Electronic Colloquium on Computational Complexity, Report No. 56 (2025)

A Near-Optimal Polynomial Distance Lemma Over Boolean Slices

Prashanth Amireddy* Amik Raj Behera† Srikanth Srinivasan ‡ Madhu Sudan§

April 28, 2025

Abstract

The celebrated Ore-DeMillo-Lipton-Schwartz-Zippel (ODLSZ) lemma asserts that n-variate non-zero polynomial functions of degree d over a field $\mathbb{F}$, are non-zero over any "grid" (points of the form S^n for finite subset $S \subseteq \mathbb{F}$) with probability at least $\max\{|S|^{-d/(|S|-1)}, 1 - d/|S|\}$ over the choice of random point from the grid. In particular, over the Boolean cube ($S = \{0,1\} \subseteq \mathbb{F}$), the lemma asserts non-zero polynomials are non-zero with probability at least 2^{-d}. In this work we extend the ODLSZ lemma optimally (up to lower-order terms) to "Boolean slices" i.e., points of Hamming weight exactly k. We show that non-zero polynomials on the slice are non-zero with probability $(t/n)^d(1 - o_n(1))$ where $t = \min\{k, n-k\}$ for every $d \leqslant k \leqslant (n-d)$. As with the ODLSZ lemma, our results extend to polynomials over Abelian groups. This bound is tight upto the error term as evidenced by multilinear monomials of degree d, and it is also the case that some corrective term is necessary. A particularly interesting case is the "balanced slice" ($k = n/2$) where our lemma asserts that non-zero polynomials are non-zero with roughly the same probability on the slice as on the whole cube.

The behaviour of low-degree polynomials over Boolean slices has received much attention in recent years. However, the problem of proving a tight version of the ODLSZ lemma does not seem to have been considered before, except for a recent work of Amireddy, Behera, Paraashar, Srinivasan and Sudan (SODA 2025), who established a sub-optimal bound of approximately $((k/n) \cdot (1 - (k/n)))^d$ using a proof similar to that of the standard ODLSZ lemma.

While the statement of our result mimics that of the ODLSZ lemma, our proof is significantly more intricate and involves spectral reasoning which is employed to show that a natural way of embedding a copy of the Boolean cube inside a balanced Boolean slice is a good sampler.

*School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Supported in part by a Simons Investigator Award and NSF Award CCF 2152413 to Madhu Sudan and a Simons Investigator Award to Salil Vadhan. Email: pamireddy@g.harvard.edu

†Department of Computer Science, University of Copenhagen, Denmark. Supported by Srikanth Srinivasan's start-up grant from the University of Copenhagen. Email: ambe@di.ku.dk

‡Department of Computer Science, University of Copenhagen, Denmark. This work was funded by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA). Email: srsr@di.ku.dk

§School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Supported in part by a Simons Investigator Award, NSF Award CCF 2152413 and AFOSR award FA9550-25-1-0112. Email: madhu@cs.harvard.edu

ISSN 1433-8092

Contents

1 Introduction

The Ore-DeMillo-Lipton-Schwartz-Zippel (ODLSZ) [Ore22; DL78; Zip79; Sch80] lemma captures the basic algebraic fact that a low-degree polynomial does not have many roots on a "nice set" of points. The standard nice set for this lemma is a grid S^n (where S is a finite subset of a field) and a version of this lemma states that no non-zero degree-d polynomial can vanish on more that $d|S|^{n-1}$ points. This is easily seen to be tight: Take, for example, a univariate polynomial that has d roots in S.

There also exist useful variants of this lemma for the case where $|S| < d$. The example above shows that in general a degree-d polynomial can vanish over all of S^n and so some further condition is necessary. The most obvious condition is to simply force the polynomial to be non-zero on the grid S^n. In the setting of the Boolean cube, i.e. $S = \{0, 1\}$, which is the setting we study, this is equivalent to considering non-zero *multilinear* polynomials of degree d. In this setting (a variant of) the ODLSZ lemma states that a non-zero multilinear polynomial of degree d is non-zero on at least 2^{n-d} points of $\{0, 1\}^n$. Again, this is tight: Take, e.g., a multilinear monomial of degree d.

Though both these forms of the ODLSZ lemma are simple statements with easy inductive proofs, they have many different applications in the design of randomized algorithms [RV89], probabilistically checkable proofs [BFLS91; ALMSS98], pseudorandom constructions [DKSS13; GRS23], Boolean function analysis [NS92], data communication [ABCO88], small-depth circuit lower bounds [PS90; HRRY19] and extremal combinatorics [SS08].

In this paper, we extend the ODLSZ lemma to a different nice set namely the *Boolean slice*, which is an important subset of the Boolean cube $\{0, 1\}^n$. For a parameter k, we use $\{0, 1\}^n_k$ to denote the kth Boolean slice, i.e., the set of points in the cube of Hamming weight exactly k. The behavior of low-degree polynomials on Boolean slices has received quite a bit of attention recently with motivations from learning theory [OW13], Boolean function analysis [Wim14; Fil16; FKMW18; FI19], property testing [DDGKS17; KLMZ24], circuit lower bounds [HRRY19], and local decoding algorithms [ABPSS24]. However, as far as we know, the natural question of finding a tight version of the ODLSZ lemma over Boolean slices has not been considered before. This is the question we address in this paper.

More precisely, we consider the following question:

Given a polynomial P of degree at most d that does not vanish on $\{0, 1\}^n_k$, how many zeroes can have P have in this set?

This question makes sense when $d \leqslant t := \min\{k, n - k\}$, since any function on $\{0, 1\}^n_k$ can be expressed as a polynomial of degree t.

We give a near-optimal answer to this question for low-degree polynomials. More precisely, our main theorem is stated below. It holds for polynomials over any field and even in the case where the coefficients come from an *Abelian group*[1] (as is also true of the standard version of ODLSZ lemma over the Boolean cube).

[1] A multilinear polynomial over an Abelian group G is of the form $\sum_{S\subseteq[n]} a_S \prod_{i\in S} x_i$ where $a_S \in G$ for each S. Polynomials over such domains appear naturally in applications to circuit complexity [BHLR19] and additive combinatorics [TZ12].

Theorem 1.1 (Main Theorem). *There exists an absolute constant $\varepsilon > 0$ so that the following holds. Fix an arbitrary Abelian group G and a degree parameter $d \in \mathbb{N}$. For all natural numbers n and k such that $d \leqslant k \leqslant n - d$, the following holds whenever $d \leqslant t^{\varepsilon}$ where $t = \min\{k, n-k\}$:*
For any degree-d polynomial $P : \{0,1\}_k^n \to G$ that does not vanish on $\{0,1\}_k^n$, we have

$$\Pr_{\mathbf{x} \sim \{0,1\}_k^n}[P(\mathbf{x}) \neq 0] \geqslant \left(\frac{t}{n}\right)^d \cdot \left(1 - \frac{1}{t^{\varepsilon}}\right).$$

We prove Theorem 1.1 in Section 4.2.
At a high level, the uniform distribution on $\{0,1\}_k^n$ is similar to the (k/n)-biased distribution, i.e. the distribution where each coordinate is independently 1 with probability k/n. With slight modifications to the proof of ODLSZ lemma, one can show (see for example [DFH17, Claim 6.8]) that the probability of sampling a non-zero point from (k/n)-biased distribution is $(t/n)^d$, where $t = \min\{k, n-k\}$. The bound given by Theorem 1.1 is equal to this bound up to small error terms.

Tightness. The bound given is easily seen to be nearly tight using essentially the same example as in the case of the Boolean cube. For $k \leqslant n/2$, the monomial $x_1 \cdots x_d$ is non-zero with probability approximately $(k/n)^d$, and there is a similar example for $k > n/2$. Moreover, it is also possible to see that for certain k, an error term is required. For example, assume that G is the finite field $\mathbb{F}_2$, $k = n/2$ and $d = 1$. Then the linear polynomial $x_1 + x_2 + 1$ is non-zero with probability $1/2 - \Theta(1/n)$, implying that the monomial does not yield exactly the optimal bound. In the case that the degree $d = 1$, we can improve the error parameter and show a bound of $t/n - 1/n$ (Theorem 6.1 in Section 6).

Proof Techniques. The standard proofs of the ODLSZ lemma follow a simple inductive strategy, using the obvious univariate case for both the base case and each inductive step of the argument. The recent work of Amireddy, Behera, Paraashar, Srinivasan and Sudan [ABPSS24] used a similar idea to show the following sub-optimal bound. Unfortunately, it is not clear how to make the inductive strategy work for the slice to get a tight answer.

Lemma 1.2 (Suboptimal distance lemma for slices). *[ABPSS24, Lemma 5.1.6]. For every Abelian group G and non-negative integers d, k, n with $n \geqslant 1$ and $d \leqslant k \leqslant n - d$ the following holds: For every degree-d polynomial $P : \{0,1\}_k^n \to G$ that does not vanish on $\{0,1\}_k^n$, we have*

$$\Pr_{\mathbf{x} \sim \{0,1\}_k^n}[P(\mathbf{x}) \neq 0] \geqslant \binom{n-2d}{k-d} \Big/ \binom{n}{k}.$$

In particular, for $k = n/2$ (for an even n), the above probability is at least 4^{-d}.

A computation shows that for small d, the above implies that the fraction of points in $\{0,1\}_k^n$ where P does not vanish is at least $((k/n) \cdot (1 - (k/n)))^d$ (up to small error terms). When $k = n/2$, for example, this bound is 4^{-d} which is quadratically worse than Theorem 1.1.

To get the tight bound, we use a very different approach. We start with the above suboptimal bound, but combine it with spectral techniques, which we elaborate on next. Note that if the slice

$k = n/2$, i.e. the balanced slice, then we get a bound of nearly $1/2^d$ which is essentially the same as the ODLSZ lemma over the Boolean cube $\{0,1\}^n$ (Theorem 2.1). To prove this, the high-level idea is to consider the process of choosing a random subcube in the balanced Boolean slice $\{0,1\}^n_{n/2}$ as follows: pair the n coordinates into $n/2$ pairs uniformly at random, and in each such pair $\{x_i, x_j\}$, identify x_i with the Boolean negation of x_j, i.e. $1 - x_j$. This gives us a random embedding of an $n/2$-dimensional cube in the slice $\{0,1\}^n_{n/2}$ and the polynomial P restricts to a degree-d polynomial Q on this subcube. If we could guarantee that Q was always non-zero, then the standard ODLSZ lemma on the cube would give us the desired statement. Unfortunately, there are subcubes on which P could be identically zero. The main technical lemma is to show that Q is non-zero with high probability: intuitively, this is because the random process above is a good *sampler* of the balanced slice, i.e. the points in the randomly chosen subcube behave essentially like independent samples of the balanced slice.

Formally, the technical lemma is a statement about the approximate pairwise independence of two random points of the chosen subcube. We show (see Lemma 3.2) that the probability that two random points of this subcube lie in a set of density ρ is roughly ρ^2. This is done by analyzing a natural weighted graph Γ on the balanced slice defined by the above sampling process. We show this via two arguments, depending on the regime of the degree parameter d.

For $d \leqslant C \log n$ for a constant $C > 0$, the main technical lemma follows from the use of the Expander mixing lemma [AC88], which implies such a statement using bounds on the second-largest eigenvalue of the graph. To analyze the second-largest eigenvalue of Γ, we show that it can be embedded (as an induced subgraph) in a Cayley graph defined on the subgroup of $\mathbb{F}_2^n$ defined by points of even Hamming weight. The latter is easier to analyze using Boolean Fourier analysis, and an application of the eigenvalue interlacing theorem allows us to bound the eigenvalues of Γ. See Section 3.1 for more details. This easier case of the lemma is already interesting: for instance, it yields a different (arguably easier) proof of a junta theorem on the Boolean slice [FI19], analogous to a well-known theorem of Nisan and Szegedy [NS92].

For $d = n^\gamma$ for a small constant $\gamma > 0$, we need to strengthen the guarantee of the sampler. To do so, we use the fact that the adjacency matrix for Γ can be spectrally upper-bounded by another matrix[2] that satisfies a *Hypercontractive inequality* (see Lemma 3.11). Intuitively, this is stronger than an eigenvalue bound, as the latter measures only the worst-case expansion of the underlying graph, while the former gives us stronger bounds on the expansion of smaller sets. Using this inequality alongside the Expander mixing lemma yields the desired pairwise independence. See Section 3.2 for more details.

For imbalanced slices, i.e. $k \neq n/2$, we reduce to the balanced case via a random restriction idea (see Section 4). The main conceptual idea is to obtain a basis for the space of polynomial functions on a slice. We note, essentially using an argument of Wilson [Wil90], that for many distinct slices, the space of homogeneous multilinear monomials of degree d forms a basis for the space of polynomials of degree d on the slice (see Claim 4.5). Unlike other known bases for this space [Fil16], this idea also works over fields of positive characteristic and even over cyclic groups of prime power order. For such 'good' slices $k \leqslant n/2$, we reduce to a $2k$-dimensional cube via a random restriction (see Lemma 4.6), which can easily be seen to leave the polynomial non-zero with probability $(2k/n)^d$.

[2]The technically accurate descriptor for this matrix is the 'Noise operator in the Bose-Mesner algebra of the Johnson scheme.' See Section 3 for details.

Invoking the balanced case now concludes the lemma for the good slices.

Finally, to extend the main theorem to all slices, we note that for any slice k, there is a good slice not too 'far away' (in the range $[k - \mathcal{O}(d), k]$) (see Lemma 4.8). By setting a few variables at random to 1, we are able to reduce to a good slice.

Related Work. As mentioned above, the study of low-degree polynomials over Boolean slices has received much attention in recent years. Closely related to this work is the work of Filmus [Fil16] that constructs a basis for the space of real-valued degree-d polynomial functions over general Boolean slices. A recent result of Kalai, Lifshitz, Minzer and Ziegler [KLMZ24] constructs a *dense model* for the balanced slice $\{0,1\}^n_{n/2}$ under the Gowers norm U_d; in particular, this implies that there is a subset S of $\{0,1\}^n$ of constant density such that any polynomial of degree-d has the same density over S as it does over the balanced slice. In principle, both these works should be useful in order to prove a version of the ODLSZ lemma over Boolean slices. However, we note that each of these results is applicable over different domains ($\mathbb{R}$ or $\mathbb{F}_2$) while we prove a unified statement that holds over any Abelian group (and in particular over all fields).

1.1 Applications of Optimal Distance Lemma

To give some idea of the applicability of the ODLSZ lemma over the slice, we prove some variants of well-known theorems in combinatorics and Boolean function analysis.

Hyperplane covering. Given a subset S of the cube $\{0,1\}^n$, we define the *exact cover number of* S, denoted $\mathrm{ec}_n(S)$ to be the minimum number of hyperplanes (over some field $\mathbb{F}$) such that their union intersects $\{0,1\}^n$ exactly in the set S. A classical result of Alon and Füredi shows that for S being the cube with a single point removed, $\mathrm{ec}_n(S) = n$. This combinatorial result, which easily follows from with ODLSZ lemma over the cube, has seen many subsequent generalizations (e.g. [CH20; SW22; BBDM23]).

Using just the sub-optimal version of the ODLSZ lemma (Lemma 1.2), we immediately get an optimal version of the hyperplane covering over a Boolean slice $\{0,1\}^n_k$ with a missing point, instead of the whole Boolean cube $\{0,1\}^n$. More precisely, for $S \subseteq \{0,1\}^n_k$, let $\mathrm{ec}_{n,k}(S)$ be the minimum number of hyperplanes (over some fixed field $\mathbb{F}$) such that their union intersects $\{0,1\}^n_k$ exactly in the set S. Following the idea of [AF93], we have the following.

Theorem 1.3. *Let n, k be natural numbers with $k \in [n]$. Fix an arbitrary point $\mathbf{a} \in \{0,1\}^n_k$. Then $\mathrm{ec}_{n,k}(\{0,1\}^n_k \backslash \{\mathbf{a}\}) = \min\{k, n-k\}$.*

Proof of Theorem 1.3. Without loss of generality, we assume that $k \leqslant n/2$ and $\mathbf{a} = 1^k 0^{n-k}$. Let S denote $\{0,1\}^n_k \backslash \{\mathbf{a}\}$.

It is easy to see that $\mathrm{ec}_{n,k}(S) \leqslant k$. The hyperplanes $H_i = \{x_i = 0\}$ for $i \in [k]$ cover exactly the points in S.

For the lower bound, assume for the sake of contradiction that there exists $m < k$ hyperplanes $H_i = \{\ell_i(\mathbf{x}) = 0\}$ (here $\ell_i(\mathbf{x})$ denotes a degree-1 polynomial and $i \in [m]$) covering exactly the points in S. Then the polynomial $P(\mathbf{x}) := \prod_{i=1}^m \ell_i(\mathbf{x})$ is non-zero at exactly one point of $\{0,1\}^n_k$.

However, by Lemma 1.2, P must be non-zero at at least $\binom{n-2m}{k-m} > 1$ points (since $m < k \leqslant n/2$) of $\{0,1\}^n_k$. Hence we arrive at a contradiction. ■

A junta theorem for the slice. Nisan and Szegedy [NS92] showed that any Boolean function on $\{0,1\}^n$ that has degree d over $\mathbb{R}$ depends on $\mathcal{O}(d2^d)$ variables, i.e. it is a $\mathcal{O}(d2^d)$-junta. Chiarelli, Hatami, and Saks [CHS20] improved the bound to $\mathcal{O}(2^d)$. Filmus and Ihringer [FI19] extended this result to slices and showed that for a suitable range of k, any degree-d (over $\mathbb{R}$) Boolean function on the slice k is a restriction of a degree-d function on $\{0,1\}^n$. Along with the result of [CHS20], this implies that such a function is an $\mathcal{O}(2^d)$-junta. While the results of [NS92; CHS20] are fairly elementary, the theorem of [FI19] is more involved, relying on the Log-Sobolev inequality and Hypercontractivity for the Boolean slice [LY98; DS96].

Using Theorem 1.1, we show that we can avoid the use of advanced analytic techniques[3] in the proof of [FI19], and give a direct proof (following the proof of [NS92]) of the fact that any degree-d Boolean function on the balanced slice $\{0,1\}^n_{n/2}$ depends on $\mathcal{O}(d2^d)$ variables (see Lemma 5.3). Plugging this into the proof of [FI19], we can again recover the optimal bound of $\mathcal{O}(2^d)$. More details can be found in Section 5.

Organization of the paper. We provide basic definitions and other preliminaries in Section 2. We give the proof details of the distance lemma over the balanced slice (i.e., $k = n/2$) in Section 3 and then use this to get the same over *all* slices in Section 4, thus finishing the proof of our main theorem Theorem 1.1. In Section 5, we present our alternate proof of the junta theorem over slices. Finally, we obtain an improvement of our main theorem for the case of linear functions (i.e., $d = 1$) as Theorem 6.1 in Section 6.

2 Preliminaries

Notations. Let $(G, +)$ denote an Abelian group G with addition as the binary operation. For any $g \in G$, let $-g$ denote the inverse of $g \in G$. For any $g \in G$ and integer $a \geqslant 0$, $a \cdot g$ (or simply ag) is the shorthand notation of $g + \ldots + g$ (taken a times), and $-ag$ denotes $a \cdot (-g)$.

For any $\mathbf{x} \in \{0,1\}^n$, $|\mathbf{x}|$ denotes the Hamming weight of $\mathbf{x}$. For any $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$, let $\Delta(\mathbf{x}, \mathbf{y})$ denote the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$, i.e. $\Delta(\mathbf{x}, \mathbf{y}) = |\{i \in [n] \mid x_i \neq y_i\}|$. For natural numbers n and $k \leqslant n$, let $\{0,1\}^n_k$ denote the subset of strings in $\{0,1\}^n$ of Hamming weight exactly k.
We denote the set of functions $f : \{0,1\}^n \to G$ that can be expressed as a multilinear polynomial of degree d, with the coefficients being in G by $\mathcal{P}_d(n, G)$. We also consider functions $f : \{0,1\}^n_k \to G$. We denote the set of functions on $\{0,1\}^n_k$ that can be expressed as a multilinear polynomial of degree d with the coefficients in G by $\mathcal{P}_d(n, k, G)$. We will simply write $\mathcal{P}_d(n, k)$ when G is clear from the context.

For any natural numbers n and $k \leqslant n$, U_n denotes the uniform distribution on $\{0,1\}^n$ and $U_{n,k}$ denotes the uniform distribution on $\{0,1\}^n_k$. For a growing parameter n, $o_n(1)$ denotes a function that goes to 0 as n grows large.

[3] We have two proofs of our main theorem. In the general case where d can be as large as $n^{\Omega(1)}$, our proof also relies on hypercontractivity. However, in the case that $d \leqslant C \log n$, which is also the main case of interest for junta theorems, our proof needs only basic Fourier analysis over the Boolean cube and the eigenvalue interlacing theorem.

Basic Tools. We start with the standard ODLSZ lemma over the Boolean cube.

Theorem 2.1 (ODLSZ lemma over $\{0,1\}^n$)**.** *Let G be any Abelian group and let $P \in \mathcal{P}_d(n, G)$ be any non-zero polynomial. Then*

$$\Pr_{\mathbf{x} \sim U_n}[P(\mathbf{x}) \neq 0] \geqslant \frac{1}{2^d}.$$

Another important tool we require is Lucas's theorem which allows us to compute binomial coefficients modulo a prime p.

Lemma 2.2 (Lucas's Theorem [Luc78])**.** *Let p be a prime number and a and b be any two natural numbers. Denote a and b in their unique p-ary representations as:*

$$a = \sum_{i=0}^{\ell-1} a_i p^i, \qquad b = \sum_{i=0}^{\ell-1} b_i p^i, \qquad a_i, b_i \in \{0, 1, \ldots, p-1\}$$

Then,

$$\binom{a}{b} \equiv \prod_{i=0}^{\ell-1} \binom{a_i}{b_i} \mod p,$$

where we define $\binom{x}{y}$ to be 0 if $x < y$.

We will need the following standard facts about expanders and Cayley graphs. We refer the reader to the survey [HLW06] for more details.

Definition 2.3 (Weighted Cayley Graph)**.** *Let $(G, +)$ be a finite Abelian group and $w : G \to \mathbb{R}^{\geqslant 0}$ be a* weight *function (we refer to the elements of non-zero weight as* generators*). We say that a weighted graph $\Gamma = \Gamma(G, w)$ defined as follows is a* weighted Cayley graph *over G.*

- *The vertices of Γ are the elements of G.*
- *For every $g, g' \in G$, we add an edge $(g, g + g')$ with weight $w(g')$ to Γ.*

The following lemma gives us a way of computing the eigenvalues of the adjacency matrix of weighted Cayley graphs over Abelian groups.

Lemma 2.4 (Eigenvalues of Cayley graphs, see e.g. [HLW06])**.** *Let $\Gamma = \Gamma(G, w)$ be a weighted Cayley graph over a finite Abelian group G, where $w : G \to \mathbb{R}^{\geqslant 0}$ is the corresponding weight function. Let $\chi : G \to \mathbb{C}^\times$ be an arbitrary group homomorphism (which we will refer to as a* character*). Then, χ is an eigenvector of the adjacency matrix of Γ with eigenvalue equal to $\sum_{g \in G} w(g)\chi(G)$.*

Following is a consequence of the *expander mixing lemma.*

Lemma 2.5 (Expander mixing lemma, see e.g. [HLW06] Lemma 2.5)**.** *For a symmetric random walk matrix W over vertices V and every subset $S \subseteq V$, it holds that*

$$\Pr_{u \sim V, v \sim N(u)}[u \in S \text{ and } v \in S] \leqslant \left(\frac{|S|}{|V|}\right)^2 + \mu(W) \cdot \frac{|S|}{|V|},$$

where $N(u)$ denotes the distribution over V corresponding to taking a step from u according to W (i.e., the u-th row of W).

3 Distance Lemma for the Balanced Slice

In this section, we state the main technical lemma of our proof for Theorem 1.1. It is a statement on the "expansion" property of a graph $\{0,1\}^n_{n/2}$, where the edge weights are given by a random process. We start by describing a random process that maps a string in $\{0,1\}^n_{n/2}$ to a string in $\{0,1\}^n_{n/2}$. In this section, we will always assume that n is an even number.

Definition 3.1 (The map Γ)**.** *Let $\mathbf{a} \in \{0,1\}^{n/2}$ and $\mathbf{u} \in \{0,1\}^n_{n/2}$. Let $\mathbf{u}^{-1}\{0\}$ denote the set of coordinates where $\mathbf{u}$ is 0, i.e. $\mathbf{u}^{-1}\{0\} = \{i \in [n] \mid u_i = 0\}$. Similarly we have $\mathbf{u}^{-1}\{1\}$. let $\mathbf{u}^{-1}\{1\} = \{i_1, \ldots, i_{n/2}\}$.*
For any perfect matching $\mathcal{M}$ between $\mathbf{u}^{-1}\{0\}$ and $\mathbf{u}^{-1}\{1\}$ ($\mathcal{M}$ is a bijection between these two sets), the function $\Gamma(\mathbf{u}, (\mathcal{M}, \mathbf{a}))$ is a balanced string $\mathbf{v} \in \{0,1\}^n_{n/2}$ defined as follows:
For every $k \in [n/2]$, $v_{i_k} = u_{i_k} \oplus a_k$ and $v_{\mathcal{M}(i_k)} = u_{\mathcal{M}(i_k)} \oplus a_k$.

In simple words, for every matching between the 0-coordinates and 1-coordinates of $\mathbf{u}$ and a string $\mathbf{a} \in \{0,1\}^{n/2}$, we get a new balanced string $\mathbf{v}$ by flipping the endpoints of a subset of matching edges. Here the subset of matching edges whose endpoints are flipped is given by the string $\mathbf{a}$. Following is an example for $n = 8$.

Example. Let $\mathbf{u} = 10101010$. Here $\mathbf{u}^{-1}\{0\} = \{2,4,6,8\}$ and $\mathbf{u}^{-1}\{1\} = \{1,3,5,7\}$. Let $\mathcal{M} = ((2,3),(6,1),(4,5),(8,7))$ and $\mathbf{a} = 0110$. Then $\Gamma(\mathbf{u}, (\mathcal{M}, \mathbf{a})) = \mathbf{v} = 00110110$ (endpoints of the 2^{nd} matching edge $(6,1)$ and the 3^{rd} matching edge $(4,5)$ are flipped).

Next, we define a weighted graph on all the balanced strings with weights representing the probability of going from one balanced string to another for a random matching $\mathcal{M}$ and a random string $\mathbf{a}$ (using the map Γ).
Let $n' = |\binom{n}{n/2}|$ denote the cardinality of the set of balanced strings $\{0,1\}^n_{n/2}$. Let G denote a weighted complete graph on n' vertices, where the vertices denote strings in $\{0,1\}^n_{n/2}$. For any two distinct balanced strings $\mathbf{u}, \mathbf{v} \in \{0,1\}^n_{n/2}$, the weight of the edge $(\mathbf{u}, \mathbf{v})$, denoted by $w(\mathbf{u}, \mathbf{v})$ is:

$$w(\mathbf{u}, \mathbf{v}) \;:=\; \Pr_{\mathcal{M}, \mathbf{a}}[\Gamma(\mathbf{u}, (\mathcal{M}, \mathbf{a})) = \mathbf{v}],$$

where the probability is over the choice of a random perfect labeled matching $\mathcal{M}$ between $\mathbf{u}^{-1}\{0\}$ and $\mathbf{u}^{-1}\{1\}$, and a uniformly random string $\mathbf{a} \in \{0,1\}^{n/2}$. For every balanced string $\mathbf{u} \in \{0,1\}^n_{n/2}$, we will denote by $W(\mathbf{u})$ the distribution on $\{0,1\}^n_{n/2}$ where the probability of sampling $\mathbf{v}$ is equal to $w(\mathbf{u}, \mathbf{v})$. Let $W \in \mathbb{R}^{n' \times n'}$ denote the weighted adjacency matrix of G, i.e.

$$W[\mathbf{u}, \mathbf{v}] \;=\; w(\mathbf{u}, \mathbf{v}), \qquad \text{for all } \mathbf{u}, \mathbf{v} \in \{0,1\}^n_{n/2}$$

We are now ready to state the main technical lemma of our proof. It roughly says that if we sample a random vertex (which is a random balanced string) and its neighbour in the above-mentioned

graph, then the two balanced strings behave "almost like pairwise-independent" points. In other words, the above-mentioned graph is a good sampler for the balanced slice $\{0,1\}^n_{n/2}$.

Lemma 3.2 (Main Lemma)**.** *There exists a constant $\varepsilon > 0$ for which the following holds. Let G and W be as mentioned above and let $S \subseteq \{0,1\}^n_{n/2}$ be an arbitrary subset of vertices with $|S| \geqslant 4^{-d} \cdot \binom{n}{n/2}$. Let ρ denote the density of the set S. Then,*

$$\Pr_{\substack{\mathbf{x}\sim U_{n,n/2}\\ \mathbf{y}\sim W(\mathbf{x})}}[\mathbf{x} \in S \;\; \textit{and} \;\; \mathbf{y} \in S] \;\leqslant\; \rho^2 \cdot \left(1 + \frac{1}{n^{\varepsilon}}\right).$$

We will give two proofs for Lemma 3.2, for two regimes of the degree d:

1. For degree $d \leqslant C \log n$ for some absolute constant $C > 0$, we give a simple argument using the spectral expansion properties of Cayley graphs and the expander mixing lemma. We prove this in Section 3.1.
2. For degree $d \leqslant n^{\gamma}$ for some absolute constant $\gamma > 0$, we rely on the spectrum of Johnson association schemes and use hypercontractivity for slice functions. We prove this in Section 3.2.

We will also need a lower bound on the probability in Lemma 3.2. This will hold for all degree d. Combining the upper and lower bounds (i.e., Lemma 3.2 and Lemma 3.3 gives the final bound: see Section 3.3.)

Lemma 3.3 (The lower bound)**.** *Let G be the graph as mentioned above and fix a degree parameter $d \in \mathbb{N}$. Let $P(\mathbf{x}) : \{0,1\}^n_{n/2} \to \mathbb{R}$ be a non-zero polynomial on the balanced slice $\{0,1\}^n_{n/2}$ with $\deg(P) \leqslant d$. If $S \subseteq \{0,1\}^n_{n/2}$ denote the set of non-zeroes of $P(\mathbf{x})$, then,*

$$\Pr_{\substack{\mathbf{x}\sim U_{n,n/2}\\ \mathbf{y}\sim W(\mathbf{x})}}[\mathbf{x} \in S \;\; \textit{and} \;\; \mathbf{y} \in S] \;\geqslant\; \frac{|S|}{\binom{n}{n/2}} \cdot \frac{1}{2^d}.$$

Proof of Lemma 3.3. Note that it is sufficient to show that

$$\Pr_{\mathbf{y}\sim W(\mathbf{x})}[\mathbf{y} \in S \,|\, \mathbf{x} \in S] \;\geqslant\; \frac{1}{2^d}, \qquad \text{for all} \;\; \mathbf{x} \in S.$$

Fix an arbitrary point $\mathbf{u} \in S$ and fix an arbitrary matching $\mathcal{M}$ between $\mathbf{u}^{-1}\{0\}$ and $\mathbf{u}^{-1}\{1\}$. We will show that for $1/2^d$-fraction of $\mathbf{a} \in \{0,1\}^{n/2}$, the string $\Gamma(\mathbf{u}, (\mathcal{M}, \mathbf{a})) \in S$.

Define the polynomial $Q(z_1, \ldots, z_{n/2}) := P(\Gamma(\mathbf{u}, (\mathcal{M}, \mathbf{z})))$. Note that $\deg(Q) \leqslant \deg(P) \leqslant d$ and $Q(\mathbf{0}) = P(\mathbf{u}) \neq 0$. Now using the standard ODLSZ lemma (Theorem 2.1) on $Q(\mathbf{z})$, we get,

$$\Pr_{\mathbf{z}\sim\{0,1\}^{n/2}}[Q(\mathbf{z}) \neq 0] \;\geqslant\; \frac{1}{2^d} \quad \Rightarrow \quad \Pr_{\mathbf{a}\sim\{0,1\}^{n/2}}[\Gamma(\mathbf{u}, (\mathcal{M}, \mathbf{a})) \in S] \;\geqslant\; \frac{1}{2^d}$$

Since the above lower bound holds for every matching $\mathcal{M}$ between $\mathbf{u}^{-1}\{1\}$ and $\mathbf{u}^{-1}\{0\}$, we have,

$$\Pr_{\mathcal{M},\mathbf{a}}[\Gamma(\mathbf{u},(\mathcal{M},\mathbf{a})) \in S] \geqslant \frac{1}{2^d}$$

Since the above lower bound holds for arbitrary choice of $\mathbf{u} \in S$, this completes the proof of Lemma 3.3.

■

Next, we observe that the random process mentioned above is "S_n-invariant"[4], i.e. the probabilities do not change even if we simultaneously permute the coordinates of $\mathbf{u}$ and $\mathbf{v}$ (using the same permutation for both of them).

Observation 3.4. *For any $\mathbf{u},\mathbf{v} \in \{0,1\}^n_{n/2}$, the weight $w(\mathbf{u},\mathbf{v})$ depends only on[5] $\Delta(\mathbf{u},\mathbf{v})$. We have,*

$$w(\mathbf{u},\mathbf{v}) = \frac{\Delta!(n/2-\Delta)!}{2^{n/2}\cdot(n/2)!} = \frac{1}{2^{n/2}\cdot\binom{n/2}{\Delta}}, \qquad \text{where } 2\Delta = \Delta(\mathbf{u},\mathbf{v}) \in [0,n]$$

To see the above probability, observe that the $\frac{1}{2^{n/2}}$ factor corresponds to sampling the right $\mathbf{a}$ and the $\frac{\Delta!(n/2-\Delta)!}{(n/2)!}$ factor corresponds to picking a matching $\mathcal{M}$ that results in the output $\mathbf{v}$).

Note that the above observation in particular implies that the weighted adjacency matrix W is a real symmetric matrix, and thus has real eigenvalues. Both of our proofs for Lemma 3.2 will be based on upper bounding the eigenvalues of W.

3.1 Simple Proof Using Cayley Graphs

In this section, we prove a version of Lemma 3.2 using a simple and (mostly) self-contained argument. This version holds for degrees $d \leqslant C\log n$ for some absolute constant C.

Let $1 = \mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_{n'}$ be the eigenvalues of W and let $\mu(W)$ denote the second largest eigenvalue in absolute value, i.e. $\mu(W) := \max(|\mu_2|,|\mu_{n'}|)$. A small value of $\mu(W)$ suggests that the random walk represented by W is "expanding" (see Lemma 2.5). The main lemma of this subsection is the following, which shows that $\mu(W)$ is small, i.e. W is a good expander. In the rest of the subsection, let $m = n/2$.

Lemma 3.5 (W is a good expander)**.** *Let W denote the $n' \times n'$ matrix as described before. Then,*

$$\mu(W) \leqslant \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right).$$

We now prove Lemma 3.5, i.e., we show that W is a good expander. The idea of the proof is to show that one can turn W into a (weighted) Cayley graph by adding additional edges and vertices and deduce that the original graph is an expander by using the expansion of the Cayley graph. In

[4] S_n is the group of permutations on n elements.
[5] Recall that $\Delta(\cdot,\cdot)$ represents the Hamming distance.

particular, we will show that W is an induced subgraph of a Cayley graph and use the interlacing of eigenvalues to prove the expansion.

Proof of Lemma 3.5. For the proof, we will assume that m is even; the odd case is handled similarly. Let $\{0,1\}^n_{odd}$ and $\{0,1\}^n_{even}$ denote the sets of points in $\{0,1\}^n$ that are of odd Hamming weight and even Hamming weight respectively. We will now define the weighted Cayley graph.

Let $V' = \{0,1\}^n_{even}$ (note that $\{0,1\}^n_{n/2} \subseteq V'$ as n is assumed to be even). Note that V' is an Abelian group with addition defined by performing coordinate sums modulo 2 (in particular, we may identify {0,1} with $\mathbb{F}_2$). We shall define a weighted Cayley graph W' over vertices V' by specifying its generators (and their weights) as follows. The set of generators is $\mathcal{S} = \{0,1\}^n_{even}$ and a generator $\mathbf{x} \in \mathcal{S}$ has weight

$$w(\mathbf{x}) \;=\; \frac{1}{2^m \cdot \binom{m}{\Delta}}, \qquad \text{where } \; |\mathbf{x}| = 2\Delta \qquad \text{and} \qquad 0 \leqslant \Delta \leqslant m$$

With the above definition for V' and W', we note that the induced subgraph of W' when restricted to the balanced slice $\{0,1\}^n_{n/2} \subseteq V'$, is identical to W. Hence, by applying the eigenvalue interlacing theorem, we have the following.

Claim 3.6 (Eigenvalue interlacing, see e.g. [HJ91])**.** *Let $\mu'_1 \geqslant \mu'_2 \geqslant \cdots \geqslant \mu'_{|V'|}$ be the eigenvalues of W'. Then $\mu_2 \leqslant \mu'_2$ and $\mu'_{|V'|} \leqslant \mu_{n'}$. Hence, $\mu(W) \leqslant \max(|\mu'_2|, |\mu'_{|V'|}|)$.*

The above claim allows us to bound $\mu(W)$ by bounding the absolute values of the eigenvalues of W' (except the largest). To do this, we will first fix an eigenbasis for W'. The characteristic vectors of the first $(n-1)$ variables forms such an eigenbasis (because if $\mathbf{x} \in V'$, then x_n can be expressed as a $\mathbb{F}_2$-linear combination of $x_1, \ldots, x_{n-1}$). That is, for $A \subseteq [n-1]$, the characteristic vector $\chi_A \in \mathbb{R}^{2^{n-1}}$ is defined as $\chi_A(\mathbf{x}) := (-1)^{\sum_{i \in A} x_i}$ for $\mathbf{x} \in V'$. The corresponding eigenvalue of χ_A is denoted by μ'_A (with a slight abuse of notation), and by Lemma 2.4, is equal to

$$\mu'_A = \sum_{\mathbf{y} \in \mathcal{S}} w(\mathbf{y}) \chi_A(\mathbf{y}).$$

It will be convenient to normalize the weights of the generators in $\mathcal{S}$ to make it a probability distribution. More formally, let $\mathcal{D}$ be the probability distribution over $\mathcal{S}$ where the probability of sampling a point $\mathbf{x} \in \mathcal{S}$ is equal to $w(\mathbf{x}) / \sum_{\mathbf{y} \in \mathcal{S}} w(\mathbf{y})$. Thus, $\mu'_\varnothing = \sum_{\mathbf{y} \in \mathcal{S}} w(\mathbf{y})$ and $\mu'_A / \mu'_\varnothing = \mathbb{E}_{\mathbf{x} \sim \mathcal{D}}[\chi_A(\mathbf{x})]$.

We will now show that $0 < \mu'_\varnothing \leqslant \mathcal{O}(\sqrt{m})$ and $|\mathbb{E}_{\mathbf{x} \sim \mathcal{D}}[\chi_A(\mathbf{x})]| = |\mu'_A / \mu'_\varnothing| \leqslant O\big(\frac{\log m}{m}\big)$ for all *non-empty* $A \subseteq [2m-1]$. This would in turn give that $\mu(W) \leqslant \max_{A \neq \varnothing}(|\mu'_A|) \leqslant \mathcal{O}\left(\frac{\sqrt{m} \log m}{m}\right) = \mathcal{O}(\log m / \sqrt{m})$, finishing the proof of Lemma 3.5.

The proof of Claim 3.7 follows from a simple counting argument and can be found in Appendix A.1.

Claim 3.7. $0 < \mu'_\varnothing \leqslant \mathcal{O}(\sqrt{m})$.

It remains to show that $|\mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[\chi_A(\mathbf{x})]| \leqslant \mathcal{O}\big(\frac{\log m}{m}\big)$ for all *non-empty* $A \subseteq [2m-1]$. Note that the distribution $\mathcal{D}$ has some symmetry in the sense that all the points of a given Hamming weight have the same probability. Furthermore, two given points, one of weight 2Δ and the other of weight $(2m-2\Delta)$ also have the same probability mass (for arbitrary $0 \leqslant \Delta \leqslant m$). This leads to $\mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[\chi_A(\mathbf{x})]$ being equal to 0 if $|A| = 1$ or $2m-1$, and hence it suffices to focus on the regime $2 \leqslant |A| \leqslant 2m-2$.

We show the following concentration inequality for the distribution $\mathcal{D}$. The proof of Claim 3.8 can be found in Appendix A.1.

Claim 3.8. $\Pr_{\mathbf{x}\sim\mathcal{D}}[||\mathbf{x}| - m| > \sqrt{50m\log m}] \leqslant \mathcal{O}(1/m^2)$.

Assuming Claim 3.8, it suffices to show that $\mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[\chi_A(\mathbf{x}) \;\mid\; |\mathbf{x}| \in m \pm \sqrt{50m\log m}] \leqslant \mathcal{O}\big(\frac{\log m}{m}\big)$ to conclude that $\mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[\chi_A(\mathbf{x})] \leqslant \mathcal{O}\big(\frac{\log m}{m}\big)$. We will show that this holds even conditioned on $|\mathbf{x}| = 2\Delta$ for every $\Delta \in (m \pm \sqrt{50m\log m})/2$. However, recall that $\mathcal{D}$ is uniform when restricted to $\{0,1\}^{2m}_{2\Delta}$. Therefore, we can equivalently upper bound the quantity $|\mathbb{E}_{\mathbf{x}\in\{0,1\}^{2m}_{2\Delta}}[\chi_A(\mathbf{x})]|$ to conclude the proof. Now, we note that $\mathbb{E}_{\mathbf{x}\in\{0,1\}^{2m}_{2\Delta}}[\chi_A(\mathbf{x})] = \mathbb{E}_{B\sim\binom{[2m]}{|A|}}[\chi_B(\mathbf{c})]$, where $\mathbf{c}$ is an arbitrary point in $\{0,1\}^{2m}_{2\Delta}$ (we will fix it to be $0^{2m-2\Delta}1^{2\Delta}$). Hence, it suffices to show that

$$\left|\mathop{\mathbb{E}}_{B\sim\binom{[2m]}{k}}[\chi_B(\mathbf{c})]\right| \;\leqslant\; \mathcal{O}\left(\frac{\log m}{m}\right), \tag{1}$$

for every $2 \leqslant k \leqslant (2m-2)$ (since we assumed that $2 \leqslant |A| \leqslant 2m-2$). We may further assume that $k \leqslant m$ without loss of generality, as $\chi_B(\mathbf{c}) = \chi_{\overline{B}}(\mathbf{c})$.

To help with the analysis, we will choose $B \sim \binom{[2m]}{k}$ by first choosing a subset C of $[2m]$ of size $(k-2)$ (which is non-negative) uniformly at random and then choosing two elements $b_1 \neq b_2$ from $\overline{C} = [2m]\backslash C$ uniformly at random.

For a subset $C \subseteq [2m]$, we will use the notation $\mathrm{wt}(C)$ to denote the number of 1's in $\mathbf{c}$ when restricted to the coordinates indexed by C. Let $p := \mathrm{wt}([2m])/(2m)$ denote the fractional Hamming weight of $\mathbf{c}$. We say that a subset $C \subseteq [2m]$ is *good*, if $\big||\overline{C}| - 2\mathrm{wt}(\overline{C})\big| \leqslant \sqrt{2000\log m}$. We claim that $\Pr_{C\sim\binom{[2m]}{k-2}}[C \text{ is not good}] \leqslant \mathcal{O}(1/m^2)$. This essentially follows from standard tail bounds for the hypergeometric distribution but needs some care to handle small k. We divide this analysis into two cases.

- **Case 1:** $k \leqslant \sqrt{50m\log m}$**.** We note that $|\overline{C}| = 2m-k+2 \in 2m \pm \sqrt{50m\log m}$, and similarly $\mathrm{wt}(\overline{C}) \in \mathrm{wt}([2m]) \pm \sqrt{50m\log m} \subseteq m \pm 2\sqrt{50m\log m}$. Using these bounds, it follows that $\big||\overline{C}| - 2\mathrm{wt}(\overline{C})\big| \leqslant \sqrt{2000\log m}$ for sufficiently large m (for every choice of $C \in \binom{[2m]}{k-2}$). Hence all choices of C are good in this case.
- **Case 2:** $k > \sqrt{50m\log m}$**.** We note that $\mathrm{wt}(C)$ is distributed according to a hypergeometric distribution — it corresponds to the number of successes in $k-2$ draws with replacement from a population of $2m$ total states and $\mathrm{wt}([2m]) = 2d$ success states. Using a standard tail bound [Hoe94], we obtain that $\Pr_{C\sim\binom{[2m]}{k-2}}\big[\big|\frac{\mathrm{wt}(C)}{k-2} - p\big| > \sqrt{\frac{4\log k}{k}}\big] = \mathcal{O}(1/k^4) = \mathcal{O}(1/m^2)$. Using $\big|p - \frac{1}{2}\big| \leqslant \sqrt{\frac{50\log m}{m}}$, we thus get that $||C| - 2\mathrm{wt}(C)| \leqslant 4\sqrt{50m\log m}$ with probability

at least $1 - \mathcal{O}(1/m^2)$. Because $2m = |C| + \overline{C}$ and $2d = \mathrm{wt}(C) + \mathrm{wt}(\overline{C})$, with probability at least $1 - \mathcal{O}(1/m^2)$, we have $||\overline{C}| - 2\mathrm{wt}(\overline{C})| \leqslant 6\sqrt{50m \log m}$, i.e., C is good.

We now show that conditioned on C being good, the expectation of $\chi_B(\mathbf{c})$ is upper bounded by $\mathcal{O}\left(\frac{\log m}{m}\right)$ in absolute value. This would then prove (1). For ease of notation, let $n_0 = |\overline{C}| = 2m - k + 2 \geqslant m$ and $d_0 = \mathrm{wt}(\overline{C}) \in \frac{n_0}{2} \pm \Theta(\sqrt{n_0 \log n_0})$ (since C is good). We now note that $\chi_B(\mathbf{c}) = \chi_C(\mathbf{c}) \cdot \chi_{\{b_1,b_2\}}(\mathbf{c})$, so it suffices to bound $|\mathbb{E}_{b_1,b_2}[(-1)^{c_{b_1}+c_{b_2}}]|$. The idea now is that c_{b_1} and c_{b_2} almost behave like two independent draws, so the expectation is *roughly* the square of $|\mathbb{E}_b[(-1)^{c_b}]|$ for a uniformly random coordinate $b \in \overline{C}$, which is equal to $\left|\frac{n_0 - 2d_0}{n_0}\right| \leqslant O(\sqrt{\frac{\log n_0}{n_0}}) \leqslant O(\sqrt{\frac{\log m}{m}})$ as $n_0 \geqslant m$. More precisely, we have the following:

$$\begin{aligned}
|\mathop{\mathbb{E}}_{b_1,b_2}[(-1)^{c_{b_1}+c_{b_2}}]| &= \frac{|\binom{d_0}{2} + \binom{n_0-d_0}{2} - d_0(n_0-d_0)|}{\binom{n_0}{2}} \\
&= \frac{|(n_0 - 2d_0)^2 - n_0|}{n_0(n_0-1)} \\
&\leqslant \mathcal{O}\left(\frac{\log n_0}{n_0}\right) \leqslant \mathcal{O}\left(\frac{\log n}{n}\right). \qquad \text{(as } d_0 \in n_0/2 \pm \Theta(\sqrt{n_0 \log n_0})\text{)}
\end{aligned}$$

This finishes the proof of Lemma 3.5. ■

Proof of Lemma 3.2 for $d \leqslant C \log n$. We use Lemma 3.5 to prove an upper bound for Lemma 3.2 that holds for all $d \leqslant C \log n$ for an absolute constant $C > 0$. Using the expander mixing lemma (Lemma 2.5), we obtain

$$\Pr_{\mathbf{x} \sim U_{n,n/2}, \mathbf{y} \sim N(\mathbf{x})}[\mathbf{x} \in S \text{ and } \mathbf{y} \in S] \leqslant \rho^2 + \rho \cdot \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right),$$

where S denotes the non-zeroes of P in $\{0,1\}^n_{n/2}$ and $\rho = |S|/\binom{n}{n/2}$. Thus assuming $d \leqslant C \log n$ for small enough constant C and using $\rho \geqslant 4^{-d}$ (Lemma 1.2), we get that the above probability is at most $\rho^2(1 + 1/n^{\varepsilon})$ for sufficiently small constant ε. Hence, this finishes the proof of Lemma 3.2 in the regime $d \leqslant C \log n$.

■

3.2 Proof via Johnson Association Schemes

In this section, we prove a stronger version of Lemma 3.2, i.e. we will give a tighter upper bound. This will hold for all $d \leqslant n^{\gamma}$ for some absolute constant $\gamma > 0$. In this subsection, we will always assume that n is divisible by 2. We will first give some preliminaries on *Johnson association schemes* and functions on the balanced slice.

3.2.1 Preliminaries for Johnson Association Schemes

We now discuss that the matrix W has some useful properties. Recall $n' = \binom{n}{n/2}$. Consider the set of $n' \times n'$ dimensional matrices satisfying the following property: For every entry $(\mathbf{u}, \mathbf{v})$, the entry

only depends on the Hamming distance $\Delta(\mathbf{u}, \mathbf{v})$ as mentioned in Observation 3.4. The set of all such matrices forms an algebra known as *Bose-Mesner algebra of the* $(n, n/2)$ *Johnson association scheme*. Note that the matrices in this algebra are invariant under the action of the symmetric group S_n on the coordinates, i.e., if we apply a permutation $\pi \in S_n$ on the coordinates of the rows and columns simultaneously, the matrix remains invariant. Filmus [Fil16] showed that there exist vector spaces that form orthogonal eigenspaces for matrices in Bose-Mesner algebra of the $(n, n/2)$ Johnson association scheme with certain useful properties.

We equip the space of functions on the balanced slice $\{0,1\}^n_{n/2}$ with an inner product. We consider the following inner product on functions on the balanced slice: For any two functions $f, g : \{0,1\}^n_{n/2} \to \mathbb{R}$

$$\langle f, g \rangle \;=\; \mathbb{E}_{\mathbf{x} \sim U_{n,n/2}}[f(\mathbf{x}) \cdot g(\mathbf{x})]$$

The p^{th} norm of a function $f : \{0,1\}^n_{n/2} \to \mathbb{R}$ is defined as follows:

$$\|f\|_p \;=\; \mathbb{E}_{\mathbf{x} \sim U_{n,n/2}}[|f(\mathbf{x})|^p]^{1/p} \;=\; \frac{1}{n'} \left(\sum_{\mathbf{x} \in \{0,1\}^n_{n/2}} |f(\mathbf{x})|^p \right)^{1/p}$$

Let us define a particular function on the balanced slice $\{0,1\}^n_{n/2}$. For every $t \in \{0, 1, \ldots, n/2\}$, let $f_t(\mathbf{x}) : \{0,1\}^n_{n/2} \to \mathbb{R}$ be the following function: For $t = 0$, $f_t = 1$ and for $t \geqslant 1$,

$$f_t(x_1, \ldots, x_n) \;=\; (x_1 - x_2) \cdot (x_3 - x_4) \cdots (x_{2t-1} - x_{2t}) \tag{2}$$

We are going to interpret $f_t(\mathbf{x})$ as a vector, i.e. for every $\boldsymbol{\alpha} \in \{0,1\}^n_{n/2}$, the $\boldsymbol{\alpha}^{th}$ entry of the vector is $f_t(\boldsymbol{\alpha})$. We will use the following result from [Fil16].

Lemma 3.9. *[Fil16, Lemma 18]. There exists orthogonal*[6] *vector spaces* $\mathcal{V}_{n,0}, \ldots, \mathcal{V}_{n,n/2}$ *for which the following holds. Let* W *be any matrix in the Bose-Mesner algebra for the* $(n, n/2)$ *Johnson association scheme. Then,*

1. *The spaces* $\mathcal{V}_{n,0}, \ldots, \mathcal{V}_{n,n/2}$ *are orthogonal eigenspaces for the matrix* W *with corresponding eigenvalues (not necessarily distinct)* $\lambda_0, \lambda_1, \ldots, \lambda_{n/2}$.
2. *For every integer* $t \in \{0, 1, \ldots, n/2\}$, *the function* $f_t(\mathbf{x})$, *as described in Equation (2) lies in* $\mathcal{V}_{n,t}$. *In other words,* f_t *is an eigenvector of* W *in the* t^{th} *eigenspace with eigenvalue* λ_t.

For a function $f : \{0,1\}^n_{n/2} \to \mathbb{R}$ and $0 \leqslant t \leqslant n/2$, let $f^{=t}$ denote the component of f in the t^{th} eigenspace $\mathcal{V}_{n,t}$ (see Lemma 3.9). We now define the following noise operator for functions on the slice.

[6] The orthogonality is with respect to the inner product defined in the previous paragraph.

Definition 3.10 (Noise operator for functions on slice)**.** *(See [FI19, Section 2]). For a parameter* $\rho \in (0,1]$*, the* noise operator T_ρ *maps functions on the slice* $\{0,1\}^n_{n/2}$ *to functions on the slice* $\{0,1\}^n_{n/2}$*, defined as:*

$$T_\rho f \;=\; \sum_{t=0}^{n/2} \rho^{t\left(1-\frac{t-1}{n}\right)}\, f^{=t}.$$

Now we state a hypercontractive inequality for the noise operator T_ρ. Lee and Yau [LY98] proved a log-Sobolev inequality which combined with a result of Diaconnis and Saloff-Coste [DS96] implies the following hypercontractive inequality. The following lemma is a simplified form of the more general inequality. For more details, refer to [FI19, Section 2] and [Fil16, Section 2].

Lemma 3.11 (Hypercontractive inequality using log-Sobolev inequality)**.** *Fix any multilinear polynomial* f *on the balanced slice* $\{0,1\}^n_{n/2}$*. Then there exists a constant* c *such that for every* $1 \leqslant p \leqslant q \leqslant \infty$ *with* $\frac{q-1}{p-1} \leqslant \exp\left(c \log 1/\rho\right)$*,*

$$\|T_\rho f\|_q \;\leqslant\; \|f\|_p.$$

Remark 3.12. *In [Fil16] and [FI19], the above lemma is stated in a slightly different way, so we take a moment to clarify that here. As mentioned in [FI19, Page 2] (see the line after Equation (2)), the noise operator* T_ρ *is equivalent to the expected value after applying random* $\mathrm{Po}(\frac{n-1}{2}\log(1/\rho))$ *transpositions. Using the comment in [Fil16] (after Definition 26),* T_ρ *is equivalent to* $H_{g(\rho)}$*, where* $H_\cdot$ *is the Heat operator and* $g(\rho) = \frac{n-1}{2}\log(1/\rho)$*. We get Lemma 3.11 by using [Fil16, Lemma 27] (the reader should be careful that the* ρ *in that lemma is the* log Sobolev constant *and is different from the* ρ *in Lemma 3.11).*

3.2.2 Proof of the Sampling Guarantee

One of the key steps in our proof of Lemma 3.2 will be Lemma 3.13 which gives an upper bound on the eigenvalues of the matrix W. It says that for small t, the eigenvalue λ_t is quite small (roughly $1/n^t$) and for larger t, the eigenvalue is still exponentially small (roughly $1/2^t$). To upper bound the eigenvalues, the idea is to choose a suitable eigenvector and argue about its non-zero coordinates. In particular, we will work with the eigenvector $f_t(\mathbf{x})$ as stated in Equation (2) and Lemma 3.9.

Lemma 3.13. *Let* $\lambda_0, \lambda_1, \ldots, \lambda_{n/2}$ *denote the eigenvalues of* W *and let* $\tau := n^\delta$ *for a sufficiently small constant* $\delta > 0$*. Then,*

- *For* $1 \leqslant t \leqslant \tau$*,* $\lambda_t \leqslant 1/n^{\Omega(t)}$
- *For* $t > \tau$*,* $\lambda_t \leqslant 1/2^{n^{\Omega(1)}}$*.*

Before going into the proof of Lemma 3.13, we will define a property on bipartite matching. For a matching $\mathcal{M}$, if $(i,j) \in \mathcal{M}$, then we will use the notation $\mathcal{M}(i) = j$ and $\mathcal{M}(j) = i$.

Definition 3.14 (Good and self-good matching)**.** *Consider a complete bipartite graph $K_{n/2,n/2}$ on vertex set $(L \bigcup R)$, where $L = \{1, 3, \ldots, n-1\}$ and $R = \{2, 4, \ldots, n\}$. We will refer to a matching $\mathcal{M}$ between L and R as* t-good *if the following holds:*

$$\mathcal{M}(2i-1) \in \{2, 4, \ldots, 2t\} \quad \text{or} \quad \mathcal{M}(2i) \in \{1, 3, \ldots, 2t-1\}, \qquad \text{for all } i \in [t]$$

We will call a good matching $\mathcal{M}$ a t-self good *matching if for every subset $T \subseteq [t]$ of size $t/2$, there exists $i \in T$ such that*

$$\mathcal{M}(2i-1) = 2i$$

We will simply refer to matchings that are not t-good as *t-bad* matchings. We will refer to matchings that are t-good but not t-self good as *t non-self good* matchings. In our proof of Lemma 3.13, it will be useful to have an upper bound on the probability that a random matching $\mathcal{M}$ is a t-good matching or a t-self good matching. We upper bound these probabilities in the following claim. The proof is a straightforward counting argument with standard binomial estimations. We omit the proof here and it can be found in Appendix A.2.

Claim 3.15 (Upper bound on probability of (self) good matchings)**.** *Consider the complete bipartite graph $K_{n/2,n/2}$ on $L \bigcup R$ where $L = \{1, 3, \ldots, n-1\}$ and $R = \{2, 4, \ldots, n\}$. Let $\tau = n^{\delta}$ for a sufficiently small $\delta > 0$. Then,*

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a } t\text{-good matching}] \leqslant \frac{1}{n^{\Omega(t)}}, \qquad \text{for all } t \leqslant \tau,$$

where the above probability is over the choice of a uniformly random matching $\mathcal{M}$. Also,

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a } t\text{-self good matching}] \leqslant \frac{1}{n^{\Omega(t)}}, \qquad \text{for all } t > \tau,$$

where the above probability is over the choice of a uniformly random matching $\mathcal{M}$.

Now we have all the essentials with us to prove the Lemma 3.13. As mentioned earlier, the idea would be to fix a non-zero coordinate of the eigenvector $f_t(\mathbf{x})$ and consider that coordinate in Wf_t.

Proof of Lemma 3.13. Recall that the rows and columns of the matrix W are indexed by points of $\{0,1\}^n_{n/2}$. Fix any particular $t \in \{1, \ldots, n/2\}$. We know that f_t is an eigenvector of the matrix M with eigenvalue λ_t, i.e. for every $\mathbf{u} \in \{0,1\}^n_{n/2}$, we have,

$$(Wf_t)[\mathbf{u}] \;=\; \lambda_t \cdot f_t(\mathbf{u}) \quad \Rightarrow \quad \mathbb{E}_{\mathbf{v} \sim W(\mathbf{u})}[f_t(\mathbf{v})] \;=\; \lambda_t \cdot f_t(\mathbf{u})$$

Fix any $\mathbf{u}$ for which $f_t(\mathbf{u}) = 1$. For convenience, let $\mathbf{u} = 1010\ldots10$. Then we have,

$$\lambda_t \;=\; \mathbb{E}_{\mathbf{v} \sim W(\mathbf{u})}[f_t(\mathbf{v})]$$

We make the following observation about bad matchings.

Observation 3.16. *Fix $\mathcal{M}$ to be any t-bad matching. For simplicity in notation, assume that $\mathcal{M}(1) \notin \{2, 4, \ldots, 2t\}$ and $\mathcal{M}(2) \notin \{1, 3, \ldots, 2t-1\}$. Let $\mathcal{M}(2) = (2j-1)$ for some $j > t$. Then we have $v_1 = u_1 \oplus a_1$ and $v_2 = u_2 \oplus a_j$. Since a_1 and a_j are mutually independent, this implies that v_1 and v_2 are mutually independent too. This implies that the expected value of $(v_1 - v_2)$ over the random choice of $\mathbf{a}$ is 0 (conditioned on a bad matching $\mathcal{M}$). Using the independence of bits in $\mathbf{a}$, we have,*

$$\mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-bad}] \;=\; \mathbb{E}_{(a_1,a_j)}\left[(v_1 - v_2)\right] \cdot \mathbb{E}_{a_2,\ldots,a_{n/2}\setminus a_j}\left[\prod_{k>1}^{t}(v_{2k-1} - v_{2k})\right]$$
$$\Rightarrow \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-bad}] \;=\; 0$$

This gives us the following:

$$\lambda_t \;=\; \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-good}] \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-good}] \;+\; \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-bad}] \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-bad}]$$

From Observation 3.16, we know that the expected value of $f_t(\mathbf{v})$ for $\mathbf{v} \in N(\mathbf{u})$ conditioned on a bad matching $\mathcal{M}$ is 0. This gives us

$$\lambda_t \;=\; \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_d(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-good}] \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-good}]$$

Case 1 - $t \leqslant \tau$: Note that the expectation is over a random choice of $\mathbf{a} \sim \{0,1\}^{n/2}$. For any $\mathbf{v} \in \{0,1\}^n_{n/2}$, $f_t(\mathbf{v}) \in \{-1, 0, 1\}$. In other words, the absolute value of the expectation is at most 1. Using this, we now have,

$$\lambda_t \;\leqslant\; |\mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-good}]| \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-good}] \;\leqslant\; \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-good}]$$

For $t \leqslant \tau$, Claim 3.15 implies that $\lambda_t \leqslant 1/n^{\Omega(t)}$. This shows the first item of Lemma 3.13.

Case 2 - $t > \tau$: We make the following observation about non-self good matchings.

Observation 3.17. *Fix a t non-self good matching $\mathcal{M}$. By definition of t-self good matchings, we know that there exists a set $T \subseteq [t]$ of size $t/2$ such that for every $i \in T$, $\mathcal{M}(2i-1) \neq 2i$. Assume without loss of generality that $T = [t/2]$ and $\mathcal{M}(2i) = (2j_i - 1)$ for $i \in T$. Note that for every $i \in T$, if $a_i \neq a_{j_i}$, then $f_t(\mathbf{x}) = 0$. In other words, for f_t to be non-zero, it is necessary that for every $i \in [t/2]$, $a_i = a_{j_i}$. This implies that f_t is non-zero for at most $\frac{1}{2^{t/4}}$ choices of $\mathbf{a}$.*

Using Observation 3.17, we get the following upper bound on the expected value conditioned on a t non-self good matching $\mathcal{M}$:

$$\mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t \text{ non-self good}] \;\leqslant\; |\mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t \text{ non-self good}]| \;\leqslant\; \frac{1}{2^{t/4}}$$

Thus finally we have,

$$\lambda_t \;=\; \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-good}] \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-good}]$$

$$= \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t\text{-self good}] \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-self good}]$$
$$+ \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t \text{ non-self good}] \cdot \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t \text{ non-self good}]$$

$$\leqslant \Pr_{\mathcal{M}}[\mathcal{M} \text{ is } t\text{-self good}] + \mathbb{E}_{\mathbf{v}\sim W(\mathbf{u})}[f_t(\mathbf{v}) \,|\, \mathcal{M} \text{ is } t \text{ non-self good}]$$
$$\Rightarrow \lambda_t \leqslant \frac{1}{2^{t/4}} + \frac{1}{n^{\Omega(t)}} \leqslant \frac{1}{2^{n^{\Omega(1)}}}$$

This shows the second item of Lemma 3.13 and completes the proof of Lemma 3.13.

■

We are now ready to prove our main lemma (Lemma 3.2) in the setting when $d \leqslant n^{\kappa}$. We recall the statement below.

Lemma 3.2 (Main Lemma)**.** *There exists a constant $\varepsilon > 0$ for which the following holds. Let G and W be as mentioned above and let $S \subseteq \{0,1\}^n_{n/2}$ be an arbitrary subset of vertices with $|S| \geqslant 4^{-d} \cdot \binom{n}{n/2}$. Let ρ denote the density of the set S. Then,*

$$\Pr_{\substack{\mathbf{x}\sim U_{n,n/2}\\ \mathbf{y}\sim W(\mathbf{x})}}[\mathbf{x} \in S \;\; \textit{and} \;\; \mathbf{y} \in S] \;\leqslant\; \rho^2 \cdot \left(1 + \frac{1}{n^{\varepsilon}}\right).$$

Proof of Lemma 3.2 for $d \leqslant n^{\gamma}$. Let $\mathbb{1}_S$ denote the n'-dimensional characteristic vector for the subset S. Then,

$$\Pr_{\substack{\mathbf{x}\sim U_{n,n/2}\\ \mathbf{y}\sim W(\mathbf{x})}}[\mathbf{x} \in S \;\text{ and }\; \mathbf{y} \in S] \;=\; \langle \mathbb{1}_S,\, W\mathbb{1}_S \rangle,$$

where $\langle f, g\rangle = \mathbb{E}_{\mathbf{x}\sim U_{n,n/2}}[f(\mathbf{x})g(\mathbf{x})]$. Let $\mathcal{V}_{n,0}, \ldots, \mathcal{V}_{n,n/2}$ be the orthogonal basis for the space of functions on $\{0,1\}^n_{n/2}$ as stated in Lemma 3.9. Let $\mathbb{1}_S^{=t}$ denote the component of $\mathbb{1}_S$ in the t^{th} eigenspace $\mathcal{V}_{n,t}$. We have,

$$W\mathbb{1}_S \;=\; \sum_{t=0}^{n/2+1} W\mathbb{1}_S^{=t} \;=\; \sum_{t=0}^{n/2+1} \lambda_t \mathbb{1}_S^{=t},$$

where for the final equality we used the fact that $\mathbb{1}_S^{=t}$ is an eigenvector for W (first item of Lemma 3.9). Using this, we have,

$$\langle \mathbb{1}_S,\, W\mathbb{1}_S \rangle \;=\; \left\langle \sum_{t=0}^{n/2+1} \mathbb{1}_S^{=t},\; \sum_{t=0}^{n/2+1} \lambda_t \mathbb{1}_S^{=t} \right\rangle \;=\; \sum_{t=0}^{n/2+1} \lambda_t \|\mathbb{1}_S^{=t}\|_2^2,$$

where for the final equality we used the orthogonality of $\mathcal{V}_{n,t}$'s.

Using the orthogonality $\mathcal{V}_{n,t}$'s, we have,

$$\|T_\rho \mathbb{1}_S\|_2^2 \;=\; \sum_{t=0}^{n/2+1} \rho^{2t\left(1-\frac{t-1}{n}\right)} \|\mathbb{1}_S^{=t}\|_2^2$$

Let $\tau := n^\delta$ for small enough $\delta > 0$ such that Lemma 3.13 holds. If we have $\rho = 1/n^{\delta'}$ for a small enough δ' depending on δ, then for $t \leqslant \tau$, we have,

$$\lambda_t \;\leqslant\; \rho^{2t} \;\leqslant\; \rho^{2t\left(1-\frac{t-1}{n}\right)}$$

This implies that for $t \leqslant \tau$, we have,

$$\sum_{t=0}^{\tau} \lambda_t \|\mathbb{1}_S^{=t}\|_2^2 \;\leqslant\; \sum_{t=0}^{\tau} \rho^{2t\left(1-\frac{t-1}{n}\right)} \|\mathbb{1}_S^{=t}\|_2^2 \;\leqslant\; \|T_\rho \mathbb{1}_S\|_2^2 \tag{3}$$

For $t > \tau$, we have,

$$\sum_{t=\tau+1}^{n/2+1} \lambda_t \|\mathbb{1}_S^{=t}\|_2^2 \;\leqslant\; \sum_{t=\tau}^{n/2+1} \frac{1}{2^{n^{\Omega(1)}}} \|\mathbb{1}_S^{=t}\|_2^2 \;\leqslant\; n \cdot \frac{1}{2^{n^{\Omega(1)}}}, \tag{4}$$

where we upper bounded $\|\mathbb{1}_S^{=t}\|_2^2$ by 1. Combining these two together, we get,

$$\sum_{t=0}^{n/2+1} \lambda_t \|\mathbb{1}_S^{=t}\|_2^2 \;\leqslant\; \|T_\rho \mathbb{1}_S\|_2^2 \;+\; \frac{n}{2^{n^{\Omega(1)}}}$$

Now applying the hypercontractivity theorem for the noise operator T_ρ, we get,

$$\|T_\rho \mathbb{1}_S\|_2 \;\leqslant\; \|\mathbb{1}_S\|_p,$$

where $1/(p-1) \leqslant \exp(c \log 1/\rho) = \exp(c\delta' \log n)$. We also have that for any p, the norm $\|\mathbb{1}_S\|_p$ is equal to $(|S|/n')^{1/p}$. Using this, we get,

$$\|T_\rho \mathbb{1}_S\|_2^2 \;\leqslant\; \left(\frac{|S|}{n'}\right)^{2/p} \;=\; \left(\frac{|S|}{n'}\right)^{2(1-1/n^\kappa)},$$

for some constant κ depending on the constant c from the hypercontractive inequality Lemma 3.11 and ρ. Plugging this back in, we have,

$$\begin{aligned}\Pr_{\substack{\mathbf{x}\sim U_{n,n/2}\\ \mathbf{y}\sim W(\mathbf{x})}}[\mathbf{x}\in S \;\text{ and }\; \mathbf{y}\in S] \;&=\; \langle \mathbb{1}_S,\, W\mathbb{1}_S\rangle = \sum_{t=0}^{n/2+1} \lambda_t \|\mathbb{1}_S^{=t}\|_2^2 \\ \leqslant\; \left(\frac{|S|}{n'}\right)^{2(1-1/n^\kappa)} \;+\; \frac{1}{2^{n^{\Omega(1)}}} &= \left(\frac{|S|}{n'}\right)^2 \cdot \left(\left(\frac{|S|}{n'}\right)^{-2/n^\kappa} + \frac{(n'/|S|)^2}{2^{n^{\Omega(1)}}}\right).\end{aligned}$$

From the hypothesis of Lemma 3.2, we know that $|S|/n' \geqslant 4^{-d}$. Using this in the parenthetical term above, we get

$$\Pr_{\substack{\mathbf{x}\sim U_{n,n/2}\\ \mathbf{y}\sim W(\mathbf{x})}}[\mathbf{x}\in S \;\text{ and }\; \mathbf{y}\in S] \;\leqslant\; \left(\frac{|S|}{n'}\right)^2 \cdot \left(4^{\mathcal{O}(d/n^\kappa)} + \frac{4^{\mathcal{O}(d)}}{2^{n^{\Omega(1)}}}\right) \;\leqslant\; \left(\frac{|S|}{n'}\right)^2 \cdot \left(1 + \frac{1}{n^\varepsilon}\right).$$

for a small enough absolute constant $\varepsilon > 0$ as long as $d \leqslant n^\gamma$ for a small enough absolute constant $\gamma > 0$. This concludes the proof of Lemma 3.2. ∎

3.3 Putting Everything Together

We now use the above bounds to show that over the balanced slice (the set of points with Hamming weight $n/2$), we have the optimal distance lemma for low-degree polynomials. The main result of this section is the following lemma.

Theorem 3.18 (Distance lemma over the balanced slice)**.** *There exists an absolute constant $\varepsilon > 0$ so that the following holds. Fix an arbitrary Abelian group G and fix a degree parameter $d \in \mathbb{N}$ where $d \leqslant n^{\varepsilon}$. For every even natural number n, and for every non-zero degree-d polynomial $P(\mathbf{x}) \in \mathcal{P}_d(n, n/2, G)$,*

$$\Pr_{\mathbf{x} \sim U_{n,n/2}}[P(\mathbf{x}) \neq 0] \geqslant \frac{1}{2^d} \cdot \left(1 - \frac{1}{n^{\Omega(1)}}\right)$$

Proof of Theorem 3.18. Letting $S \subseteq \{0,1\}^n_{n/2}$ denote the set of points on the balanced slice on which P evaluates to a non-zero value. From Lemma 1.2, the set S satisfies the density lower required in Lemma 3.2. Combining Lemma 3.2 and Lemma 3.3, we obtain

$$\frac{|S|}{\binom{n}{n/2}} \cdot \frac{1}{2^d} \leqslant \Pr_{\substack{\mathbf{x} \sim U_{n,n/2} \\ \mathbf{y} \sim W(\mathbf{x})}}[\mathbf{x} \in S \text{ and } \mathbf{y} \in S] \leqslant \left(\frac{|S|}{\binom{n}{n/2}}\right)^2 \cdot \left(1 + \frac{1}{n^{\varepsilon}}\right),$$

where ε is a sufficiently small constant. Hence, $|S|/\binom{n}{n/2} \geqslant \frac{1}{2^d} \cdot \left(1 - \frac{1}{n^{\Omega(1)}}\right)$. ∎

4 Arbitrary Slices

In this section we prove a distance lemma for slices over arbitrary Abelian groups G. As discussed in the proof overview, the proof has three key steps:

1. First, we prove a distance lemma over cyclic groups of prime power order for some fixed set of slices, which we refer to as "good" slices. We prove this in Lemma 4.6.
2. Secondly, we prove a distance lemma over cyclic groups of prime power order for any slice by reducing it to one of the good slices. We prove this in Lemma 4.8.
3. In the end, we show that any Abelian group can be assumed to be a finitely generated Abelian group. To prove the distance lemma for a finitely generated Abelian group, it suffices to have the distance lemma over cyclic groups of prime power order, which we prove in Lemma 4.1.

We will start by proving a distance lemma for slices over cyclic groups of prime power order.

4.1 Cyclic Groups of Prime Power Order

In this subsection we will prove the distance lemma for slices over $\mathbb{Z}_{p^\ell}$ for some prime p and natural number ℓ. The main result of this subsection is the following lemma.

Lemma 4.1 (Distance lemma for cyclic groups)**.** *There exists an absolute constant $\varepsilon > 0$ so that the following holds. Fix a cyclic group $\mathbb{Z}_q$ where $q = p^\ell$ for some prime p and a degree parameter $d \in \mathbb{N}$. For all natural numbers n and k such that $1 + d \leqslant k^\varepsilon$ and $k \leqslant n/2$, the following is true.*
For a non-zero degree-d polynomial $P : \{0,1\}_k^n \to \mathbb{Z}_q$,

$$\Pr_{\mathbf{x} \sim U_{n,k}} [P(\mathbf{x}) \neq 0] \;\geqslant\; \alpha^d \left(1 - \frac{1}{k^{\Omega(1)}}\right), \qquad \text{where } \alpha := \frac{k}{n}.$$

We will start with the definition of good slices. The idea is that these slices admit a nice basis which can be used to reduce the problem from a good slice to the balanced slice over a smaller dimensional cube, which we have proved in Theorem 3.18.

Definition 4.2 (Good slices)**.** *Fix a prime p and a degree parameter $d \in \mathbb{N}$. An integer $k \geqslant d$ is said to be (d,p)-good if the p-ary expansion of k agrees with that of d in all the digits up to the leading digit of d, and is greater than equal to d in the leading digit. More formally, if $k = \sum_{j=0}^{m} a_j p^j$ and $d = \sum_{j=0}^{\ell} b_j p^j$ with $a_j, b_j \in \{0, \ldots, p-1\}$, with $b_\ell > 0$, then $a_j = b_j$ for all $j < \ell$ and $a_\ell \geqslant b_\ell$.*

For a degree parameter d, let $\mathcal{H}_d$ denote the set of homogeneous monomials in $\{x_1, \ldots, x_n\}$, i.e.

$$\mathcal{H}_d \;=\; \left\{ \prod_{i \in T} x_i \;\middle|\; T \subseteq [n], |T| = d \right\}$$

Since every $x_i \in \{0,1\}$, we only work with multilinear monomials. So for convenience, we will identify monomials with sets and vice-versa and for a set $T \subseteq [n]$, let $x^T := \prod_{i \in T} x_i$.

We will now show that if k is (d,p)-good, then the set of degree-d homogeneous multilinear monomials $\mathcal{H}_d$ form a 'basis' for degree-d polynomials on the k^{th} slice in the following sense.

Lemma 4.3. *Let q be a power of prime p as above. Fix n, k, d such that $d \leqslant k \leqslant n - d$ and assume that k is (d,p)-good. Then, any function in $\mathcal{P}_d(n, k, \mathbb{Z}_q)$ can be written uniquely as a linear combination of the monomials in $\mathcal{H}_d$.*

To prove the above lemma, we will first show that there are at least $q^{\binom{n}{d}}$ distinct degree-d polynomial functions on the slice, and then that the monomials in $\mathcal{H}_d$ span all these functions. Since $|\mathcal{H}_d|$ is exactly $\binom{n}{d}$, these two statements immediately implies Lemma 4.3.

We start with the lower bound on $|\mathcal{P}_d(n, k, \mathbb{Z}_q)|$. The proof is implicit in the work of Wilson [Wil90]. For the sake of completeness, we give a proof in Appendix A.3.

Lemma 4.4 (Number of degree-d polynomials on the slice)**.** *[Wil90]. For every degree parameter $d \in \mathbb{N}$ and for every slice parameter k such that $d \leqslant \min\{k, n-k\}$, the number of distinct degree-d polynomial functions on $\{0,1\}_k^n$ is at least $q^{\binom{n}{d}}$.*

We now show that $\mathcal{H}_d$ is a spanning set for $\mathcal{P}_d(n, k, \mathbb{Z}_q)$, which is our next claim. This proof is also inspired by [Wil90] who proves this in the case when the polynomials have coefficients that are real numbers.

Claim 4.5. *Fix a cyclic group $\mathbb{Z}_q$ where q is a power of prime p. Fix a degree parameter $d \in \mathbb{N}$. For all natural numbers n, k such that $d \leqslant k \leqslant (n-d)$ and k is (d,p)-good, the following holds. The set $\mathcal{H}_d$ of homogeneous degree-d monomials is a spanning set for $\mathcal{P}_d(n,k,\mathbb{Z}_q)$.*

Proof Idea: Every monomial of degree strictly less than d can be expressed as a linear combination of monomials of degree exactly equal to d using the fact that we are working over the slice $\{0,1\}^n_k$. As we are working over a group of prime power order, we have to be careful about the coefficients arising while expressing lower degree monomials using homogeneous monomials. We will use that k is (d,p)-good and Lucas's theorem (Lemma 2.2) in a crucial way to argue about the coefficients.

Proof of Claim 4.5. Consider any monomial $\mathfrak{m}$ of degree $0 \leqslant i < d$. For simplicity in notations, assume without loss of generality that $\mathfrak{m} = x_1 x_2 \cdots x_i$. Let $\mathcal{H}_d|_{\mathfrak{m}}$ denote the subset of $\mathcal{H}_d$ which is divisible by $\mathfrak{m}$, i.e.

$$\mathcal{H}_d|_{\mathfrak{m}} \;:=\; \left\{ x^T \in \mathcal{H}_d \;\middle|\; \{1,\ldots,i\} \subset T \right\}$$

Since every monomial in $\mathcal{H}_d|_{\mathfrak{m}}$ is divisible by $\mathfrak{m}$, it is easy to verify that over the slice $\{0,1\}^n_k$, the following identity holds:

$$\sum_{x^T \in \mathcal{H}_d|_{\mathfrak{m}}} x^T \;=\; \mathfrak{m} \sum_{\substack{T' \subseteq [n] \setminus [i] \\ |T'| = d-i}} x^{T'} \;=\; \mathfrak{m} \cdot \binom{k-i}{d-i}$$

To write the monomial $\mathfrak{m}$ as a linear combination of monomials in $\mathcal{H}_d$, we need the integer $\binom{k-i}{d-i}$ to be invertible in the ring $\mathbb{Z}_q$. This happens, exactly when $\binom{k-i}{d-i}$ is non-zero modulo the prime p.

Using Lucas's theorem Lemma 2.2, we now argue that if the slice k is (d,p)-good, then for all $0 \leqslant i < d$,

$$\binom{k-i}{d-i} \not\equiv 0 \mod p$$

Let $r = p^\ell$ be the smallest power of p strictly greater than d. From Definition 4.2, we know that $k \equiv d \mod p^\ell$. Fix any $0 \leqslant i \leqslant (d-1)$. If we represent $(k-i)$ and $(d-i)$ in p-ary representation, then, we get,

$$(d-i) \;=\; \sum_{j=0}^{\ell-1} b_j p^j + \sum_{j=\ell}^{m} 0 p^j, \qquad (k-i) \;=\; \sum_{j=0}^{\ell-1} a_j p^j + \sum_{j=\ell}^{m} a_j p^j$$

One can verify that $a_j = b_j$ for all $0 \leqslant j \leqslant (\ell-2)$ and $a_{\ell-1} \geqslant b_{\ell-1}$.[7] Thus by Lucas's theorem Lemma 2.2, $\binom{k-i}{d-i} \not\equiv 0 \mod p$. This concludes our proof that a degree strictly less than d monomial can be expressed as a linear combination of monomials in $\mathcal{H}_d$. ■

[7] This is true by assumption for $i = 0$. The fact that it is also true when $i > 0$ follows from elementary properties of subtraction.

Note that Lemma 4.4 and Claim 4.5 together imply Lemma 4.3.

Now we are ready to prove the following distance lemma for degree-d polynomials for good slices. The proof will be by a random restriction, which allows us to reduce to the case of Theorem 3.18. It random restriction sets a uniformly random set of $(n-2k)$ variables to 0 (i.e. we are reducing from $\{0,1\}_k^n$ to $\{0,1\}_k^{2k}$). The main step is to argue that a non-zero polynomial on the k^{th} slice continues to be a non-zero polynomial on the balanced slice (in a smaller dimension) with good enough probability. For this, we use $\mathcal{H}_d$ as a basis $\mathcal{P}_d(n,k,\mathbb{Z}_q)$ for good slices.

Lemma 4.6 (Distance Lemma over good slices)**.** *There exists an absolute constant $\gamma > 0$ so that the following holds. Fix a cyclic group $\mathbb{Z}_q$ where q is a power of a prime number p. Fix a degree parameter $d \in \mathbb{N}$. For all natural numbers n, k such that $1 + d \leqslant k^\gamma$ and $k \leqslant n/2$ and k is (d,p)-good, the following holds.*
For every non-zero degree-d polynomial $P(\mathbf{x}) \in \mathcal{P}_d(n,k,\mathbb{Z}_q)$,

$$\Pr_{\mathbf{x}\sim U_{n,k}}[P(\mathbf{x}) \neq 0] \geqslant \alpha^d \left(1 - \frac{1}{k^{\Omega(1)}}\right), \text{ where } \alpha := \frac{k}{n}$$

Proof of Lemma 4.6. We start by describing the random process to reduce the problem from k^{th} slice to the balanced slice in a smaller dimension cube.

Random process and the new polynomial Sample a random subset $T \subseteq [n]$ of size exactly $2k$ and set all the variables NOT in T to 0. Let $\tilde{P}(y_1,\ldots,y_{2k})$ be the resulting polynomial in $2k$ variables.[8] Note that $\deg(\tilde{P}) \leqslant \deg(P) = d$.
We will now argue that if $P(x_1,\ldots,x_n)$ is a non-zero degree-d polynomial in $\mathcal{P}_d(n,k,\mathbb{Z}_q)$, then $\tilde{P}(y_1,\ldots,y_{2k})$ is a non-zero degree-d polynomial in $\mathcal{P}_d(2k,k,\mathbb{Z}_q)$ with some good enough probability.

Claim 4.7. *Let $P(\mathbf{x}) \in \mathcal{P}_d(n,k,\mathbb{Z}_q)$ be a non-zero polynomial. Then $\tilde{P}$, as defined above, is a non-zero polynomial over $\{0,1\}_k^{2k}$ with probability at least $(2k/n)^d \cdot (1 - (d^2/2k))$ over the randomness of set T.*

Proof of Claim 4.7. Firstly, represent the polynomial $P(x_1,\ldots,x_n)$ as a unique linear combination of monomials in $\mathcal{H}_d$ using Claim 4.5. Let $\mathfrak{m}$ be a monomial in $\mathcal{H}_d$ which has a non-zero coefficient in the polynomial $P(\mathbf{x})$. Assume without loss of generality that $\mathfrak{m} = x_1 x_2 \cdots x_d$. The probability over the choice of T that $\{1,\ldots,d\} \subset T$ is:

$$\binom{n-d}{2k-d} \Big/ \binom{n}{2k}$$

Since $d \leqslant 2k \leqslant n$, we have the following inequality:

$$\binom{n-d}{2k-d} \geqslant \binom{n}{2k} \cdot \left(\frac{2k-d}{n-d}\right)^d$$

[8] We identify the elements in T with $[2k]$ in a canonical way.

Using the inequality $(1-x)^m \geqslant (1-mx)$ and upper bounding $(n-d)^d$ by n^d, we get the following inequality:

$$\binom{n-d}{2k-d} \geqslant \left(\frac{2k}{n}\right)^d \cdot \left(1 - \frac{d^2}{2k}\right)$$

By Lemma 4.3, the probability of $\tilde{P}$ is a non-zero polynomial function over $\{0,1\}_k^{2k}$ is at least the probability that the monomial $\mathfrak{m}$ has non-zero coefficient in $\tilde{P}$, and this is at least $(2k/n)^d \cdot (1 - (d^2/2k))$. This finishes the proof of Claim 4.7. ■

Note that if $\tilde{P}(\tilde{\mathbf{a}}) \neq 0$ for some $\tilde{\mathbf{a}} \in \{0,1\}_k^{2k}$, then $P(\mathbf{a}) \neq 0$ where $\mathbf{a}$ is obtained from $\tilde{\mathbf{a}}$ and fixing the coordinates not in T to 0. It is also easy to see that for a random choice of T, if $\tilde{\mathbf{a}} \sim U_{2k,k}$, then the corresponding $\mathbf{a} \sim U_{n,k}$. Thus we get,

$$\Pr_{\mathbf{x} \sim U_{n,k}}[P(\mathbf{x}) \neq 0] \geqslant \Pr_T[\tilde{P} \text{ doesn't vanish on } \{0,1\}_k^{2k}] \cdot \Pr_{\mathbf{y} \sim U_{2k,k}}[\tilde{P}(\mathbf{y}) \neq 0 \mid \tilde{P} \text{ doesn't vanish on } \{0,1\}_k^{2k}].$$

By using $d \leqslant k^{\varepsilon'}$ for sufficiently small ε' and applying the distance lemma for the balanced slices Theorem 3.18 on $\tilde{P} \in \mathcal{P}_d(2k,k,d)$ and Claim 4.7, we get,

$$\Pr_{\mathbf{x} \sim U_{n,k}}[P(\mathbf{x}) \neq 0] \geqslant \frac{1}{2^d}\left(1 - \frac{1}{k^{\Omega(1)}}\right) \cdot \left(\frac{2k}{n}\right)^d \cdot \left(1 - \frac{d^2}{2k}\right) \geqslant \left(\frac{k}{n}\right)^d \left(1 - \frac{1}{k^{\Omega(1)}}\right).$$

This finishes the proof of Lemma 4.6. ■

Now we will show how to reduce a non-good slice to a good slice by randomly fixing some $\mathcal{O}(d)$ many variables to 1. We will prove the following lemma.

Lemma 4.8 (Reducing any slice to a good slice)**.** *Fix a cyclic group $\mathbb{Z}_q$ where q is a power of a prime number p, a degree parameter $d \in \mathbb{N}$. Let n and $k \in [n^{1/3}, n/2]$ be positive integers, and let $0 \leqslant c \leqslant 2d \leqslant k^{\varepsilon}$ for a sufficiently small constant $\varepsilon > 0$. Let $\beta \in (0, (k/n)^d)$ be such that for every non-zero polynomial $Q(\mathbf{x}) \in \mathcal{P}_d(n-c, k-c, \mathbb{Z}_q)$, it holds that*

$$\Pr_{\mathbf{x} \sim U_{(n-c),(k-c)}}[Q(\mathbf{x}) \neq 0] \geqslant \beta.$$

Then for every non-zero polynomial $P(\mathbf{x}) \in \mathcal{P}_d(n,k,\mathbb{Z}_q)$, it holds that

$$\Pr_{\mathbf{x} \sim U_{n,k}}[P(\mathbf{x}) \neq 0] \geqslant \beta\left(1 - \frac{c}{n^{0.1}}\right).$$

Together with Lemma 4.6, this completes the proof of Lemma 4.1 as shown below. We recall the statement first.

Lemma 4.1 (Distance lemma for cyclic groups)**.** *There exists an absolute constant $\varepsilon > 0$ so that the following holds. Fix a cyclic group $\mathbb{Z}_q$ where $q = p^\ell$ for some prime p and a degree parameter $d \in \mathbb{N}$. For all natural numbers n and k such that $1 + d \leqslant k^\varepsilon$ and $k \leqslant n/2$, the following is true.*
For a non-zero degree-d polynomial $P : \{0,1\}_k^n \to \mathbb{Z}_q$,

$$\Pr_{\mathbf{x} \sim U_{n,k}} [P(\mathbf{x}) \neq 0] \geqslant \alpha^d \left(1 - \frac{1}{k^{\Omega(1)}}\right), \qquad \text{where } \alpha := \frac{k}{n}.$$

Proof of Lemma 4.1. We first argue that we can assume that $k \geqslant n^{1/3}$. Otherwise, Lemma 1.2 suffices to give us the desired bound. Indeed, if $k \leqslant n^{1/3}$, we have

$$\frac{\binom{n-2d}{k-d}}{\binom{n}{k}} \geqslant \left(\frac{k-d}{n}\right)^d \left(\frac{n-2k}{n}\right)^d \geqslant \left(\frac{k}{n}\right)^d \left(1 - \frac{d^2}{k}\right)\left(1 - \frac{2kd}{n}\right) \geqslant \left(\frac{k}{n}\right)^d \left(1 - \frac{1}{k^{\Omega(1)}}\right).$$

Hence, for the rest of the proof, we will assume that $k \leqslant n^{1/3}$. We show below that it suffices to show that for every slice $k > d$, there exists $c \in [0, 2d]$ such that the slice $(k - c)$ is (d, p)-good (see Definition 4.2). Assuming this, the premise of Lemma 4.8 is true with

$$\beta = \left(\frac{k-c}{n-c}\right)^d \left(1 - \frac{1}{(k-c)^\gamma}\right),$$

for some constant $\gamma \in (0, 1)$, by the distance lemma for good slices (Lemma 4.6).

The conclusion of Lemma 4.8 then implies that

$$\Pr_{\mathbf{x} \sim U_{n,k}} [P(\mathbf{x}) \neq 0] \geqslant \left(\frac{k-c}{n-c}\right)^d \left(1 - \frac{1}{(k-c)^\gamma}\right)\left(1 - \frac{c}{n^{0.1}}\right) \geqslant \left(\frac{k}{n}\right)^d \left(1 - \frac{1}{k^{\Omega(1)}}\right),$$

using $d \leqslant k^\varepsilon$ and $c \leqslant 2d$. Hence, it only remains to show that there exists a $c \in [0, 2d]$ such that the slice $(k - c)$ is (d, p)-good.

Let $d = \sum_{j=0}^{\ell} b_j p^j$ and $k = \sum_{j=0}^{m} a_j p^j$ be the p-ary representations of d and k respectively, with $b_\ell, a_m > 0$ and $m \geqslant \ell$ (since $k \geqslant d$). We first note that for $c_1 \in [0, p^\ell - 1]$ such that $c_1 \equiv k - d \mod p^\ell$, we have $k_1 := k - c_1 \equiv d \mod p^\ell$. Or equivalently for $k_1 = \sum_{j=0}^{m} u_j p^j$ in p-ary representation, we have

$$u_j = b_j, \text{ for all } j < \ell. \tag{5}$$

We now show that there exists a $c_2 \in [0, d]$ such that when $k_2 := k_1 - c_2$ is expressed in its p-ary representation as $k_2 = \sum_{j=0}^{m} v_j p^j$, it holds that $v_\ell \geqslant b_\ell$ and $v_j = u_j$, for all $j < \ell$. We have two cases:

- **Case 1:** $u_\ell \geqslant b_\ell$**.** In this case, we can take $c_2 = 0$ and $k_2 = k_1$.
- **Case 2:** $u_\ell < b_\ell$**.** We take $c_2 = (u_\ell + 1)p^\ell \leqslant b_\ell \cdot p^\ell \leqslant d$. In this case, the p-ary representation of k_2 as described above satisfies $v_j = u_j$ for all $j < \ell$ and $v_\ell = p - 1$.

Taking $c = c_1 + c_2$, we see that $k_2 = k - c$ is indeed (d, p)-good, with $c \leqslant 2d$.

This finishes the proof of Lemma 4.1. ■

We now prove Lemma 4.8.

Proof of Lemma 4.8. Note that the proof is trivial for $c = 0$. We will now show the lemma for $c \geqslant 1$. Consider the following way of sampling a point $\mathbf{x}$ from $U_{n,k}$.

1. Initialize $S = \varnothing$.
2. Choose $s \in [n] \backslash S$ uniformly at random, set $x_s = 1$, and add s to S.
3. Repeat Step 2 until $|S| = c$.
4. Set the co-ordinates of $\mathbf{x}$ not in S according to the distribution $U_{n-c,k-c}$.
5. Output $\mathbf{x}$.

Let $P_1 \in \mathcal{P}_d(n-1, k-1, \mathbb{Z}_q), P_2 \in \mathcal{P}_d(n-2, k-2, \mathbb{Z}_q), \ldots, P_c \in \mathcal{P}_d(n-c, k-c, \mathbb{Z}_q)$ be the restrictions of the polynomial P on the respective domains obtained by setting $x_{s_1} = 1$, $x_{s_2} = 1, \ldots, x_{s_c} = 1$ successively by the above random procedure, so that $S = \{s_1, s_2, \ldots, s_c\} \in \binom{[n]}{c}$.

We claim that P_1 is non-zero on $\{0,1\}^{n-1}_{k-1}$, with high probability over the choice of s_1. To prove this, let us call an index $s \in [n]$ "bad" if for all $\mathbf{a} \in \{0,1\}^n_k$ such that $a_s = 1$, we have that $P(\mathbf{a}) = 0$. We observe that the probability of P_1 being entirely zero over $\{0,1\}^{n-1}_{k-1}$ is equal to the probability of a uniformly random $s \in [n]$ being bad. We show below that the number of such bad indices $s \in [n]$ is at most $\ell := \lfloor n/\sqrt{k} \rfloor$. Towards a contradiction, suppose there are some ℓ bad indices $i_1, i_2, \ldots, i_\ell \in [n]$. This means that if at least one of these co-ordinates takes value 1, $P(\mathbf{x})$ evaluates to 0. Thus the number of non-zeroes of P in $\{0,1\}^n_k$ is upper bounded by the number of points that take the value 0 on all these bad indices, i.e., $\binom{n-\ell}{k}$. However, by the weak distance lemma (Lemma 1.2), we know that the number of non-zeroes has to be at least $\binom{n-2d}{k-d}$. This yields a contradiction as we have $\binom{n-\ell}{k} < \binom{n-2d}{k-d}$ by the following claim.

Claim 4.9. *For every sufficiently large integer n and arbitrary integers $k \in [n^{1/4}, n/2]$, $d \in [1, k^{0.1}]$, and $\ell = \lfloor n/\sqrt{k} \rfloor$, we have*

$$\binom{n-\ell}{k} < \binom{n-2d}{k-d}.$$

Proof. We have

$$\binom{n-\ell}{k} \leqslant \binom{n-2d}{k} \cdot \left(1 - \frac{k}{n}\right)^{\ell-2d} < \binom{n-2d}{k-d} \cdot n^d \cdot \left(1 - \frac{k}{n}\right)^{\ell-2d}.$$

Hence, it suffices to show that $(1 - \frac{k}{n})^{\ell-2d} \leqslant \frac{1}{n^d}$. Using $\ell - 2d \geqslant n/(4\sqrt{k})$, indeed we have that $(1 - \frac{k}{n})^{\ell-2d} \leqslant (1 - \frac{k}{n})^{n/(4\sqrt{k})} \leqslant e^{-\sqrt{k}/4} \leqslant \frac{1}{n^d}$, using $d \leqslant k^{0.1}$ and $k \geqslant n^{1/4}$. ■

By using the notation $P' \not\equiv 0$ to denote that a polynomial $P' \in \mathcal{P}_d(n', k', \mathbb{Z}_q)$ has at least one non-zero evaluation over the underlying slice, we have

$$\Pr_{s_1 \sim [n]}[P_1 \not\equiv 0] \geqslant 1 - \frac{\ell}{n} \geqslant 1 - \frac{1}{\sqrt{k}} \geqslant 1 - \frac{1}{n^{1/6}}.$$

By the same argument for $P_2, \ldots, P_c$, we get with probability at least $1 - \frac{1}{(n-1)^{1/6}}$ (over the choice of S), that P_c has at least one non-zero evaluation over $\{0,1\}^{n-c}_{k-c}$. Here, we note that we will need to show that Claim 4.9 is also applicable if we replace n and k with $n-i$ and $k-i$ for every $i \in [c]$; this follows as $c \leqslant 2d \leqslant k^{\varepsilon}$, so $d \leqslant (k-i)^{0.1}$ and $k - i \in [(n-i)^{1/4}, (n-i)/2]$. Hence, at each step $i \in [1, c-1]$, conditioned on P_i being non-zero over $\{0,1\}^{n-i}_{k-i}$, we get that P_{i+1} is non-zero over $\{0,1\}^{n-i-1}_{k-i-1}$ with probability at least $1 - \frac{1}{(n-i)^{1/6}} \geqslant 1 - \frac{1}{n^{0.1}}$.

Finally, applying the premise of the lemma statement (Lemma 4.8) for $Q = P_c$, we get the desired bound. More formally, we obtain

$$\begin{aligned}
\Pr_{\mathbf{x} \sim U_{n,k}}[P(\mathbf{x}) \neq 0] &\geqslant \Pr_{s_1 \sim [n]}[P_1 \not\equiv 0] \cdot \Pr_{s_2 \sim [n] \setminus \{s_1\}}[P_2 \not\equiv 0 \mid P_1 \not\equiv 0] \cdot \ldots \\
&\quad \cdots \Pr_{s_c \sim [n] \setminus \{s_1, \ldots, s_{c-1}\}}[P_c \not\equiv 0 \mid P_{c-1} \not\equiv 0] \cdot \Pr_{\mathbf{y} \sim U_{n-c,k-c}}[P_c(\mathbf{y}) \neq 0 \mid P_c \not\equiv 0] \\
&\geqslant \left(1 - \frac{1}{n^{0.1}}\right)^c \cdot \beta \\
&\geqslant \beta \left(1 - \frac{c}{n^{0.1}}\right).
\end{aligned}$$

This concludes the proof of Lemma 4.8. ■

4.2 General Abelian Groups

In this subsection, we use the distance lemma for slices over cyclic groups of prime power order Lemma 4.1 to get a distance lemma for slices over arbitrary Abelian groups. We prove Theorem 1.1 now.

Proof of Theorem 1.1. By negating all the variables if necessary, we can assume that $k \leqslant n/2$. Let $P(\mathbf{x}) \in \mathcal{P}_d(n, k, G)$ be the following polynomial:

$$P(\mathbf{x}) = \sum_{\substack{I \subseteq [n] \\ |I| \leqslant d}} a_I \cdot x^I,$$

where $x^I := \prod_{j \in I} x_j$. We first argue that it suffices to prove Theorem 1.1 for **finite** Abelian groups. Then we will use the fundamental theorem of finite Abelian groups and our distance lemma for cyclic groups (Lemma 4.1) to finish the proof.

Reducing to finitely generated Abelian groups Define G' to be the subgroup of G generated by the coefficients of P, i.e.

$$G' = \langle \{a_I \mid I \subseteq [n], |I| \leqslant d\} \rangle$$

Observe that we can treat the polynomial $P(\mathbf{x})$ as a degree-d polynomial over the group G' with a non-zero evaluation on the slice $\{0,1\}^n_k$, i.e. $P(\mathbf{x}) \in \mathcal{P}_d(n,k,G')$. Thus we have reduced to the case of finitely generated Abelian groups. Using the structure theorem of finitely generated Abelian groups, we know that G' is isomorphic to

$$G' \;\cong\; \mathbb{Z}_{p_1^{\ell_1}} \times \cdots \times \mathbb{Z}_{p_s^{\ell_s}} \times \mathbb{Z}^r$$

for some prime numbers $p_1, \ldots, p_s$, positive integers $\ell_1, \ldots, \ell_s$ and $r \in \mathbb{Z}_{\geqslant 0}$.

Reducing to finite Abelian groups We now show that there exists a *finite* Abelian group G'' and a polynomial $P'' \in \mathcal{P}_d(n,k,G'')$ such that the following holds:

$$\Pr_{\mathbf{x}\sim U_{n,k}}[P(\mathbf{x}) \neq 0] = \Pr_{\mathbf{x}\sim U_{n,k}}[P''(\mathbf{x}) \neq 0]$$

If $r = 0$, we can take $G'' = G'$ and we are done.
Otherwise, let M be a prime number greater than the absolute values of the last r co-ordinates (using the isomorphism between G and $\mathbb{Z}_{p_1^{\ell_1}} \times \cdots \times \mathbb{Z}_{p_s^{\ell_s}} \times \mathbb{Z}^r$) of the evaluations $P(\mathbf{x})$ for all $\mathbf{x} \in \{0,1\}^n_k$. Then, we take $G'' := \mathbb{Z}_{p_1^{\ell_1}} \times \cdots \times \mathbb{Z}_{p_s^{\ell_s}} \times \mathbb{Z}^r_M$. Let $a_I = (a_I(1), \ldots, a_I(s+r)) \in G'$ be a coefficient of $P(\mathbf{x})$. Define the coefficient $a''_I \in G''$ as follows:

$$a''_I := (a_I(1), a_I(2), \ldots, a_I(s), a_I(s+1) \mod M, \ldots, a_I(s+r) \mod M)$$

Now define the polynomial $P''(\mathbf{x})$ as follows:

$$P''(\mathbf{x}) = \sum_{\substack{I\subseteq[n]\\ |I|\leqslant d}} a''_I \cdot x^I,$$

where $x^I := \prod_{j\in I} x_j$. For every $\mathbf{x} \in \{0,1\}^n_k$. we have that $P(\mathbf{x}) = 0 \iff P''(\mathbf{x}) = 0$, since by the definition of M, the last r co-ordinates of $P(\mathbf{x})$ can only take values strictly in between $-M$ and M. Thus we have reduced to finite Abelian group G''.

Cyclic groups of prime power order We will now argue that we can further reduce it to the case of cyclic groups of prime power order. For simplicity of notation, let the primes $p_{s+1} = \cdots = p_{s+r} = M$, exponents $\ell_{s+1} = \cdots = \ell_{s+r} = 1$. We thus have $G'' \cong \mathbb{Z}_{p_1^{\ell_1}} \times \cdots \times \mathbb{Z}_{p_{s+r}^{\ell_{s+r}}}$ and $P''(\mathbf{x}) \in \mathcal{P}_d(n,k,G'')$ is non-zero on $\{0,1\}^n_k$. This means that there exists $j \in [s+r]$ such that the polynomial $P''(\mathbf{x})$ is a non-zero degree-d polynomial on $\{0,1\}^n_k$ over the cyclic group $\mathbb{Z}_{p_j^{\ell_j}}$.

Using Lemma 4.1, we know there exists a constant ε such that $P''(\mathbf{x})$ is non-zero on at least $\alpha^d(1-\mathcal{O}(1/k^\varepsilon)) \cdot \binom{n}{k}$ points over the cyclic group $\mathbb{Z}_{p_j^{\ell_j}}$ where $\alpha = k/n$. This implies that the polynomial $P''(\mathbf{x})$ is non-zero on at least $\alpha^d(1-\mathcal{O}(1/k^\varepsilon)) \cdot \binom{n}{k}$ points over the group G''. As we argued above, this in particular implies that the polynomial polynomial $P(\mathbf{x})$ is non-zero on at least $\alpha^d(1-\mathcal{O}(1/k^\varepsilon))$ fraction of points on the slice $\{0,1\}^n_k$ over the Abelian group G. This finishes the proof of Theorem 1.1. ■

5 Low-degree Functions Over Slices

In this section we will give a simple proof of a lemma of Filmus and Ihringer [FI19] following the proof idea of Nisan and Szegedy [NS92]. We will first give a couple of definitions and set the notations for this section.

For a function $f(x_1, \ldots, x_n)$ on the slice $\{0,1\}^n_k$ with coefficients in $\mathbb{R}$, and for any two coordinates $i, j \in [n] \times [n]$, define $f^{(ij)}$ to be the function where we swap the i^{th} variable with the j^{th} variable in the function f.

Definition 5.1 (Influence)**.** *Let $f : \{0,1\}^n \to \mathbb{R}$ be a function on the slice $\{0,1\}^n_k$. For $(i,j) \in [n] \times [n]$, the $(i,j)^{th}$-influence of f, denoted by $\mathrm{Inf}_{ij}(f)$, is defined as,*

$$\mathrm{Inf}_{ij}(f) := \frac{1}{4} \Pr_{\mathbf{x} \sim U_{n,k}} [f(\mathbf{x}) \neq f^{(ij)}(\mathbf{x})]$$

The total influence of f, denoted by $\mathrm{Inf}(f)$, is defined as,

$$\mathrm{Inf}(f) := \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} \mathrm{Inf}_{ij}(f)$$

Note that if $i = j$ in the above definition, then $\mathrm{Inf}_{ii}(f) = 0$ and it does not contribute anything towards the total influence.

A key lemma in the proof of [FI19] is a lower bound on every non-zero influence (see [FI19, Lemma 3.1]). They showed that there exists a constant α such that every non-zero influence of a degree-d polynomial on the balanced slice is at least α^d. The proof of this lemma in [FI19] uses analytic techniques such as the Log-Sobolev inequality on the Boolean slice [LY98] and the Hypercontractive inequality [DS96]. Using our distance lemma for the balanced slices (Theorem 3.18), we can improve the lower bound to almost $1/2^d$ (which is easily seen to be tight up to constant factors). Note that the main result of this section only holds for degree $d \leqslant C \log n$ for some absolute constant $C > 0$, it suffices to use the simpler proof of Theorem 3.18. We state the lemma below.

Lemma 5.2 (Lower bound on influences)**.** *Let $f(x_1, \ldots, x_n)$ be a non-zero degree-d function on the balanced slice $\{0,1\}^n_{n/2}$. Then for every $(i,j) \in [n] \times [n]$ for which $\mathrm{Inf}_{ij}(f) > 0$, the following holds:*

$$\mathrm{Inf}_{ij}(f) \geqslant \frac{1}{4} \cdot \frac{1}{2^d} \cdot \left(1 - \frac{1}{n^\varepsilon}\right)$$

for some absolute constant $\varepsilon > 0$.

Proof. Fix some pair $(i,j) \in [n] \times [n]$ with $\mathrm{Inf}^{(ij)}(f) > 0$ and consider the polynomial g_{ij} on the balanced slice, defined as follows: $g_{ij}(\mathbf{x}) := f(\mathbf{x}) - f^{(ij)}(\mathbf{x})$.

Observe that since f is a degree-d polynomial, $f^{(ij)}$ is also a degree-d polynomial, which means g_{ij} is also a degree-d polynomial on the balanced slice. Since the influence $\mathrm{Inf}_{ij}(f) > 0$, this means that g_{ij} is non-zero on the slice $\{0,1\}^n_{n/2}$. Now using our distance lemma on the balanced slice Theorem 3.18,

$$\Pr_{\mathbf{x}\sim U_{n,n/2}}[f(\mathbf{x}) \neq f^{(ij)}(\mathbf{x})] \;=\; \Pr_{\mathbf{x}\sim U_{n,n/2}}[g_{ij}(\mathbf{x}) \neq 0] \;\geqslant\; \frac{1}{2^d}\cdot\left(1-\frac{1}{n^{\varepsilon}}\right)$$
$$\Rightarrow \mathrm{Inf}_{ij}(f) \;\geqslant\; \frac{1}{4}\cdot\frac{1}{2^d}\cdot\left(1-\frac{1}{n^{\varepsilon}}\right)$$

for some absolute constant $\varepsilon > 0$. ■

Filmus and Ihringer [FI19, Lemma 3.3] use this lower bound on non-zero influences to get a bound on junta of degree-d polynomials on the balanced slice. Using the above-mentioned improved lower bound on non-zero influence, we can also improve the bounds in [FI19, Lemma 3.3].

Lemma 5.3. *There exists an absolute constant $C > 0$ such that for all degree parameters $d \in \mathbb{N}$ such that $d \leqslant C\log n$, the following holds. Every degree-d polynomial on the slice $\{0,1\}^n_{n/2}$ is a $\eta(d)$-junta, where*

$$\eta(d) = \mathcal{O}(d\cdot 2^d).$$

Proof. The proof is essentially the same proof as in [FI19], except for one inequality which can be improved using Lemma 5.2. Let $f(\mathbf{x}) \in \mathcal{P}_d(n, n/2, \mathbb{R})$ and let G be a graph on the vertex set $[n]$ where (i,j) is an edge if $\mathrm{Inf}_{ij}(f) \geqslant 1/2^d\cdot(1-1/n^{\varepsilon})$, where ε is the absolute constant from Lemma 5.2. Let M be a maximal matching in G. We now proceed similar to the proof in [FI19, Lemma 3.3] and we request the reader to refer [FI19] as we just highlight the changes in the proof here.

Using Lemma 5.2, we get the following two inequalities upper and lower bounding the influence:

$$\left(\frac{1}{4}\cdot\frac{1}{2^d}\cdot\left(1-\frac{1}{n^{\varepsilon}}\right)\right)\cdot\left(1-\frac{1}{n}\right)\cdot M \;\leqslant\; \mathrm{Inf}(f) \;\leqslant\; d$$
$$\Rightarrow M \leqslant \mathcal{O}(d\cdot 2^d),$$

where we used the assumption $d \leqslant C\log n$ in upper bounding $1/n^{\varepsilon}$ by $\frac{1}{10}\cdot\frac{1}{2^d}$. Following the argument of [FI19], this gives us that f is a $2M$-junta, i.e., a $\mathcal{O}(d\cdot 2^d)$-junta. ■

As already noted in the introduction, a stronger upper bound of $\eta(d) = \mathcal{O}(2^d)$ follows from the work of [FI19; CHS20] (and can also be obtained by plugging Lemma 5.2 in place of [FI19, Lemma 3.3] in the proof of [FI19]). The advantage here is the relatively simple proof following exactly the template of [NS92].

6 Improved Bound for Linear Functions

In this section, we show how to obtain an improvement over our distance lemma (Theorem 1.1) for the case of linear polynomials, i.e., $d = 1$. In particular, we will show the following.

Theorem 6.1. *Let G be an arbitrary Abelian group, and $n \geqslant 8$ and $k \in [n-1]$ be positive integers. Then, for every polynomial $P(\mathbf{x}) \in \mathcal{P}_1(n,k,G)$ that is non-zero on $\{0,1\}^n_k$, we have*

$$\Pr_{\mathbf{x}\sim\{0,1\}^n_k}[P(\mathbf{x}) \neq 0] \geqslant \frac{t-1}{n},$$

where $t = \min\{k, n-k\}$.

This is an improvement over Theorem 1.1 as we have an additive term of $1/n$ as opposed to $1/n^{\varepsilon}$ for some constant $\varepsilon \in (0,1)$. In particular, in the regime $k \leqslant n^{\delta}$ for small enough $\delta \in (0,1)$, the above theorem gives a lower bound of $1/n^{1-\delta}$, while Theorem 1.1 fails to give anything non-trivial.

We note that this is also an improvement over the weak distance lemma shown by [ABPSS24] (i.e., Lemma 1.2) as it gives a lower bound of $\frac{k(n-k)}{n(n-1)}$ which is less than $\frac{k-1}{n}$ for $k \geqslant \sqrt{n}+1$. In particular, taking $P(\mathbf{x}) = x_1$, we see that our bound is almost tight. Furthermore, the additive term of $1/n$ in the above theorem cannot be avoided (at least for $k = n/2$ and up to a constant factor) as the polynomial $P(\mathbf{x}) = x_1 + x_2 + 1$ in $\mathcal{P}_1(n, k = n/2, \mathbb{Z}_2)$ which is non-zero with probability $\frac{1}{2} - \frac{1}{2(n-1)}$.

We will now prove Theorem 6.1.

Proof of Theorem 6.1. Let $P(\mathbf{x}) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + c$, where $a_1, \ldots, a_n, c \in G$. We may (and we will) assume that $a_1, a_2, \ldots, a_n$ are not all equal (abbreviated as NAE from now), as otherwise the polynomial always evaluates to a constant over $\{0,1\}^k$ since $\sum_{i\in[n]} x_i = k$ for all $\mathbf{x} \in \{0,1\}^n_k$. Hence the desired probability is equal to 1 since we are guaranteed that $P(\mathbf{x}) \neq 0$ for at least one point $\mathbf{x} \in \{0,1\}^n_k$. We may further assume that $k \leqslant n/2$, as otherwise, we can consider the evaluations of the polynomial $P(1-x_1, \ldots, 1-x_n)$ over $\mathbf{x} \in \{0,1\}^n_{n-k}$.

We divide the proof into two cases depending on whether k divides n:

Case 1: k divides n. That is, $n = mk$ for some integer $m \geqslant 2$. Let $M \in [n]^{k\times m}$ be the matrix formed by arranging $[n]$ according to a uniformly random permutation. Denoting the j-th column of M by M_j, let $\mathrm{sum}(M_j) = \sum_{i\in M_j} a_i$. We note that

$$\Pr_{x\sim\{0,1\}^n_k}[P(\mathbf{x}) \neq 0] = \Pr_{M}[\mathrm{sum}(M_1) \neq -c], \tag{6}$$

as both LHS and RHS are essentially picking k elements (without replacement) from $a_1, \ldots, a_n$ uniformly at random and checking whether their sum is not equal to $-c$. We construct M by the following random process:

- Partition $[n]$ into k buckets, $B_1, B_2, \ldots, B_k \subseteq [n]$, each of size m, uniformly at random.
- Set the i-th row of M to be a uniformly random permutation of B_i, for each $i \in [k]$ independently.

We say that an index $i \in [k]$ is "good" if the elements $(a_j)_{j\in B_i}$ are NAE and "bad" otherwise. Since $a_1, \ldots, a_n$ are NAE, one might expect that there is at least one good index with high probability

(over the randomness of the first step i.e., choosing the buckets $B_1, \ldots, B_k$). That is, we give an upper bound on

$$\Pr_{B_1,\ldots,B_k}[\text{all the indices in } [k] \text{ are bad}].$$

Note that for the above probability to be positive, the number of times each a_i appears in $(a_i)_{i\in[n]}$ must be a multiple of m. Thus we may assume that the multiset $(a_i)_{i\in[n]}$ is of the form: b_1 (taken $f_1 m$ times), b_2 (taken $f_2 m$ times), $\ldots b_\ell$ (taken $f_\ell m$ times), where $\ell \geqslant 2$ and $b_1, \ldots, b_\ell$ are mutually distinct elements of G and $f_1, f_2, \ldots, f_\ell \geqslant 1$.

We handle the case of $\ell = m = 2$ and either $f_1 = 1$ or $f_2 = 1$ separately by the following claim.

Claim 6.2. *If $\ell = m = 2$ and at least one of f_1 or f_2 is equal to 1, then*

$$\Pr_{\mathbf{x}\sim\{0,1\}^n_k}[P(\mathbf{x}) \neq 0] \geqslant \frac{1}{2} - \frac{1}{2(n-1)} \geqslant \frac{k-1}{n}.$$

Hence, for the rest of the argument, we will assume that m, ℓ and f_i's are such that the premise of Claim 6.2 are not met. We then bound the bad probability as follows:

Claim 6.3.

$$\Pr_{B_1,\ldots,B_k}[\textit{all the indices in } [k] \textit{ are bad}] = \frac{\binom{k}{f_1,\ f_2,\ \ldots,\ f_\ell}}{\binom{n}{f_1 m,\ f_2 m,\ \ldots,\ f_\ell m}} \leqslant \frac{1}{n}.$$

Given Claim 6.3, we conclude that with probability at least $1 - 1/n$, there exists at least one index $i \in [k]$ such that $(a_j)_{j\in B_i}$ are NAE; suppose $h_1 \neq h_2 \in (a_j)_{j\in B_i} \subseteq G$ be such elements. Recall that we permute the elements corresponding to B_i randomly to form the i-th row of the matrix M. As this is performed independently across the rows, we may fix an arbitrary permutation of all the rows except the i-th one and argue a lower bound on the probability of $P(\mathbf{x})$ being non-zero by (6). Let $g_1, \ldots, g_i, \ldots, g_k \in G$ be the corresponding group elements in the first column, where $(g_j)_{j\neq i}$'s are some constants and g_i is picked uniformly at random from the multiset $(a_j)_{j\in B_i}$. We have $\mathrm{sum}(M_1) = g_1 + \cdots + g_k$. Notice that the sum corresponding to $g_i = h_1$ and $g_i = h_2$ are not equal as $h_1 \neq h_2$. Hence,

$$\Pr_{g_i}[g_1 + \cdots + g_k \neq -c] \geqslant 1/m.$$

Therefore, we get

$$\Pr_{M}[\mathrm{sum}(M_1) \neq -c \mid \text{there exists a good index } i \in [k]] \geqslant 1/m = k/n.$$

Combining with Claim 6.3, this gives us that $\Pr_M[\mathrm{sum}(M_1) \neq -c] \geqslant (k-1)/n$, thus finishing the proof of the theorem (when k divides n).

We now move to the case when k does not divide n.

Case 2: k does not divide n. Suppose $n = mk + k'$ for some positive integers m and $k' < k$. Since $k \leqslant n/2$, this implies $m \geqslant 2$. Similar to the previous argument, we will analyze the probability of $P(\mathbf{x})$ being non-zero using the group elements. In particular, we would like to lower bound the probability of $\sum_{i\in A} a_i \neq -c$, where $A \subseteq [n]$ is subset of size k chosen uniformly at random. We will first sample A by the following process.

– Choose a uniformly random subset B of $[n]$ of size mk.

– Choose a uniformly random subset A of B of size k.

We claim that with probability at least mk/n, the elements $(a_i)_{i\in B}$ are NAE. More formally, we claim that $\Pr_{B\sim\binom{[n]}{mk}}[(a_i)_{i\in B} \text{ are NAE}] \geqslant mk/n$. To show this, suppose $g_1 \neq g_2 \in G$ be two distinct elements in $(a_i)_{i\in[n]}$. We have two cases.

- **Case (i): There exists a $g \in G$ that occurs at least mk times in $(a_i)_{i\in[n]}$.** Since $g_1 \neq g_2$, at least one of these two elements is not equal to g; say $g \neq g_1$ without loss of generality. We claim that $g_1 \in (a_i)_{i\in B}$ implies $(a_i)_{i\in B}$ are NAE. To see this, we note that $n - mk = k' < mk$, so g must always appear in $(a_i)_{i\in B}$. Hence, $(a_i)_{i\in B}$ are NAE whenever $g_1 \in (a_i)_{i\in B}$. As B is a uniformly random subset of size mk, this event (i.e., $g_1 \in (a_i)_{i\in B}$) happens with probability at least mk/n.
- **Case (ii): No such element exists.** In this case, $(a_i)_{i\in B}$ are NAE for all choices of $B \in \binom{[n]}{mk}$.

Therefore, $\Pr_{B\sim\binom{[n]}{mk}}[(a_i)_{i\in B} \text{ are NAE}] \geqslant mk/n$. Conditioned on B satisfying the condition that $(a_i)_{i\in B}$ are NAE, we note that $\sum_{i\in A} a_i \neq -c$ with probability at least $(k-1)/(mk)$ by using the fact that we have already established in Theorem 6.1 for the case when k divides n. Hence, the final probability is

$$\begin{aligned}\Pr_{A\sim\binom{[n]}{k}}\left[\sum_{i\in A} a_i \neq -c\right] &\geqslant \Pr_{B\sim\binom{[n]}{mk}}\left[(a_i)_{i\in B} \text{ are NAE}\right] \cdot \Pr_{A\sim\binom{B}{k}}\left[\sum_{i\in A} a_i \neq -c \;\middle|\; (a_i)_{i\in B} \text{ are NAE}\right] \\ &\geqslant \frac{mk}{n}\cdot\frac{k-1}{mk} = \frac{k-1}{n}.\end{aligned}$$

■

We end with the proofs of Claim 6.2 and Claim 6.3.

Proof of Claim 6.2. We have $k = n/2$. Without loss of generality, suppose $f_1 = 1$. Recall that the multiset $(a_i)_{i\in[n]}$ is equal to b_1 repeated $2f_1 = 2$ times and b_2 repeated $2f_2 = n-2$ times, for some $b_1 \neq b_2 \in G$. We may further assume that $P(\mathbf{x}) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + c$ where $a_1 = a_2 = b_1$ and $a_3 = a_4 = \cdots = a_n = b_2$. As $\sum_{i=1}^n x_i = k$ for all points $\mathbf{x} \in \{0,1\}^n_k$, we have

$$\begin{aligned}P(\mathbf{x}) &= b_1(x_1 + x_2) + b_2(x_3 + x_4 + \cdots + x_n) + c \\ &= b_1(x_1 + x_2) + b_2(k - x_1 - x_2) + c \\ &= (b_1 - b_2)(x_1 + x_2) + kb_2 + c.\end{aligned}$$

Let $b := b_1 - b_2$ and $b' := kb_2 + c$. Then we have, $P(\mathbf{x}) = \begin{cases} b', & \text{if } x_1 = x_2 = 0, \\ b + b', & \text{if } x_1 + x_2 = 1, \\ 2b + b', & \text{otherwise.}\end{cases}$

Since $b \neq 0$, we have the implication

$$b + b' = 0 \implies (b' \neq 0 \text{ and } 2b + b' \neq 0).$$

Hence,

$$\Pr_{\mathbf{x}\sim\{0,1\}^n_{n/2}}[P(\mathbf{x}) \neq 0] \geqslant \min\left\{\Pr_{\mathbf{x}\sim\{0,1\}^n_{n/2}}[x_1 + x_2 = 1], \Pr_{\mathbf{x}\sim\{0,1\}^n_{n/2}}[x_1 + x_2 \neq 1]\right\} = \frac{1}{2} - \frac{1}{2(n-1)}.$$

■

Now, we prove Claim 6.3.

Proof of Claim 6.3. We consider two cases depending on the value of m.

- **Case 1:** $m = 2$**.** We have $n = 2k$. We note that at least one of the two conditions below must be met:
  - $\ell \geqslant 3$, or
  - $\ell = 2$ and $f_1, f_2 \geqslant 2$.

  Regardless of which of the above two conditions is satisfied, we can always partition the multiset $(a_i)_{i\in[n]}$ into two sub(multi)sets $S_1 \subseteq G$ and $S_2 \subseteq G$, such that $4 \leqslant |S_1| \leqslant n/2$ and for every $g_1 \in S_1$ and $g_2 \in S_2$, we have that $g_1 \neq g_2$. We now note that a necessary condition for an index $i \in [k]$ to be bad is $\{a_j | j \in B_i\}$ being a subset (as a multiset) of S_1 or S_2. Hence, the probability that all $i \in [k]$ are bad is at most $\binom{k}{f}/\binom{n}{2f}$, where $f := |S_1|/2 \in [2, k/2]$. As $\binom{k}{f}/\binom{n}{2f}$ is an increasing function of f when $f \leqslant k/2$, we can lower bound it by the value corresponding to $f = 2$, i.e., $\frac{\binom{k}{2}}{\binom{n}{4}} = \frac{3}{(n-1)(n-3)} \leqslant \frac{1}{n}$.
- **Case 2:** $m \geqslant 3$**.** We note that the numerator of the fraction in Claim 6.3 is $A := \binom{k}{f_1,\ f_2,\ \ldots,\ f_\ell} \geqslant k$ (as each $f_i \geqslant 1$) and the denominator is $\binom{km}{f_1 m,\ f_2 m,\ \ldots,\ f_\ell m} \geqslant \binom{k}{f_1,\ f_2,\ \ldots,\ f_\ell}^m = A^m$ by a simple counting argument. Hence, we have

$$\frac{\binom{k}{f_1,\ f_2,\ \ldots,\ f_\ell}}{\binom{km}{f_1 m,\ f_2 m,\ \ldots,\ f_\ell m}} \leqslant \frac{A}{A^m} \leqslant \frac{1}{k^{m-1}} \leqslant \frac{1}{mk} = \frac{1}{n}.$$

  The last inequality follows using $k^{m-2} \geqslant m$ for $k, m \geqslant 3$ and for $k = 2, m = 4$.

■

References

[AC88] N. Alon and F.R.K. Chung. "Explicit construction of linear sized tolerant networks". In: *Discrete Mathematics* 72.1 (1988), pp. 15–19. ISSN: 0012-365X. DOI: https://doi.org/10.1016/0012-365X(88)90189-6. URL: https://www.sciencedirect.com/science/article/pii/0012365X88901896 (cit. on p. 5).

[ABCO88] Noga Alon, Ernest E. Bergmann, Don Coppersmith, and Andrew M. Odlyzko. "Balancing sets of vectors". In: *IEEE Trans. Inf. Theory* 34.1 (1988), pp. 128–130. DOI: 10.1109/18.2610. URL: https://doi.org/10.1109/18.2610 (cit. on p. 3).

[AF93] Noga Alon and Zoltán Füredi. "Covering the Cube by Affine Hyperplanes". In: *Eur. J. Comb.* 14 (1993), pp. 79–83. URL: https://api.semanticscholar.org/CorpusID:36232875 (cit. on p. 6).

[ABPSS24] Prashanth Amireddy, Amik Raj Behera, Manaswi Paraashar, Srikanth Srinivasan, and Madhu Sudan. "Low Degree Local Correction Over the Boolean Cube ". In: *Electron. Colloquium Comput. Complex.* TR24-164 (2024). ECCC: TR24-164. URL: https://eccc.weizmann.ac.il/report/2024/164 (cit. on pp. 3, 4, 32).

[ALMSS98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. "Proof Verification and the Hardness of Approximation Problems". In: *J. ACM* 45.3 (1998), pp. 501–555. DOI: 10.1145/278298.278306. URL: https://doi.org/10.1145/278298.278306 (cit. on p. 3).

[BFLS91] László Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. "Checking Computations in Polylogarithmic Time". In: *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*. Ed. by Cris Koutsougeras and Jeffrey Scott Vitter. ACM, 1991, pp. 21–31. DOI: 10.1145/103418.103428. URL: https://doi.org/10.1145/103418.103428 (cit. on p. 3).

[BHLR19] Abhishek Bhrushundi, Kaave Hosseini, Shachar Lovett, and Sankeerth Rao. "Torus Polynomials: An Algebraic Approach to ACC Lower Bounds". In: *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*. Ed. by Avrim Blum. Vol. 124. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, 13:1–13:16. DOI: 10.4230/LIPICS.ITCS.2019.13. URL: https://doi.org/10.4230/LIPIcs.ITCS.2019.13 (cit. on p. 3).

[BBDM23] Anurag Bishnoi, Simona Boyadzhiyska, Shagnik Das, and Tamás Mészáros. "Subspace coverings with multiplicities". In: *Combinatorics, Probability and Computing* 32.5 (2023), pp. 782–795. DOI: 10.1017/S0963548323000123 (cit. on p. 6).

[CHS20] John Chiarelli, Pooya Hatami, and Michael Saks. "An Asymptotically Tight Bound on the Number of Relevant Variables in a Bounded Degree Boolean function". In: *Combinatorica* 40.2 (Apr. 2020), pp. 237–244. ISSN: 0209-9683. DOI: 10.1007/s00493-019-4136-7. URL: https://doi.org/10.1007/s00493-019-4136-7 (cit. on pp. 7, 31).

[CH20] Alexander Clifton and Hao Huang. "On almost k-covers of hypercubes". In: *Combinatorica* 40.4 (2020), pp. 511–526 (cit. on p. 6).

[DDGKS17] Roee David, Irit Dinur, Elazar Goldenberg, Guy Kindler, and Igor Shinkar. "Direct Sum Testing". In: *SIAM Journal on Computing* 46.4 (2017), pp. 1336–1369. DOI: 10.1137/16M1061655 (cit. on p. 3).

[DL78] Richard A. DeMillo and Richard J. Lipton. "A probabilistic remark on algebraic program testing". In: *Information Processing Letters* 7.4 (1978), pp. 193–195. DOI: 10.1016/0020-0190(78)90067-4 (cit. on p. 3).

[DS96] Persi Diaconis and Laurent Saloff-Coste. "Logarithmic Sobolev inequalities for finite Markov chains". In: *The Annals of Applied Probability* 6.3 (1996), pp. 695–750 (cit. on pp. 7, 16, 30).

[DFH17] Irit Dinur, Yuval Filmus, and Prahladh Harsha. "Agreement tests on graphs and hypergraphs". In: *Electron. Colloquium Comput. Complex.* TR17 (2017). URL: https://api.semanticscholar.org/CorpusID:452567 (cit. on p. 4).

[DKSS13] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. "Extensions to the Method of Multiplicities, with Applications to Kakeya Sets and Mergers". In: *SIAM J. Comput.* 42.6 (2013), pp. 2305–2328. DOI: 10.1137/100783704. URL: https://doi.org/10.1137/100783704 (cit. on p. 3).

[Fil16] Yuval Filmus. "An Orthogonal Basis for Functions over a Slice of the Boolean Hypercube". In: *The Electronic Journal of Combinatorics* 23 (2016), p. 1. DOI: https://doi.org/10.37236/4567 (cit. on pp. 3, 5, 6, 15, 16).

[FI19] Yuval Filmus and Ferdinand Ihringer. "Boolean constant degree functions on the slice are juntas". In: *Discrete Mathematics* 342.12 (2019), p. 111614. ISSN: 0012-365X. DOI: https://doi.org/10.1016/j.disc.2019.111614. URL: https://www.sciencedirect.com/science/article/pii/S0012365X19302766 (cit. on pp. 3, 5, 7, 16, 30, 31).

[FKMW18] Yuval Filmus, Guy Kindler, Elchanan Mossel, and Karl Wimmer. "Invariance Principle on the Slice". In: *ACM Trans. Comput. Theory* 10.3 (2018), 11:1–11:37. DOI: 10.1145/3186590. URL: https://doi.org/10.1145/3186590 (cit. on p. 3).

[GRS23] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential Coding Theory (Book draft)*. 2023. URL: http://www.cse.buffalo.edu/atri/courses/coding-theory/book (cit. on p. 3).

[Hoe94] Wassily Hoeffding. "Probability inequalities for sums of bounded random variables". In: *The collected works of Wassily Hoeffding* (1994), pp. 409–426 (cit. on p. 13).

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. "Expander Graphs and their Applications". In: *Bulletin of the American Mathematical Society* 43 (2006), pp. 439–561 (cit. on p. 8).

[HJ91] Roger A Horn and Charles R Johnson. "Topics in matrix analysis, 1991". In: *Cambridge University Presss, Cambridge* 37 (1991), p. 39 (cit. on p. 12).

[HRRY19] Pavel Hrubes, Sivaramakrishnan Natarajan Ramamoorthy, Anup Rao, and Amir Yehudayoff. "Lower Bounds on Balancing Sets and Depth-2 Threshold Circuits". In: *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*. Ed. by Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi. Vol. 132. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, 72:1–72:14. DOI: 10.4230/LIPICS.ICALP.2019.72. URL: https://doi.org/10.4230/LIPIcs.ICALP.2019.72 (cit. on p. 3).

[KLMZ24] Gil Kalai, Noam Lifshitz, Dor Minzer, and Tamar Ziegler. *A Dense Model Theorem for the Boolean Slice*. To appear in the Proceedings of the 65th IEEE Symposium of Foundations of Computer Science (FOCS) 2024, Chicago, USA. 2024. arXiv: 2402.05217 [math.CO]. URL: https://arxiv.org/abs/2402.05217 (cit. on pp. 3, 6).

[LY98] Tzong-Yow Lee and Horng-Tzer Yau. "Logarithmic Sobolev inequality for some models of random walks". In: *The Annals of Probability* 26.4 (1998), pp. 1855–1873 (cit. on pp. 7, 16, 30).

[Luc78] Edouard Lucas. "Théorie des Fonctions Numériques Simplement Périodiques". In: *American Journal of Mathematics* 1.2 (1878), pp. 184–196. ISSN: 00029327, 10806377. URL: http://www.jstor.org/stable/2369308 (visited on 11/04/2024) (cit. on p. 8).

[NS92] Noam Nisan and Mario Szegedy. "On the degree of boolean functions as real polynomials". In: *Computational Complexity* 4 (1992), pp. 301–313. URL: https://api.semanticscholar.org/CorpusID:6919144 (cit. on pp. 3, 5, 7, 30, 31).

[OW13] Ryan O'Donnell and Karl Wimmer. "KKL, Kruskal-Katona, and Monotone Nets". In: *SIAM J. Comput.* 42.6 (2013), pp. 2375–2399. DOI: 10.1137/100787325. URL: https://doi.org/10.1137/100787325 (cit. on p. 3).

[Ore22] Øystein Ore. "Über höhere kongruenzen". In: *Norsk Mat. Forenings Skrifter* 1.7 (1922), p. 15 (cit. on p. 3).

[PS90] Ramamohan Paturi and Michael E. Saks. "On Threshold Circuits for Parity". In: *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, USA, October 22-24, 1990, Volume I*. IEEE Computer Society, 1990, pp. 397–404. DOI: 10.1109/FSCS.1990.89559. URL: https://doi.org/10.1109/FSCS.1990.89559 (cit. on p. 3).

[RV89] Michael O Rabin and Vijay V Vazirani. "Maximum matchings in general graphs through randomization". In: *Journal of Algorithms* 10.4 (1989), pp. 557–567. ISSN: 0196-6774. DOI: https://doi.org/10.1016/0196-6774(89)90005-9. URL: https://www.sciencedirect.com/science/article/pii/0196677489900059 (cit. on p. 3).

[SS08] Shubhangi Saraf and Madhu Sudan. "An improved lower bound on the size of Kakeya sets over finite fields". In: *Analysis and PDE* 1.3 (2008), pp. 375–379. DOI: 10.2140/apde.2008.1.375. URL: https://doi.org/10.2140/apde.2008.1.375 (cit. on p. 3).

[SW22] Lisa Sauermann and Yuval Wigderson. "Polynomials that vanish to high order on most of the hypercube". In: *Journal of the London Mathematical Society* 106.3 (2022), pp. 2379–2402. DOI: https://doi.org/10.1112/jlms.12637. eprint: https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/jlms.12637. URL: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/jlms.12637 (cit. on p. 6).

[Sch80] Jacob T. Schwartz. "Fast Probabilistic Algorithms for Verification of Polynomial Identities". In: *J. ACM* 27.4 (1980), pp. 701–717. DOI: 10.1145/322217.322225 (cit. on p. 3).

[TZ12] Terence Tao and Tamar Ziegler. "The inverse conjecture for the Gowers norm over finite fields in low characteristic". In: *Annals of Combinatorics* 16.1 (2012), pp. 121–188 (cit. on p. 3).

[Wil90] Richard M. Wilson. "A Diagonal Form for the Incidence Matrices of t-Subsets vs.k-Subsets". In: *European Journal of Combinatorics* 11.6 (1990), pp. 609–615. ISSN: 0195-6698. DOI: https://doi.org/10.1016/S0195-6698(13)80046-7. URL: https://www.sciencedirect.com/science/article/pii/S0195669813800467 (cit. on pp. 5, 22, 41).

[Wim14] Karl Wimmer. "Low Influence Functions over Slices of the Boolean Hypercube Depend on Few Coordinates". In: *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*. IEEE Computer Society, 2014, pp. 120–131. DOI: 10.1109/CCC.2014.20. URL: https://doi.org/10.1109/CCC.2014.20 (cit. on p. 3).

[Zip79] Richard Zippel. "Probabilistic algorithms for sparse polynomials". In: *Symbolic and Algebraic Computation*. Springer Berlin Heidelberg, 1979, pp. 216–226. DOI: 10.1007/3-540-09519-5_73 (cit. on p. 3).

A Appendix

A.1 Proofs of Claim 3.7 and Claim 3.8

Proof of Claim 3.7. By the definition of the weight function w of the generators, we have

$$\mu'_{\varnothing} = \sum_{\mathbf{y}\in\{0,1\}^{2m}_{e}} w(\mathbf{y}) = \sum_{d=0}^{m} \binom{2m}{2d} \cdot \frac{1}{2^m \cdot \binom{m}{d}} = \sum_{d=0}^{m} \frac{\binom{2m}{2d}}{\binom{m}{d}^2} \cdot \frac{\binom{m}{d}}{2^m} \leqslant \max_{d\in[0..m]} \left(\frac{\binom{2m}{2d}}{\binom{m}{d}^2} \right) \cdot \sum_{d=0}^{m} \frac{\binom{m}{d}}{2^m}$$

$$= \max_{d\in[0..m]} \left(\frac{\binom{2m}{2d}}{\binom{m}{d}^2} \right).$$

Now, for every $d \in [0..m]$,

$$\frac{\binom{2m}{2d}}{\binom{m}{d}^2} = \frac{(2m)!d!^2(m-d)!^2}{(2d)!(2m-2d)!m!^2} = \frac{\binom{2m}{m}}{\binom{2d}{d}\binom{2m-2d}{m-d}} = O\left(\frac{2^{2m}}{\sqrt{m}} \cdot \frac{\sqrt{d}}{2^{2d}} \cdot \frac{\sqrt{m-d}}{2^{2m-2d}}\right) = O\left(\sqrt{\frac{d(m-d)}{m}}\right) \leqslant O(\sqrt{m}),$$

where the last inequality uses the AM-GM inequality. Therefore, $0 < \mu'_{\varnothing} \leqslant O(\sqrt{m})$. ■

Finally, we prove Claim 3.8.

Proof of Claim 3.8. Letting $B := \{d \in [0..m] \mid |2d - m| > \sqrt{50m \log m}\}$, we can explicitly express the probability as

$$\Pr_{\mathbf{x}\sim\mathcal{D}}[||\mathbf{x}| - m| > \sqrt{50m\log m}] = \sum_{d\in B} \binom{2m}{2d} \cdot \frac{1}{2^m \binom{m}{d}} = \sum_{d\in B} \frac{\binom{2m}{2d}}{\binom{m}{d}^2} \cdot \frac{\binom{m}{d}}{2^m} \leqslant O(\sqrt{m}) \cdot \sum_{d\in B} \frac{\binom{m}{d}}{2^m},$$

by using the bound $\frac{\binom{2m}{2d}}{\binom{m}{d}^2} \leqslant O(\sqrt{m})$ from the proof of Claim 3.7. To bound the second factor $\sum_{d\in B} \frac{\binom{m}{d}}{2^m}$, we use a Chernoff bound for the sum of m i.i.d. copies of a uniformly random Boolean variable. In particular, we get $\sum_{d\in B} \frac{\binom{m}{d}}{2^m} \leqslant O(1/m^3)$. Hence $\Pr_{\mathbf{x}\sim\mathcal{D}}[||x| - m| > \sqrt{50m\log m}] \leqslant O(1/m^2)$. ■

A.2 Proof of Claim 3.15

Claim 3.15 (Upper bound on probability of (self) good matchings)**.** *Consider the complete bipartite graph $K_{n/2,n/2}$ on $L \bigcup R$ where $L = \{1, 3, \ldots, n-1\}$ and $R = \{2, 4, \ldots, n\}$. Let $\tau = n^{\delta}$ for a sufficiently small $\delta > 0$. Then,*

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a } t\text{-good matching}] \leqslant \frac{1}{n^{\Omega(t)}}, \qquad \text{for all } t \leqslant \tau,$$

where the above probability is over the choice of a uniformly random matching $\mathcal{M}$. Also,

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a } t\text{-self good matching}] \leqslant \frac{1}{n^{\Omega(t)}}, \qquad \text{for all } t > \tau,$$

where the above probability is over the choice of a uniformly random matching $\mathcal{M}$.

Proof of Claim 3.15. Let us first how to describe a good matching, or in other words, how to generate a good matching. We first choose a subset $T \subseteq \mathbf{u}^{-1}\{1\}$ of size $0 \leqslant k \leqslant t$ which will be matched outside the set $\{2, 4, \ldots, 2t\}$. To satisfy the good matching condition, it enforces that for every $i \in T$, $\mathcal{M}(2i) \in \{1, 3, \ldots, 2t-1\} \setminus T$. We choose remaining k vertices from $\mathbf{u}^{-1}\{0\}$ which will be matched outside $\{1, 3, \ldots, 2t-1\}$. This also enforces that $(t-k) \geqslant k \Rightarrow k \leqslant t/2$. Finally, we also account for the possible matching. This gives us:

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a good matching}] = \sum_{k=0}^{t/2} \frac{\binom{t}{k}\binom{t-k}{k}(t-k)!(n/2-t+k)!}{(n/2)!} \tag{7}$$

We upper bound $\binom{t}{k}$, $\binom{t-k}{k}$, and $(t-k)!$ by t^t for all $0 \leqslant k \leqslant t/2$. Since $t \leqslant n^\delta$, we can upper bound $(n/2-t+k)$ by $(n/2-t/2)!$ for all $0 \leqslant k \leqslant t/2$. Using these we get,

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a good matching}] \leqslant \frac{t^{3t}}{\binom{n/2}{t/2}(t/2)!}$$

Employing the standard binomial estimate that $\binom{N}{K} \geqslant (N/K)^K$ and Stirling's approximation, we get,

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a good matching}] \leqslant \frac{t^{3t}}{n^{t/2}} = \frac{1}{n^{\Omega(t)}}$$

Now we upper bound the probability that a random matching is a t-self good matching. Observe that if a matching $\mathcal{M}$ is a t-self good matching, then it implies there exists a subset $\widetilde{T} \subset [t]$ of size $t/2$ such that $\mathcal{M}(2i-1) = 2i$ for $i \in \widetilde{T}$. We can have an arbitrary matching in the remaining $(n/2-t/2)$ vertices in both L and R. This gives us:

$$\Pr_{\mathcal{M}}[\mathcal{M} \text{ is a } t\text{-self good matching}] \leqslant \frac{\binom{t}{t/2}(n/2-t/2)!}{(n/2)!} \leqslant \frac{1}{n^{\Omega(t)}}$$

This finishes the proof of Claim 3.15. ■

A.3 Proof of Lemma 4.4

Lemma 4.4 (Number of degree-d polynomials on the slice)**.** *[Wil90]. For every degree parameter $d \in \mathbb{N}$ and for every slice parameter k such that $d \leqslant \min\{k, n-k\}$, the number of distinct degree-d polynomial functions on $\{0,1\}^n_k$ is at least $q^{\binom{n}{d}}$.*

The proof is by induction on the parameter $t = (k-d)\cdot(n-k-d)$.

The base case of the induction corresponds to the case when $t = 0$, i.e. $d = \min\{k, n-k\}$. In this case, we want to show that any function $f : \{0,1\}^n_k \to \mathbb{Z}_q$ is a degree-d polynomial. To show this, it suffices to show that any δ-function on $\{0,1\}^n_k$ can be written as a polynomial of degree at most d.

Consider without loss of generality the δ-function at the point $\mathbf{a} = 1^k 0^{n-k}$. If $d = k$, then the monomial $x_1 \cdots x_d$ computes exactly this δ-function. On the other hand, if $d = n-k$, we can instead use the polyomial $(1-x_1)\cdots(1-x_d)$. In either case, we are done. This proves the base case, i.e. that $|\mathcal{P}_d(n,k,\mathbb{Z}_q)| \geqslant q^{\binom{n}{d}}$ in this case.

For the induction, assume that $d < \min\{k, n-k\}$. We claim that

$$|\mathcal{P}_d(n,k,\mathbb{Z}_q)| \geqslant |\mathcal{P}_d(n-1,k,\mathbb{Z}_q) \times \mathcal{P}_{d-1}(n-1,k-1,\mathbb{Z}_q)| \tag{8}$$

which immediately implies the inductive case using the induction hypothesis[9], as we have

$$|\mathcal{P}_d(n-1,k,\mathbb{Z}_q) \times \mathcal{P}_{d-1}(n-1,k-1,\mathbb{Z}_q)| \geqslant |\mathcal{P}_d(n-1,k,\mathbb{Z}_q)| \cdot |\mathcal{P}_{d-1}(n-1,k-1,\mathbb{Z}_q)|$$

[9]Note that the induction hypothesis is applicable as either $k-d$ or $n-k-d$ drops by 1 while the other remains the same.

$$\geqslant q^{\binom{n-1}{d}} \cdot q^{\binom{n-1}{d-1}} = q^{\binom{n}{d}}.$$

To prove Equation (8), we give an injection ι from the set $\mathcal{P}_d(n-1,k,\mathbb{Z}_q) \times \mathcal{P}_{d-1}(n-1,k-1,\mathbb{Z}_q)$ to the set $\mathcal{P}_d(n,k,\mathbb{Z}_q)$. For each function in $\mathcal{P}_d(n-1,k,\mathbb{Z}_q)$, we fix arbitrarily a polynomial of degree at most d representing this function, and do the same for functions in $\mathcal{P}_{d-1}(n-1,k-1,\mathbb{Z}_q)$ with the degree parameter being $d-1$.

Let $(P,Q) \in \mathcal{P}_d(n-1,k,\mathbb{Z}_q) \times \mathcal{P}_{d-1}(n-1,k-1,\mathbb{Z}_q)$ be the chosen polynomial representations of a pair of functions in the corresponding sets. Define $\iota(P,Q)$ to be

$$R(x_1,\ldots,x_n) = P(x_1,\ldots,x_{n-1}) + x_n \cdot Q(x_1,\ldots,x_{n-1}).$$

The map is injective because the function computed by P (and hence the underlying polynomial which is fixed by the function) can be obtained by restricting R to the points where $x_n = 0$. Further, the function Q can be obtained by evaluating R at the points where $x_n = 1$ and subtracting the value of the polynomial P evaluated at the same point.

This proves Equation (8) and concludes the inductive case.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il