Electronic Colloquium on Computational Complexity, Revision 3 of Report No. 194 (2018)

ANTI-CONCENTRATION IN MOST DIRECTIONS

ANUP RAO AND AMIR YEHUDAYOFF

Abstract. We prove an anti-concentration bound for the inner product of two independent random vectors in the discrete cube. If, for example, A, B are subsets of the cube $\{\pm 1\}^n$ with $|A| \cdot |B| \geq 2^{1.01n}$, and $X \in A$ and $Y \in B$ are sampled independently and uniformly, then $\langle X, Y \rangle$ takes on any fixed value with probability at most $O(\frac{1}{\sqrt{n}})$. The proof provides a general framework for establishing anti-concentration in discrete domains. We describe applications to communication complexity, randomness extraction and additive combinatorics.

1. Introduction

Anti-concentration bounds establish that the distribution of outcomes of a random process is not concentrated in any small region. No single outcome is obtained too often. Such bounds play an important role in mathematics and computer science [11, 16, 1].

A well-known example is a sum of independent identically distributed random variables. If $Y \in \{\pm 1\}^n$ is uniformly distributed, then the probability that $\sum_{j=1}^n Y_j$ takes any specific value is at most $\binom{n}{\lceil n/2 \rceil}/2^n = O(\frac{1}{\sqrt{n}})$.

This was studied and generalized by Littlewood and Offord [11], Erdös [7], and many others. The classical Littlewood-Offord problem is about understanding the anti-concentration of the inner product $\langle x, Y \rangle = \sum_{j=1}^n x_j Y_j$, for arbitrary $x \in \mathbb{R}^n$ and $Y \in \{\pm 1\}^n$ chosen uniformly. Higher dimensional analogs were studied by Frankl and Furedi [8], Halász [9] and others.

It is interesting to understand the generality of this phenomenon (see [15] and references within). Anti-concentration certainly fails when the entropy of Y is not full. We can, for example, sample Y uniformly

A.R. is supported by the National Science Foundation under agreement CCF-1420268. This work was done while the author was visiting the Simons Institute for the Theory of Computing.
A.Y. is partially supported by ISF grant 1162/15. This work was done while the author was visiting the Simons Institute for the Theory of Computing.

ISSN 1433-8092

from the set of strings with exactly $\lceil n/2 \rceil$ entries that are 1. Then $\sum_j Y_j$ is always the same, yet Y has almost full entropy.

Can we somehow recover anti-concentration for high entropy samples? A natural setting is to consider the inner-product $\langle X, Y \rangle$ of two independent variables with high entropy. This indeed recovers anti-concentration, as the following theorem shows.

Theorem (Chakrabarti and Regev [4]). *There is a constant $c > 0$ such that if $A, B \subseteq \{\pm 1\}^n$ are each of size at least $2^{(1-c)n}$ and $X \in A, Y \in B$ are sampled uniformly and independently, then*

$$\Pr[|\langle X, Y \rangle| \leq c\sqrt{n}] \leq 1 - c.$$

Chakrabarti and Regev's proof uses the deep connection between the discrete cube and Gaussian space. They proved a geometric correlation inequality in Gaussian space, and translated it to the cube. Vidick [19] later simplified part of their argument, but stayed in the geometric setting. Sherstov [13] found a third proof that uses Talagrand's inequality from convex geometry [14] and ideas of Babai, Frankl and Simon from communication complexity [2].

This theorem is part of a more fundamental phenomenon. When studying anti-concentration, what we are ultimately interested in is proving point-wise estimates. Namely, we would like to control the *concentration probability* or the ℓ_∞ norm[1]

$$\|\mu\|_\infty = \max_{\omega \in \Omega} \mu(\omega)$$

(see [16] and references within).

We strengthen the theorem above. As long as the sum of the fractional min-entropies of X and Y is more than 1, the inner product is anti-concentrated.

Theorem 1. *For every $0 < \beta \leq 1$ and $0 < \delta < \frac{\beta}{6}$, there exists $C > 0$ such that the following holds. If $B \subseteq \{\pm 1\}^n$ is of size $2^{\beta n}$, then for all but $2^{n(1-\beta+\delta)}$ directions $x \in \{\pm 1\}$,*

$$\Pr_Y[\langle x, Y \rangle = k] \leq \tfrac{C}{\sqrt{n}}$$

for every number k.

Theorem 1 is essentially sharp. First, as mentioned above the $O(\frac{1}{n})$ bound is tight even when $A = B = \{\pm 1\}^n$. Second, if $B \subset \{\pm 1\}^n$ is the set of y's where the first $(1-\beta)n$ coordinates are set to 1 and

[1] We consider only finite probability spaces in this text.

$\sum_{j>(1-\beta)n} y_j = 0$, and $A \subset \{\pm 1\}^n$ is the set of x's where the last βn coordinates are set to 1 and $\sum_{j \le (1-\beta)n} x_j = 0$, then

$$|B| \approx \tfrac{1}{\sqrt{n}} 2^{\beta n} \quad \& \quad |A| \approx \tfrac{1}{\sqrt{n}} 2^{(1-\beta)n},$$

yet $\langle x, y \rangle = 0$ for every $x \in A$ and $y \in B$. There is no anti-concentration in this case, although $|A| \cdot |B|$ is roughly 2^n.

We now discuss the proof. We build a flexible framework for proving anti-concentration results in discrete domains. Think of the random variable $\langle X, Y \rangle$ as built in n steps. It starts as 0, and $X_d Y_d$ is added to $\langle X_{<d}, Y_{<d} \rangle$ to generate $\langle X_{\le d}, Y_{\le d} \rangle$. The analysis of the behavior of the system consists of three parts. In the first part, we use harmonic analysis to study the behavior of $\langle x, Y \rangle$ for x fixed and Y random. We are able to identify a collection of good x's for which the Fourier spectrum of the distribution of $\langle x, Y \rangle$ decays rapidly (Section 3 and Section 5). In the second part, we show that the number of bad x's is small by explicitly encoding them with few bits (Section 4). In the last part, we put together the harmonic analysis findings together with the encoding argument to deduce anti-concentration (Section 2).

There are several differences between our argument and the ones in [4, 19, 13]. The main difference is that the arguments from [4, 19, 13] are based, in one way or another, on the geometry of euclidean space. The arguments in [4, 19] prove a correlation inequality in gaussian space and translate it to the discrete world. It seems that such an argument can not yield effective bounds on the concentration probability in the discrete setting. A common ingredient to [4, 13] is showing that every set of large enough measure contains many almost orthogonal vectors (this is called 'identifying the hard core' in [13]). In [19] this part of the argument is replaced by a statement about a relevant matrix. Our argument does not contain any such step.

Applications.

Communication Complexity. Chakrabarti and Regev's main motivation was understanding the randomized communication complexity of the gap-hamming problem. The gap-hamming problem was introduced by Indyk and Woodruff in the context of streaming algorithms [10]. Proving lower bounds on its communication complexity was a central open problem for almost ten years, until Chakrabarti and Regev solved it [4]. Vidick [19] and Sherstov [13] later simplified the proof.

Our results also imply the lower bound for the randomized communication complexity of the gap-hamming problem (see e.g. [13]). As opposed to [4, 19, 13], the proof presented here lies entirely in the discrete domain. The underlying ideas may therefore be of independent interest.

Pseudorandomness. Randomness is a computational resource [18]. There are many sources of randomness, and some of them are *weak* or imperfect. Randomness extractors allow to use weak sources of randomness as if they were perfect.

The study of randomness extractors is about constructing explicit maps that transform weak sources of randomness to almost uniform outputs. The main goal is generating a uniform output in the most general scenario possible. This often requires ingenious constructions.

The scenario described above fits nicely in the context of *two-source extractors*. A two-source extractor maps two independent random variables X and Y with significant min-entropy to a single almost uniform output.

Chor and Goldreich [6] used Lindsey's lemma to show that inner product modulo two is a two-source extractor. The bit $\langle X, Y\rangle \bmod 2$ is close to a uniform random bit as long as $|A| \cdot |B| \gg 2^n$. Bourgain [3], Raz [12] and Chattopadhyay and Zuckerman [5] constructed two-source extractors with much better parameters.

This work can be interpreted as studying a related but somewhat different question. The high-level suggestion is to investigate what other pseudorandom properties known extractors satisfy.

We already know that inner product is an excellent two-source extractor. Now we also know that over the integers inner product is anti-concentrated. This is not as good as being uniform, but inner product is not uniform over the integers (it is binomial).

Theorem 1 in fact implies a stronger statement. It is the analog of "strong" extraction. Not only that $\langle X, Y\rangle$ is anti-concentrated, but for an overwhelming majority of fixing $X = x$, the inner product $\langle x, Y\rangle$ is anti-concentrated.

Additive Combinatorics. Additive combinatorics studies the behavior of sets under algebraic operation [17]. It has many deep results, and connections to other areas of mathematics, as well as many applications to computer science. We observe that our main result can be interpreted as saying that hamming spheres are far from sum-sets.

Replace $\{\pm 1\}$ by the field $\mathbb{F}_2$ with two elements. The set $\mathbb{F}_2^n$ becomes a vector space. The sum-set of $A \subseteq \mathbb{F}_2^n$ and $B \subseteq \mathbb{F}_2^n$ is

$$A + B = \{x + y : x \in A, y \in B\}.$$

If X is random in A and Y is random in B then $X + Y$ is distributed in $A + B$.

The cube $\mathbb{F}_2^n$ is endowed with a natural metric; the hamming distance. The sphere around 0 is the collection of all vectors with a fixed number of ones in them.

The inner product $I = \sum_j (-1)^{X_i}(-1)^{Y_i}$ is similar to the inner product studied above (here $X_i, Y_i \in \{0,1\}$). However, now I is n minus two times the hamming distance between X and Y.

We saw that if $|A| \cdot |B| > 2^{1.01n}$ then I is anti-concentrated. This means that the distribution of the hamming distance of $X + Y$ is anti-concentrated. The set $A + B$ is far from any sphere.

2. Anti-Concentration via Harmonic Analysis

We are ultimately interested in proving anti-concentration for a random variables that takes values in $\mathbb{Z}$. It is more convenient to work with a finite set instead of an infinite one. We work over the additive group $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ of integers modulo m. We aim at proving that the concentration probability is at most $O(\frac{1}{\sqrt{n}})$. So we choose $m = \lceil \sqrt{n} \rceil$.

Let Y be distributed in $\{\pm 1\}^n$. For $x \in \{\pm 1\}^n$, define the map $f_x : \mathbb{Z}_m \to \mathbb{R}$ by

$$f_x(\ell) = \Pr_Y \left[\langle x, Y \rangle = \ell \operatorname{\mathsf{mod}} m \right].$$

Our goal is to bound $\|f_x\|_\infty$ from above, for a typical choice of x.

We use harmonic analysis. The finite group $\mathbb{Z}_m$ has m characters. Each character $\psi_\ell : \mathbb{Z}_m \to \mathbb{C}$, for $\ell \in \mathbb{Z}_m$, is of the form

$$\psi_\ell(k) = \exp(-i\tfrac{2\pi}{m}\ell k).$$

The characters form a basis for the vector space of function from $\mathbb{Z}_m$ to $\mathbb{C}$. A function $f : \mathbb{Z}_m \to \mathbb{C}$ can be expressed as

$$f(k) = \tfrac{1}{\sqrt{m}} \sum_{\ell \in \mathbb{Z}_m} \widehat{f}(\ell)\psi_\ell(k).$$

The number

$$\widehat{f}(\ell) = \tfrac{1}{\sqrt{m}} \sum_{r \in \mathbb{Z}_m} f(r)\overline{\psi_\ell(r)}$$

is the ℓ'th Fourier coefficient of f.

The main technical result controls the Fourier spectrum of f_x for most x's.

Theorem 2. *For every $0 < \beta \leq 1$ and $0 < \delta < \frac{\beta}{6}$, there is $c > 0$ so that the following holds. Let $B \subseteq \{\pm 1\}^n$ be of size $2^{\beta n}$. For each $\ell \in \mathbb{Z}_m$, the number of $x \in \{\pm 1\}^n$ so that*

$$|\widehat{f_x}(\ell)| > \tfrac{1}{\sqrt{m}} \left(\exp(-cn \sin^2(\ell \tfrac{4\pi}{m})) + 2^{-\delta n} \right)$$

is at most $2^{n(1-\beta+\delta)}$.

The theorem is proved in Section 5.

Proof of Theorem 1. Theorem 2 and the union bound imply that for all but $m2^{n(1-\beta+\delta)}$ choices for x,

$$\begin{aligned}|f_x(k)| &\leq \tfrac{1}{\sqrt{m}} \cdot \sum_{\ell \in \mathbb{Z}_m} |\widehat{f}_x(\ell)| \\ &\leq 2^{-\delta n} + \tfrac{1}{m} \sum_{\ell \in \mathbb{Z}_m} \exp(-cn \sin^2(\ell \tfrac{4\pi}{m})).\end{aligned}$$

Each term in the sum occurs eight times (we are going around the circle twice). Using the inequality $\sin(\zeta) > \frac{\zeta}{\pi}$ for $0 < \zeta < \frac{\pi}{2}$, we can bound it by

$$\begin{aligned}&\leq 2^{-\delta n} + \tfrac{8}{m} \sum_{\ell=0}^{\infty} \exp(-cn\ell^2 (\tfrac{4\pi}{m})^2/\pi^2) \\ &\leq 2^{-\delta n} + \tfrac{8}{m} \sum_{\ell=0}^{\infty} \exp(-c\ell) \\ &\leq 2^{-\delta n} + \tfrac{8}{m} \cdot \frac{1}{1-\exp(-c)}.\end{aligned}$$

□

3. A Single Direction

In this section we analyze the behavior of f_x for a single $x \in \{\pm 1\}$. It is helpful to reveal the entropy of Y bit-by-bit. To keep track of it, define the following functions $\gamma_1, \ldots, \gamma_n$ from B to $\mathbb{R}$. For each $j \in [n]$, let

$$\gamma_j(y) = \gamma_j(y_{<j}) = \min_{u \in \{\pm 1\}} \Pr[Y_j = u | Y_{<j} = y_{<j}].$$

To understand the interaction between x and y, we use the following n measurements. For $j \in [n]$, define $\phi_j(x,y)$ to be half of the phase of the complex number

$$\mathop{\mathbb{E}}_{Y_{>j}|Y_j=1,Y_{<j}=y_{<j}} [\exp(i\theta \langle x_{>j}, Y_{>j}\rangle)] \cdot \overline{\mathop{\mathbb{E}}_{Y_{>j}|Y_j=-1,Y_{<j}=y_{<j}} [\exp(i\theta \langle x_{>j}, Y_{>j}\rangle)]},$$

where $\theta \in \mathbb{R}$ is some fixed angle (the choice of θ depends on the Fourier coefficient we are interested in). This quantity is not defined when $\Pr[Y_j = 1 | Y_{<j} = y_{<j}]$ is 1 or 0. In this case, set $\phi_j(x,y)$ to be 0. The number $\phi_j(x,y)$ is determined by $y_{<j}$ and $x_{>j}$. In the following we think of x as fixed, and of γ_j and ϕ_j as random variables that are determined by the random variable Y.

Lemma 3. *For each $x \in \{\pm 1\}^n$, every random variable Y over $\{\pm 1\}^n$, and every angle $\theta \in \mathbb{R}$,*

$$\left| \mathop{\mathbb{E}}_{Y} [\exp(i\theta \langle x, Y \rangle)] \right|^2 \leq \mathop{\mathbb{E}}_{Y} \left[\exp\left(-2 \sum_{j=1}^{n} \gamma_j \sin^2(\phi_j + x_j\theta) \right) \right].$$

Proof. The proof is by induction on n. We prove the induction base and step simultaneously. Express

$$\begin{aligned} \left| \mathop{\mathbb{E}}_{Y} \left[e^{i\theta\langle x, Y\rangle} \right] \right|^2 &= \left| \mathop{\mathbb{E}}_{Y_1} \left[e^{i\theta x_1 Y_1} \cdot \mathop{\mathbb{E}}_{Y_{>1}|Y_1} \left[e^{i\theta\langle x_{>1}, Y_{>1}\rangle} \right] \right] \right|^2 \\ &= \left| p_1 e^{i\theta x_1} Z_1 + p_{-1} e^{-i\theta x_1} Z_{-1} \right|^2, \end{aligned}$$

where for $u \in \{\pm 1\}$,

$$p_u = \Pr[Y_1 = u] \qquad \& \qquad Z_u = \mathop{\mathbb{E}}_{Y|Y_1=u} \left[e^{i\theta\langle x_{>1}, Y_{>1}\rangle} \right].$$

When $n = 1$, we have $Z_1 = Z_{-1} = 1$. Rearrange

$$\begin{aligned} &\left| p_1 e^{i\theta x_1} Z_1 + p_{-1} e^{-i\theta x_1} Z_{-1} \right|^2 \\ &= (p_1 e^{i\theta x_1} Z_1 + p_{-1} e^{-i\theta x_1} Z_{-1}) \cdot \overline{(p_1 e^{i\theta x_1} Z_1 + p_{-1} e^{-i\theta x_1} Z_{-1})} \\ &= p_1^2 |Z_1|^2 + p_{-1}^2 |Z_{-1}|^2 + p_1 p_{-1} (Z_1 \overline{Z_{-1}} e^{i2\theta x_1} + \overline{Z_1} Z_{-1} e^{-i2\theta x_1}) \\ &= p_1^2 |Z_1|^2 + p_{-1}^2 |Z_{-1}|^2 + 2 p_1 p_{-1} |Z_1||Z_{-1}| \cos(2\phi_1 + 2x_1\theta). \end{aligned}$$

The last equality holds by definition of ϕ_1. Using the inequality $a^2 + b^2 \geq 2ab$, we can bound this by

$$\begin{aligned} &\leq p_1^2 |Z_1|^2 + p_{-1}^2 |Z_{-1}|^2 + p_1 p_{-1} (|Z_1|^2 + |Z_{-1}|^2) \cos(2\phi_1 + 2x_1\theta) \\ &= p_1 |Z_1|^2 (p_1 + p_{-1} \cos(2\phi + 2x_1\theta)) + p_{-1} |Z_{-1}|^2 (p_{-1} + p_1 \cos(2\phi_1 + 2x_1\theta)) \\ &\leq (p_1 |Z_1|^2 + p_{-1} |Z_{-1}|^2)(1 - \gamma_1 + \gamma_1 \cos(\phi_1 + 2x_1\theta)) \\ &= \mathop{\mathbb{E}}_{Y_1} \left[|Z_{Y_1}|^2 \right] (1 - 2\gamma_1 \sin^2(\phi_1 + x_1\theta)) \\ &\leq \mathop{\mathbb{E}}_{Y_1} \left[|Z_{Y_1}|^2 \right] \exp\left(-2\gamma_1 \sin^2(\phi_1 + x_1\theta) \right). \end{aligned}$$

Above, γ_1 and ϕ_1 are constants that do not depend on Y. When $n = 1$, we proved the induction base. When $n > 1$, by induction, for each $u \in \{\pm 1\}$,

$$|Z_u|^2 \leq \mathop{\mathbb{E}}_{Y|Y_1=u} \left[\exp\left(-2 \sum_{j=2}^{n} \gamma_j \sin^2(\phi_j + x_j\theta) \right) \right].$$

□

4. A Few Bad Directions

Our goal is to control the Fourier spectrum of $\|f_x\|_\infty$. Lemma 3 suggests proving that the expression

$$\sum_{j=1}^{n} \gamma_j \sin^2(\phi_j + x_j\theta)$$

is typical large. Namely, that typically there are many j's for which both γ_j and $\sin^2(\phi_j + x_j\theta)$ are large.

Recall that Y is uniformly distributed in a set B of size $|B| = 2^{\beta n}$. Let $0 < \lambda < \frac{\beta}{2}$. Set $0 < \kappa < \frac{1}{2}$ and $\tau > 0$ to be parameters satisfying the conditions

$$H\left(\tfrac{1}{\log(1/\kappa)}\right) = \tau + H(\tau) = \lambda,$$

where H is the binary entropy function $H(\xi) = \xi \log(1/\xi) + (1 - \xi)\log(1/(1-\xi))$. Let $J(y)$ denote the set

$$J(y) = \{j \in [n] : \gamma_j(y) \geq \kappa\}.$$

Let $G(x, y)$ denote the set

$$G(x, y) = \left\{j \in J(y) : \sin^2(\phi_j(x, y) + x_j\theta) \geq \tfrac{\sin^2(2\theta)}{4}\right\}.$$

We start by showing that there are few y's with small $J(y)$.

Lemma 4. *The number of $y \in B$ with $|J(y)| \leq n(\beta - 3\lambda)$ is at most $2^{n(\beta-2\lambda)}$.*

Proof. Each $y \in B$ with $J(y) \leq n(\beta - 3\lambda)$ can be uniquely encoded by the following data:

- An element $u \in \{\pm 1\}^t$ with $t = \lfloor n(\beta - 3\lambda)\rfloor$.
- A subset $S \subseteq [n]$ of size $|S| \leq \frac{n}{\log(1/\kappa)}$.

Let us describe the encoding. The vector u is encodes the value of y in the coordinates in $J(y)$. Every coordinate $j \notin J(y)$ is included in S if and only if $\Pr[Y_j = y_j | Y_{<j} = y_{<j}] < \kappa$. Each string $y \in B$ has probability at least 2^{-n}. This implies that $\kappa^{|S|} \geq 2^{-n}$.

We can reconstruct y from u and S by iteratively computing y_1, then y_2, and so on, until we get to y_n. Whether or not $1 \in J(y)$ is determined even before we know y. If $1 \in J(y)$ then u tells us what y_1 is. If $1 \notin J(y)$ and $1 \in S$ then y_1 is the least likely value of ± 1. If $1 \notin J(y)$ and $1 \notin S$ then y_1 is the more likely value. Given the value of y_1, we can continue in the same way to compute the rest of y.

The number of choices for u is at most $2^{n(\beta-3\lambda)}$. The number of choices for S is at most

$$\sum_{0\le s\le n/\log(1/\kappa)} \binom{n}{s} \le 2^{nH\left(\frac{1}{\log(1/\kappa)}\right)} = 2^{\lambda n}.$$

□

Next, we argue that there are few x's for which there are many y's with small $G(x,y)$.

Lemma 5. *The number of $x \in \{\pm1\}^n$ for which*

$$\Pr_Y[|G(x,Y)| \le \tau n] \ge 2^{-\lambda n}$$

is at most $2^{n(1-\beta+6\lambda)}$.

Proof. The lemma is proved by double-counting the edges in a bipartite graph. Denote by $\mathcal{X}$ the set

$$\mathcal{X} = \left\{x : \Pr_Y[|G(x,Y)| \le \tau n] \ge 2^{-\lambda n}\right\}.$$

The left side of the bipartite graph is $\mathcal{X}$ and the right side is B. Connect $x \in \mathcal{X}$ to $y \in B$ by an edge if and only if $G(x,y) \le \tau n$. Denote by E the set of edges in this graph.

First, we bound the number of edges from below. The number of edges that touch each $x \in \mathcal{X}$ is at least $2^{-\lambda}|B|$. It follows that

$$|E| \ge 2^{-\lambda}|\mathcal{X}||B|.$$

Second, we bound the number of edges from above. By Lemma 4, the number of $y \in B$ so that $|J(y)| \le n(\beta-3\lambda)$ is at most $2^{-2\lambda n}|B|$. Below we prove that the number of edge that touch each y with $|J(y)| > n(\beta-3\lambda)$ is at most $2^{n(1-\beta+4\lambda)}$. It follows that

$$|E| \le 2^{-2\lambda}|\mathcal{X}||B| + |B|2^{n(1-\beta+4\lambda)}.$$

We can conclude that

$$\begin{aligned} 2^{-\lambda}|\mathcal{X}||B| &\le 2^{-2\lambda}|\mathcal{X}||B| + |B|2^{n(1-\beta+4\lambda)} \\ \Rightarrow |\mathcal{X}| &\le 2^{n(1-\beta+6\lambda)}. \end{aligned}$$

It remains to fix y so that $|J(y)| > n(\beta-3\lambda)$ and bound its degree in the graph from above. This is achieved by an encoding argument. Encode each x that is connected to y by an edge using the following data:

- A vector $u \in \{\pm1\}^t$ with $t = \lfloor n(1-\beta+3\lambda)\rfloor$.
- The set $G(x,y)$.
- A vector $v \in \{\pm1\}^s$ with $s = \lfloor \tau n\rfloor$.

Let us describe the encoding. The vector u specifies the values of x on coordinates not in $J(y)$. There are at most $n-n(\beta-3\lambda) = n(1-\beta+3\lambda)$ such coordinates. The size of $G(x,y)$ is at most τn. The vector v specifies the values of x in the coordinates of $G(x,y)$, written in descending order. There are at most τn such coordinates.

The decoding of x from u, S and v is done as follows. Decode the coordinates of x in descending order from n to 1. If $n \notin J(y)$ then we read the value of x_n from u. If $n \in J(y)$ and $n \in G(x,y)$, we decode x_n by reading its value from v. If $n \in J(y)$ and $n \notin G(x,y)$, then $\sin^2(\phi_n(x,y) + x_n\theta) < \frac{\sin^2(2\theta)}{4}$. The number $\phi_n(x,y)$ does not depend on x. The following claim implies that there is exactly one value of x_n that satisfies this property.

Claim 6.

$$\max\{|\sin(\phi_n(x,y)+\theta)|, |\sin(\phi_n(x,y)-\theta)|\} \geq \frac{|\sin(2\theta)|}{2}$$

Proof. Consider the map

$$\varphi \mapsto g(\varphi) = \max\{|\sin(\varphi+\theta)|, |\sin(\varphi-\theta)|\}.$$

The minimum of this map is attained at one of four points $0, \pi/2, \pi, 3\pi/2$. Symmetry allows to remove two of the points and obtain

$$g(\varphi) \geq \min\{g(0), g(\pi/2)\} \geq |\sin(\theta)\sin(\tfrac{\pi}{2}-\theta)| = \frac{|\sin(2\theta)|}{2}.$$

□

The claim implies that we can indeed reconstruct x_n. Given x_n, we can similarly reconstruct $x_{n-1}, \ldots, x_1$.

Finally, the total number of choices for u, S, v is at most

$$2^{n(1-\beta+3\lambda)+nH(\tau)+\tau n} = 2^{n(1-\beta+4\lambda)}.$$

□

5. A Single Fourier Coefficient

Proof of Theorem 2. Recall that

$$\begin{aligned}\widehat{f}_x(\ell) &= \tfrac{1}{\sqrt{m}} \sum_{r\in\mathbb{Z}_m} \Pr_Y[\langle x, Y\rangle = r \bmod m]\exp(i\tfrac{2\pi}{m}\ell r) \\ &= \tfrac{1}{\sqrt{m}} \mathop{\mathbb{E}}_Y\left[\exp(i\tfrac{2\pi}{m}\ell\langle x, Y\rangle)\right].\end{aligned}$$

By Lemma 3 with $\theta = \frac{2\pi}{m}\ell$,

$$|\widehat{f}_x(\ell)|^2 \leq \tfrac{1}{m}\mathop{\mathbb{E}}_Y\left[\exp\left(-2\sum_{j=1}^n \gamma_j \sin^2(\phi_j + x_j\tfrac{2\pi}{m}\ell)\right)\right].$$

Whenever x is such that

$$\Pr_Y[G(x,Y) \le \tau n] < 2^{-\lambda n},$$

we can bound

$$|\widehat{f}_x(\ell)| \le \tfrac{1}{\sqrt{m}} \left(\exp\left(-\tau n \cdot \kappa \frac{\sin^2(\frac{4\pi}{m}\ell)}{4} \right) + 2^{-\lambda n} \right).$$

Lemma 5 promises that there are at most $2^{n(1-\beta+6\lambda)}$ choices for x that does not satisfy this condition. □

Acknowledgements. We wish to thank Oded Regev, Avishay Tal and David Woodruff for helpful conversations.

References

[1] S. Aaronson and A. Arkhipov. The computational complexity of linear optics. In *STOC*, pages 333–342, 2011.

[2] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *FOCS*, pages 337–347, 1986.

[3] J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(01):1–32, 2005.

[4] A. Chakrabarti and O. Regev. An optimal lower bound on the communication complexity of gap-hamming-distance. *SIAM Journal on Computing*, 41(5):1299–1317, 2012.

[5] E. Chattopadhyay and D. Zuckerman. Explicit two-source extractors and resilient functions. In *STOC*, pages 670–683, 2016.

[6] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[7] P. Erdös. On a lemma of littlewood and offord. *Bulletin of the American Mathematical Society*, 51(12):898–902, 1945.

[8] P. Frankl and Z. Furedi. Solution of the littlewood-offord problem in high dimensions. *Annals of Mathematics*, pages 259–270, 1988.

[9] G. Halász. Estimates for the concentration function of combinatorial number theory and probability. *Periodica Mathematica Hungarica*, 8(3-4):197–211, 1977.

[10] P. Indyk and D. Woodruff. Tight lower bounds for the distinct elements problem. In *FOCS*, pages 283–288, 2003.

[11] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation (iii). *Rec. Math. (Mat. Sbornik) N.S*, 12(3):277–286, 1943.

[12] R. Raz. Extractors with weak random seeds. In *STOC*, pages 11–20, 2005.

[13] A. A. Sherstov. The communication complexity of gap hamming distance. *Theory of Computing*, 8(1):197–208, 2012.

[14] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, 81(1):73–205, 1995.

[15] T. Tao and V. Vu. A sharp inverse Littlewood-Offord theorem. *Random Structures & Algorithms*, 37(4):525–539, 2010.

[16] T. Tao and V. H. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Annals of Mathematics*, pages 595–632, 2009.
[17] T. Tao and V. H. Vu. Additive Combinatorics. *Cambridge University Press.*
[18] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. In *FOCS*, pages 32–42, 2000.
[19] T. Vidick. A concentration inequality for the overlap of a vector on a large set. *Chicago Journal of Theoretical Computer Science*, 1:1–12, 2012.
[20] A. C. Yao. Lower bounds by probabilistic arguments. In *FOCS*, pages 420–428, 1983.

Department of Computer Science, University of Washington
Email address: anuprao@cs.washinton.edu

Department of Mathematics, Technion-IIT
Email address: amir.yehudayoff@gmail.com

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il