\alpha = 1$, namely $|B_{j_0} \cap (B_{j_1} \cup B_{j_2} \cup \cdots \cup B_{j_k})| < |B_{j_0}|$.*

Support recovery from Robust-UFFs

We are now ready to prove Theorem 3 (restated below for convenience) by constructing a suitable 1-bit compressive sensing matrix.

Theorem 3. (Upper bound for Support Recovery) *There exists a 1-bit compressive sensing matrix $A \in \mathbb{R}^{m \times n}$ for support recovery of k-sparse signals that uses $m = O(k^2 \log n)$ measurements. Moreover, the recovery algorithm runs in time $O(nk \log n)$.*

Proof. Starting from any $(n, m, d, k, 1/2)$-Robust-UFF $\mathcal{F} = \{B_1, \ldots, B_n\}$, we construct a compressive sensing matrix $A \in \{0, 1\}^{m \times n}$ as follows: $A_{i,j} = \mathbf{1}_{(i \in B_j)}$. From Lemma 9, we have that such a Robust-UFF exists with $m = O(k^2 \log n)$ and $d = O(k \log n)$. On receiving input $\mathbf{b} = A\mathbf{x}^*$, the support recovery algorithm proceeds as follows: Include j into set $\widehat{S}$ if and only if at least half of the measurements corresponding to set B_j are non-zero. See Algorithm 1 for a more detailed pseudo-code.

Correctness. Suppose the k-sparse vector is supported on coordinates $x_1, \ldots, x_k$ (the proof works similarly for other supports). Firstly for any $j \notin [k]$, we have that $|B_j \cap (B_1 \cup$

[5] $[m] := \{1, \ldots, m\}$

Algorithm 1: 1-bit Compressed Sensing for Support Recovery from Robust UFFs

Input: $\mathbf{b} = \text{sign}(A\mathbf{x}^*)$

$\widehat{S} \leftarrow \emptyset$

for $j \in [n]$ **do**

  if $\left|B_j \cap \text{supp}(\mathbf{b})\right| > d/2$ **then**

    $\widehat{S} \leftarrow \widehat{S} \cup \{j\}$

  end

end

Output: $\widehat{S}$

$\cdots \cup B_k)| < |B_j|/2 = d/2$ from the definition of Robust-UFF. Thus, irrespective of the values of $x_1, \ldots, x_k$, the measurement outcomes corresponding to $B_j \setminus (B_1 \cup \cdots \cup B_k)$ will always be zero. Since more than $d/2$ of the measurements in B_j are zero, j will not be included in the set $\widehat{S}$. Next, consider any $j \in [k]$. Again, from the definition of Robust-UFF, we have that $\left|B_j \cap \left(\bigcup_{i\in[k], i\neq j} B_i\right)\right| < |B_j|/2 = d/2$. Thus, irrespective of the values of $x_1, \ldots, x_k$, the measurement outcomes corresponding to $B_j \setminus \left(\bigcup_{i\in[k], i\neq j} B_i\right)$ will be non-zero. Since more than $d/2$ of the measurements in B_j are non-zero, j will be included in the set $\widehat{S}$.

Efficiency. It easy to see that each iteration of the algorithm takes $O(k \log n)$ time, and hence overall the algorithm runs in $O(nk \log n)$ time. Note that, here we are not accounting for the time needed to construct the matrix A which is part of pre-processing. □

3 Lower Bound for Support Recovery

In this section we prove Theorem 4. We prove this lower bound in two steps,

1. we show that 1-bit compressive sensing implies the existence of a Union Free Family with similar parameters, and
2. we use known upper bounds on the size of Union Free Families to prove our lower bound.

We start with the second point, for which we simply use the upper bound on the size of UFFs due to Füredi [Für96].

Lemma 11 (Upper bound on Union-Free Families [Für96]). *Let $\mathcal{F} = \{B_1, \ldots, B_n\}$ be a family of subsets of $[m]$, and $k \geq 2$, such that for any $j_0, j_1, \ldots, j_k$, it holds that $B_{j_0} \not\subseteq B_{j_1} \cup B_{j_2} \cup \cdots \cup B_{j_k}$. Then,*

$$n \leq k + \binom{m}{t} \quad \text{where, } t = \left\lceil \frac{m-k}{\binom{k+1}{2}} \right\rceil.$$

This implies $m \geq \Omega(k^2 \log n / \log k)$.

For ease of presentation, we first prove a lower bound on the number of measurements for exact support recovery using only non-negative measurements, namely in the following theorem the entries of A are non-negative (note that the compressive sensing matrix obtained in the proof of Theorem 3 in fact had only 0-1 entries). More strongly, our lower bound works even when the matrix has to recover the support for only 0-1 vectors. We remark that this lower bound was already known in the context of Combinatorial Group Testing [DR82, DRR89, Für96]. We still present this proof first as it serves as a natural segue into our main lower bound.

Theorem 12 (Lower bound for non-negative measurements)**.** *Let $A \in \mathbb{R}_{\geq 0}^{m \times n}$ be such that the map $\psi_A : \{0,1\}^n \to \{0,1\}^m$, given by $\psi_A(\mathbf{x}) \stackrel{\text{def}}{=} \text{sign}(A\mathbf{x})$ satisfies $\psi_A(\mathbf{x}_1) \neq \psi_A(\mathbf{x}_2)$ whenever $\|\mathbf{x}_1\|_0, \|\mathbf{x}_2\|_0 \leq k$ and $\text{supp}(\mathbf{x}_1) \neq \text{supp}(\mathbf{x}_2)$. Then, $m = \Omega(k^2 \log n / \log k)$.*

Proof. Any algorithm for exact support recovery of non-negative k-sparse signals, which uses only positive measurements, can be converted into an $(n, m, k-1)$-UFF. Suppose A is a matrix achieving support recovery for non-negative k-sparse signals with m measurements. Let $B_1, B_2, \ldots, B_n \subseteq [m]$ be such that $B_j = \{i : A_{ij} > 0\}$. Suppose for contradiction that $\mathcal{B} = \{B_1, \ldots, B_n\}$ is not an $(n, m, k-1)$-UFF. Then there exists $j_0, j_1, \ldots, j_{k-1}$ such that $B_{j_0} \subseteq B_{j_1} \cup \ldots \cup B_{j_{k-1}}$. Let $\mathbf{x}_1 = \mathbf{1}(\{j_1, \ldots, j_{k-1}\})$ and $\mathbf{x}_2 = \mathbf{1}(\{j_0, j_1, \ldots, j_{k-1}\})$. It is easy to see that $\psi_A(\mathbf{x}_1) = \psi_A(\mathbf{x}_2)$, which is a contradiction. Thus, we conclude that $\mathcal{B}$ is a $(n, m, k-1)$-UFF and hence from Lemma 11, we get that $m \geq \Omega(k^2 \log n / \log k)$. □

Remarkably, we use the same technique to prove our main lower bound, i.e. Theorem 4 (restated below for convenience), on the number of measurements needed for exact support recovery using arbitrary linear threshold measurements. However, here we need to use that the algorithm returns the exact support for all $(\leq k)$-sparse vectors in $\mathbb{R}^n$ and not just those in $\{0,1\}^n$.

Theorem 4. (Lower bound for Support Recovery) *Let $A \in \mathbb{R}^{m \times n}$ be such that the map $\psi_A : \mathbb{R}^n \to \{0,1\}^m$, given by $\psi_A(\mathbf{x}) \stackrel{\text{def}}{=} \text{sign}(A\mathbf{x})$ satisfies $\psi_A(\mathbf{x}_1) \neq \psi_A(\mathbf{x}_2)$ whenever $\|\mathbf{x}_1\|_0, \|\mathbf{x}_2\|_0 \leq k$ and $\text{supp}(\mathbf{x}_1) \neq \text{supp}(\mathbf{x}_2)$. Then, $m = \Omega(k^2 \log n / \log k)$.*

Proof. Let A be a matrix for support recovery of non-negative k-sparse signals. Without loss of generality, assume that $-1 \leq A_{ij} \leq 1$ for all i, j (since scaling A by constants doesn't change the outcome of sign measurements). Similar to the proof of Theorem 12, let $B_1, B_2, \ldots, B_n \subseteq [m]$ be such that $B_j = \{i \in [m] : A_{ij} \neq 0\}$. Suppose for contradiction that $\mathcal{B} = \{B_1, \ldots, B_n\}$ is not a $(n, m, k-1)$-UFF. Hence, there exists $j_0, j_1, \ldots, j_{k-1}$ such that $B_{j_0} \subseteq B_{j_1} \cup \cdots \cup B_{j_{k-1}}$. We now construct two $(\leq k)$-sparse vectors $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ with different supports such that $\psi_A(\mathbf{x}_1) = \psi_A(\mathbf{x}_2)$.

Let $\mathbf{x}_1$ be a vector supported on $j_1, \ldots, j_{k-1}$ such that all indices of $A\mathbf{x}_1$ in $B_{j_1} \cup \cdots \cup B_{j_{k-1}}$ are ε-away from 0 for some choice of ε[6]. Let $\mathbf{x}_2 = \mathbf{x}_1 + \varepsilon \cdot e_{j_0}$. Since we assumed that $-1 \leq A_{ij} \leq 1$ for all i, j, we have that $A\mathbf{x}_2 - A\mathbf{x}_1 = A \cdot (\varepsilon e_{j_0})$ has all entries with

[6]We can first choose a random $\mathbf{x}_1'$ supported on $j_1, \ldots, j_{k-1}$, which will be such that all indices of $A\mathbf{x}_1'$ in $B_{j_1} \cup \cdots \cup B_{j_{k-1}}$ are non-zero. Now we can get $\mathbf{x}_1$ by simply scaling up this random $\mathbf{x}_1'$ by a suitably large constant.

magnitude at most ε. Since all entries of $A\mathbf{x}_1$ in $B_{j_1} \cup \cdots \cup B_{j_{k-1}}$ are ε-away from 0, and $B_{j_0} \subseteq B_{j_1} \cup \cdots \cup B_{j_{k-1}}$, we get that $\psi_A(\mathbf{x}_1) = \psi_A(\mathbf{x}_2)$ even though $\text{supp}(\mathbf{x}_1) \neq \text{supp}(\mathbf{x}_2)$. Note that both $\mathbf{x}_1$ and $\mathbf{x}_2$ are $(\leq k)$-sparse, and hence we get a contradiction.

Thus, we conclude that $\mathcal{B}$ is a $(n, m, k-1)$-UFF and hence from Lemma 11, we get that $m \geq \Omega(k^2 \log n / \log k)$. □

Thus, with Theorem 4, we get a nearly tight lower bound of $\Omega(k^2 \log n / \log k)$ on the number of measurements needed for support recovery, even if we assume that the signals are non-negative and the measurements are allowed to be arbitrary. This is nearly matching the upper bound obtained in Theorem 3, where we have a measurement matrix with $O(k^2 \log n)$ rows and only 0-1 entries, which can recover support exactly for all signals in $\mathbb{R}^n$.

We note that our lower bound proof requires that the compressive sensing matrix correctly recovers the support for signals with arbitrarily large *condition number*. The condition number (or dynamic range) of a signal $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is defined as

$$K_{\mathbf{x}} = \frac{\max_{i:\mathbf{x}_i \neq 0} |\mathbf{x}_i|}{\min_{i:\mathbf{x}_i \neq 0} |\mathbf{x}_i|},$$

which is the highest ratio of absolute values of non-zero components of the signal.

Signals with bounded condition numbers are easier to handle and are also robust to noise. For example, [GNR10] considered the case of signals with bounded condition number (in addition to presence of noise), although their measurements work in the non-universal setting. Obtaining bounds on the number of measurements required in the universal setting, as a function of the condition number, is open, even in the absence of noise. Even the case when the condition number is 1 is open (see the discussion in Section 5).

4 Approximate vector recovery

4.1 Upper Bound

For the problem of approximate vector recovery, note as in Table 1, that the two known upper bounds are $\widetilde{O}(\frac{k}{\varepsilon} \log \frac{n}{k})$, and $\widetilde{O}(k^3 \log \frac{n}{k} + \frac{k}{\varepsilon})$. We improve the second bound of $\widetilde{O}(k^3 \log \frac{n}{k} + \frac{k}{\varepsilon})$ by [GNJN13] to $\widetilde{O}(k^2 \log n + \frac{k}{\varepsilon})$ in Corollary 5 (restated for convenience).

Corollary 5. (Improved Upper Bound for Approximate Recovery) *There exists a 1-bit compressive sensing matrix $A \in \mathbb{R}^{m \times n}$ for ε-approximate recovery of k-sparse signals that uses $m = \widetilde{O}\left(k^2 \log n + \frac{k}{\varepsilon}\right)$ measurements.*

Proof. The upper bound of $\widetilde{O}(k^3 \log \frac{n}{k} + \frac{k}{\varepsilon})$ in [GNJN13] is shown by recovering the support of the vector using $O(k^3 \log \frac{n}{k})$ measurements and subsequently using $\widetilde{O}(k/\varepsilon)$ measurements to approximately recover the vector in k dimensions (this is still non-adaptive because standard Gaussian measurements suffice to approximately recover the vector).

Instead, using our improved algorithm of Theorem 3, we need only $O(k^2 \log n)$ measurements for support recovery, thereby obtaining the overall bound. □

4.2 Lower Bound

A lower bound of $\Omega(k \log \frac{n}{k} + \frac{k}{\varepsilon} - k^{3/2})$ measurements for $\varepsilon < \frac{1}{\sqrt{k}}$ was shown in [JLBB13]. We prove the same bound for all values of ε up to a constant in Theorem 6 (restated below for convenience). We essentially follow the approach of [JLBB13], but unlike their lower bound, focus on only one set of k coordinates, instead of all possible sparsity patterns. Surprisingly, this gives us a simpler proof that improves the lower bound by getting rid of the $k^{3/2}$ term.

Theorem 6. (Improved Lower Bound for Approximate Recovery) *The number of measurements for ε-approximate recovery using 1-bit compressive sensing is at least $\Omega\left(k \log \frac{n}{k} + \frac{k}{\varepsilon}\right)$.*

Proof. The first term of $k \log \frac{n}{k}$ is folklore. Nevertheless, we present the proof here for completeness. Consider the set of all k-sparse vectors of unit norm that have each non-zero entry equal to $1/\sqrt{k}$. Using the Gilbert-Varshamov bound [Gil52, Var57], within this set there is a subset of at least $M = \binom{n}{k} / \binom{n}{\varepsilon k}$ elements such that for u, and v in the set, their supports have intersection at most $(1-\varepsilon)k$. This implies that $\|u - v\|_2 \geq \Omega(\varepsilon)$. Since m sign measurements can give us only m bits of information, this gives us that $2^m \geq M$. By Stirling's approximation, for any $\varepsilon < 0.5$, we have

$$m \geq \log M \geq \Omega\left(k \log \frac{n}{k}\right).$$

This shows the first term. For the second term, we use the following lemma.

Lemma 13 (cf. Lemma 1 in [JLBB13])**.** *Let $m \geq 2k$. Then m-hyperplanes in k-dimensions divides the region into at most $2^k \binom{m}{k}$ regions.*

We now use the following well known lower bound on an ε-cover for S^{k-1} (this follows from a straight forward volume argument).

Lemma 14 (ε-cover for S^{k-1})**.** *There exists a subset $\mathcal{C} \subseteq S^{k-1}$, and a constant $c > 0$ such that, $|\mathcal{C}| \geq \left(\frac{c}{\varepsilon}\right)^k$, and for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}$, it holds that $\left\|\mathbf{x} - \mathbf{y}\right\|_2 \geq \varepsilon$.*

Now consider any 1-bit compressive sensing matrix with m rows. They will correspond to m hyperplanes. To reconstruct all the vectors in $\mathcal{C}$, each entry of the $\mathcal{C}$ must lie in a different region that the m hyperplanes slice S^{k-1} into. This in turn requires $2^k \binom{m}{k} \geq \left(\frac{c}{\varepsilon}\right)^k$. Since $\binom{m}{k} < (me/k)^k$, we get $2em/k > c/\varepsilon$, thereby proving the bound.

We remark that our analysis is very similar to that of [JLBB13]. The main difference is that instead of considering sparse signals as lying in n dimensions simultaneously, which is a union of subspaces, we just consider the set of all signals lying in k dimensions. □

Thus, combining our results with prior literature, the upper and lower bounds for ε-vector recovery stand as follows:

- Upper bound: $\min\left\{\widetilde{O}\left(\frac{k}{\varepsilon} \log \frac{n}{k}\right), \widetilde{O}\left(k^2 \log \frac{n}{k} + \frac{k}{\varepsilon}\right)\right\}$.
- Lower bound: $\Omega\left(k \log \frac{n}{k} + \frac{k}{\varepsilon}\right)$.

5 Open Problems

We point out three intriguing open problems that so far have resisted easy answers. The first problem is the support recovery problem for vectors with condition number 1.

Open Problem 1. *How many measurements are necessary and sufficient for a universal algorithm that recovers the support of all 0-1 vectors that are k-sparse using 1-bit compressive sensing?*

It can be shown that $O(k^{3/2}\log n)$ random gaussian measurements suffice to recover all 0-1 vectors. This requires a simple computation that also follows from Theorem 2 of [JLBB13]. On the other hand, the best known lower bound is the trivial $\Omega(k\log(n/k))$ measurements.

Our second open problem is about the approximate vector recovery problem.

Open Problem 2. *What is the correct complexity of ε-approximate vector recovery using 1-bit compressive sensing?*

We know from Section 4 and [JLBB13, GNJN13], that $\min\left\{\tilde{O}(\frac{k}{\varepsilon}\log\frac{n}{k}), \tilde{O}(k^2\log\frac{n}{k}+\frac{k}{\varepsilon})\right\}$ is an upper bound and $\Omega\left(k\log\frac{n}{k}+\frac{k}{\varepsilon}\right)$ is a lower bound. The bounds are within a constant factor of each other in the regime where $\varepsilon < 1/(k\log\frac{n}{k})$ and also in the regime where $\varepsilon = \Theta(1)$. However, there is still a gap in the regime where $1 \gg \varepsilon \gg 1/(k\log\frac{n}{k})$.

Our final open problem is to obtain "explicit" constructions for Robust-UFFs with the parameters that we want.

Open Problem 3. *Obtain an efficient algorithm to construct (n, m, d, k, α)-Robust-UFF with parameter $m = O\left(\frac{k^2\log n}{\alpha^2}\right)$, that is, in time that is polynomial in n and k.*

See Appendix A for some approaches to obtain explicit constructions of Robust-UFFs using explicit error correcting codes. Unfortunately, such approaches using known constructions of error correcting codes, seem to fall shy of achieving the parameters we want.

Acknowledgment

AB would like to thank Prateek Jain for preliminary discussions about these problems. AB would also like to thank Prateek Jain, Satya Lokam and N.V. Vinodchandran for discussions about Open Problem 1. The authors thank Russell Impagliazzo for pointing out the connection between Union Free Families and Nisan-Wigderson designs. The authors thank Jiawei Gao, Russell Impagliazzo and Madhu Sudan for discussions on explicit constructions of Robust Union Free Families (see Open Problem 3). Finally we thank anonymous reviewers for several useful comments that improved the presentation of the paper, and also for pointing us to relevant work in the group testing literature.

References

[AS12] G.K. Atia and V. Saligrama. Boolean compressed sensing and noisy group testing. *Information Theory, IEEE Transactions on*, 58(3):1880–1901, March 2012. 4

[BB08] Petros T Boufounos and Richard G Baraniuk. 1-bit compressive sensing. In *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*, pages 16–21. IEEE, 2008. 2

[BBTK96] D.J. Balding, W.J. Bruno, D.C Torney, and E. Knill. A comparative survey of non-adaptive pooling designs. In Terry Speed and MichaelS. Waterman, editors, *Genetic Mapping and DNA Sequencing*, volume 81 of *The IMA Volumes in Mathematics and its Applications*, pages 133–154. Springer New York, 1996. 4

[Bou10] Petros T Boufounos. Reconstruction of sparse signals from distorted randomized measurements. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 3998–4001. IEEE, 2010. 2

[CJBJ13] Sheng Cai, Mohammad Jahangoshahi, Mayank Bakshi, and Sidharth Jaggi. GROTESQUE: noisy group testing (quick and efficient). *CoRR*, abs/1307.2811, 2013. 4

[CJSA14] Chun Lam Chan, Sidharth Jaggi, Venkatesh Saligrama, and Samar Agnihotri. Non-adaptive group testing: Explicit bounds and novel algorithms. *IEEE Transactions on Information Theory*, 60(5):3019–3035, 2014. 4

[CRT06] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006. 1

[DH00] Dingzhu Du and Frank K. Hwang. *Combinatorial Group Testing and Its Applications*. Applied Mathematics. World Scientific, 2000. 4

[Don06] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006. 1

[Dor43] Robert Dorfman. The detection of defective members of large populations. *The Annals of Mathematical Statistics*, 14(4):436–440, 12 1943. 4

[DR82] Arkadii Georgievich D'yachkov and Vladimir Vasil'evich Rykov. Bounds on the length of disjunctive codes. *Problemy Peredachi Informatsii*, 18(3):7–13, 1982. 7

[DRR89] AG Dyachkov, VV Rykov, and AM Rashad. Superimposed distance codes. *Problems of Control and Information Theory-problemy Upravleniya i Teorii Informatsii*, 18(4):237–250, 1989. 7

[DSXZ16] Dao-Qing Dai, Lixin Shen, Yuesheng Xu, and Na Zhang. Noisy 1-bit compressive sensing: models and algorithms. *Applied and Computational Harmonic Analysis*, 40(1):1–32, 2016. 2

[dW12] Ronald de Wolf. Efficient data structures from union-free families of sets. 2012. 5

[EFF82] Paul Erdös, Peter Frankl, and Zoltán Füredi. Families of finite sets in which no set is covered by the union of two others. *Journal of Combinatorial Theory, Series A*, 33(2):158–166, 1982. 3

[Für96] Zoltán Füredi. On r-cover-free families. *J. Comb. Theory, Ser. A*, 73(1):172–173, 1996. 6, 7

[Gil52] E. N. Gilbert. A comparison of signalling alphabets. *Bell System Technical Journal*, 31:504–522, 1952. 9

[GNJN13] Sivakant Gopi, Praneeth Netrapalli, Prateek Jain, and Aditya Nori. One-bit compressed sensing: Provable support and vector recovery. In *Proceedings of The 30th International Conference on Machine Learning*, pages 154–162, 2013. 2, 3, 4, 8, 10

[GNR10] Ankit Gupta, Robert Nowak, and Benjamin Recht. Sample complexity for 1-bit compressed sensing and sparse classification. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 1553–1557. IEEE, 2010. 2, 3, 8

[GSTV07] Anna C Gilbert, Martin J Strauss, Joel A Tropp, and Roman Vershynin. One sketch for all: fast algorithms for compressed sensing. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 237–246. ACM, 2007. 2

[HB11] Jarvis Haupt and Richard Baraniuk. Robust support recovery using sparse compressive sensing matrices. In *Information Sciences and Systems (CISS), 2011 45th Annual Conference on*, pages 1–6. IEEE, 2011. 3

[JLBB13] Laurent Jacques, Jason N. Laska, Petros T. Boufounos, and Richard G. Baraniuk. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE Transactions on Information Theory*, 59(4):2082–2102, 2013. 2, 3, 4, 9, 10

[KSW16] Karin Knudson, Rayan Saab, and Rachel Ward. One-bit compressive sensing with norm estimation. *IEEE Transactions on Information Theory*, 62(5):2748–2758, 2016. 2

[LGX16] Wenhui Liu, Da Gong, and Zhiqiang Xu. One-bit compressed sensing by greedy algorithms. *Numerical Mathematics: Theory, Methods and Applications*, 9(02):169–184, 2016. 2

[Maz16] Arya Mazumdar. Nonadaptive group testing with random set of defectives. *IEEE Transactions on Information Theory*, 62(12):7522–7531, Dec 2016. 4

[ND00] Hung Q. Ngo and Ding-Zhu Du. A Survey on Combinatorial Group Testing Algorithms with Applications to DNA Library Screening. In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 2000. 4

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994. 14

[PV13] Yaniv Plan and Roman Vershynin. Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach. *IEEE Trans. Information Theory*, 59(1):482–494, 2013. 2, 4

[SS16] Lixin Shen and Bruce W Suter. One-bit compressive sampling via ℓ_0 minimization. *EURASIP Journal on Advances in Signal Processing*, 2016(1):71, 2016. 2

[Tre01] Luca Trevisan. Extractors and pseudorandom generators. *J. ACM*, 48(4):860–879, 2001. 14

[Var57] R. R. Varshamov. Estimate of the number of signals in error correcting codes. *Dokl. Acad. Nauk SSSR*, 117:739–741, 1957. 9

[YYO12] Ming Yan, Yi Yang, and Stanley Osher. Robust 1-bit compressive sensing using adaptive outlier pursuit. *IEEE Transactions on Signal Processing*, 60(7):3868–3875, 2012. 2

[ZYJ14] Lijun Zhang, Jinfeng Yi, and Rong Jin. Efficient algorithms for robust one-bit compressive sensing. In *ICML*, pages 820–828, 2014. 2

A Towards Explicit Constructions of Robust-UFFs

In this section, we describe some attempts to get explicit constructions of Robust-UFFs using explicit constructions of error correcting codes.

Proposition 15. *If $\mathcal{F}$ is a $(n, m, d, 1, \alpha/k)$-Robust-UFF, then $\mathcal{F}$ is also a (n, m, d, k, α)-Robust-UFF.*

Proof. Let $\mathcal{F} = \{B_1, \ldots, B_n\}$, where $B_i \subseteq [m]$ with $|B_i| = d$ for all $i \in [n]$. From the definition of $(n, m, d, 1, \alpha/k)$-Robust-UFF, we have that for any $i \neq j$, it holds that $|B_i \cap B_j| < \alpha d/k$. Thus, we also get that for any $i_0, i_1, \cdots, i_k$, it holds that,

$$|B_{i_0} \cap (B_{i_1} \cup \cdots \cup B_{i_k})| \leq \sum_{j=1}^{k} |B_{i_0} \cap B_{i_j}| < \alpha d \ .$$

Thus, $\mathcal{F}$ is also a (n, m, d, k, α)-Robust-UFF. □

We note that Lemma 9 implies the existence of $(n, m, d, 1, \alpha/k)$-Robust-UFF with roughly the same parameters as that for (n, m, d, k, α)-Robust-UFF. So in terms of probabilistic constructions at least, we aren't making our task too difficult by focusing on the $k = 1$ case.

Thus, it suffices to construct Robust-UFFs with parameter $k = 1$, which are well studied under the name of Nisan-Wigderson designs [NW94]. In particular, it is known that $(n, m, d, 1, \alpha/k)$-Robust-UFF can be constructed in time $n^{O(k^2)}$ with the parameters of our interest, namely $m = O\left((k^2 \log n)/\alpha^2\right)$ and $d = O\left((k \log n)/\alpha\right)$ (cf. [Tre01]). We now describe how explicit constructions of error correcting codes, immediately give rise to explicit constructions of Nisan-Wigderson designs.

Definition 16. *$\mathcal{C} \subseteq [q]^d$ is an error correcting code with (relative) distance $(1 - \delta)$ and rate r if*

- *for any $c_1 \neq c_2 \in \mathcal{C}$, the distance $\Delta(c_1, c_2) \geq (1 - \delta)d$,*
  where $\Delta(c_1, c_2) \stackrel{\text{def}}{=} \#\{i \in [d] : c_1(i) \neq c_2(i)\}$.
- *$r = \frac{\log |\mathcal{C}|}{d \log q}$, or equivalently, $|\mathcal{C}| = q^{rd}$.*

Proposition 17 (Robust-UFFs from error correcting codes)**.** *Given an error correcting code $\mathcal{C} \subseteq [q]^d$ with distance $(1 - \delta)$ and rate r, it is possible to construct a $(n, m, d, 1, \delta)$-Robust-UFF, with parameters, $n = |\mathcal{C}| = q^{rd}$ and $m = q \cdot d = \frac{q \log n}{r \log q}$.*

Proof. We construct a $(n, m, d, 1, \delta)$-Robust-UFF $\mathcal{F}$ from such an error correcting code $\mathcal{C}$ as follows. Consider a universe $[d] \times [q]$. For every codeword $c \in \mathcal{C}$, we include the subset $S_c = \{(i, c(i)) : i \in [d]\}$. It is easy to see that $n = |\mathcal{C}| = q^{rd}$ and $m = q \cdot d$. The only thing to verify is that for $c_1 \neq c_2 \in \mathcal{C}$, it holds that $|S_{c_1} \cap S_{c_2}| \leq \delta d$. This holds because $|S_{c_1} \cap S_{c_2}| = \#\{(i, \sigma) : i \in d,\ c_1(i) = c_2(i) = \sigma\} = d - \Delta(c_1, c_2) \leq \delta d$. □

Unfortunately, this approach of using error correcting codes falls shy of achieving the parameters we want. For example, it is known that Algebraic-Geometry codes of distance $(1 - \Theta(1/\sqrt{q}))$ with rate $\Theta(1/q)$ exist. Setting $q = O(k/\alpha)$, we get a $(n, m, d, 1, \alpha/k)$-Robust-UFF with parameter $m = \frac{q \log n}{r \log q} = \widetilde{O}\left(\frac{k^3 \log n}{\alpha^3}\right)$. On the other hand, Reed-Solomon codes with distance $(1 - \delta)$ and rate δ exist (by setting $d = q \geq 1/\delta$). Setting $\delta = \alpha/k$ and q s.t. $n = q^{\delta q}$, we get a $(n, m, d, 1, \alpha/k)$-Robust-UFF with parameter $m = \frac{q \log n}{r \log q} = \widetilde{O}\left(\frac{k^2 \log^2 n}{\alpha^2}\right)$.

In order to obtain $m = O\left(\frac{k^2 \log n}{\alpha^2}\right)$, we would require an error correcting code with distance $(1 - \delta)$ and rate $r \geq (\delta^2 q / \log q)$ (where $\delta = \alpha/k$). We are not aware of even probabilistic constructions of error correcting codes which satisfy these constraints on the parameters.