Electronic Colloquium on Computational Complexity, Report No. 78 (2018)

Small Set Expansion in The Johnson Graph

Subhash Khot * Dor Minzer † Dana Moshkovitz ‡ Muli Safra §

Abstract

This paper studies expansion properties of the (generalized) Johnson Graph. For natural numbers $t < \ell < k$, the nodes of the graph are sets of size ℓ in a universe of size k. Two sets are connected if their intersection is of size t. The Johnson graph arises often in combinatorics and theoretical computer science: it represents a "slice" of the noisy hypercube, and it is the graph that underlies direct product tests as well as a candidate hard unique game.

We prove that any small set of vertices in the graph either has near perfect edge expansion or is not pseudorandom. Here "not pseudorandom" means that the set becomes denser when conditioning on containing a small set of elements. In other words, we show that slices of the noisy hypercube – while not small set expanders like the noisy hypercube – only have non-expanding small sets of a certain simple structure.

This paper is related to a recent line of work establishing the 2-to-2 Theorem in PCP. The result was motivated, in part, by [7] which hypothesized and made partial progress on similar result for the Grassmann graph. In turn, our result for the Johnson graphs served as a crucial step towards the full result for the Grassmann graphs completed subsequently in [20].

1 Introduction

Expanding graphs are ubiquitous in mathematics, theoretical computer science, information theory and other areas of study (See, e.g., [17]). They are well-connected graphs that are robust to deletion of edges, and random walks on them mix rapidly. In an expander, for every set S of at most half of the nodes, many of S's neighbors are outside S. Formally, we define edge expansion as follows:

Definition 1.1 (Edge expansion)**.** *Let $G = (U, E)$ be a d-regular graph, and let $S \subseteq U$ be a set of vertices. The expansion of S is defined as*

$$\Phi(S) = \frac{\left|E(S, \overline{S})\right|}{d\,|S|},$$

where $E(S, \overline{S})$ is the set of edges from S outside it.

In recent years there has been considerable interest in *small set expanders*, which are graphs where only sets containing a small fraction of the nodes are required to expand. The interest in small set expansion was motivated by its connection to the Unique Games Conjecture [23, 26] and small set expanders were used in SDP integrality gaps [23], lower bounds on metric embeddings [23] and algorithms [1].

*Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. Research supported by NSF CCF-1422159, Simons Collaboration on Algorithms and Geometry, and Simons Investigator Award.

†School of Computer Science, Tel Aviv University.

‡School of Computer Science, University of Texas at Austin. This material is based upon work supported by the National Science Foundation under grants number 1218547 and 1648712.

§School of Computer Science, Tel Aviv University.

ISSN 1433-8092

An important example of a small set expander (that is not an expander) is the noisy Boolean hypercube. Its nodes are the vectors in $\{0,1\}^k$, or alternatively sets in a universe of size k. Two sets are connected, roughly, if their intersection is of size t, typically close to the size of the sets[1]. In the noisy hypercube small sets have near perfect expansion, and this property is key to many of the aforementioned results [23]. The reader is referred to [25] for information about the hypercube, noise and the rich theory underlying them.

1.1 Our Contribution

In this work we focus on slices of the hypercube, also known as Johnson graphs. For parameters $t < \ell < k$ we define the Johnson graph $J(k, \ell, t)$ to be the graph whose nodes are sets of size ℓ in a universe of size k. There is an edge between two sets if their intersection is of size t. Typically ℓ is much smaller than k and t is close to ℓ. Johnson graphs are the subject of considerable interest recently (See, e.g., [10, 12, 13]). In combinatorics they are key to the study of sharp thresholds of graph properties [14]. In theoretical computer science they underlie direct product tests [15, 5, 19, 9], and a recent candidate hard unique game [22].

Johnson graphs are not small set expanders as we demonstrate next. Consider the family S_x of all sets containing a certain element x in the size-k universe ("zoom in"). This is a very small family (The fraction of S_x is roughly ℓ/k, which is small for large k and much smaller ℓ), yet many of the edges with one endpoint in S_x have their other endpoint in S_x as well (roughly t/ℓ fraction, which is large for t close to ℓ). In this paper we show that essentially only families like S_x are small and non-expanding in Johnson graphs. Other families have near-perfect expansion, that is, almost all of their neighbors are outside them. We think of a set being "like" an S_x if it becomes substantially denser when conditioning on containment of one, or a few, elements x. We define a "pseudo-random" set as one that is unlike any S_x:

Definition 1.2 (Pseudo-randomness on the Generalized Johnson Graph)**.** *Let $J(k, \ell, t)$ be a Johnson Graph, and let $r \in \mathbb{N}, \varepsilon > 0$ where $r \leqslant \ell$. We say that a set of vertices S is (r, ε) pseudo-random if for every $R \subseteq [k]$ of size at most r,*

$$\left| \Pr_{A \supseteq R}[A \in S] - \Pr_{A}[A \in S] \right| \leqslant \varepsilon.$$

Our main theorem is that small sets of nodes in the Johnson graph that are pseudorandom have near perfect expansion.

Theorem 1.3 (Main Theorem)**.** *For every $0 < \alpha < 1$ and $0 < \eta < 1$ there exist $r \in \mathbb{N}$, $\varepsilon, \delta > 0$, such that for large enough $k > \ell$ the following holds. If S is a set of vertices of $J(k, \ell, \alpha\ell)$ of density at most δ such that $\Phi(S) \leqslant 1 - \eta$, then S is not (r, ε) pseudo-random.*

Theorem 2.16 in the body of the paper gives a concrete dependence of the parameters on α and η. Theorem 2.17 gives an improved dependence of ε at the expense of a worse dependence of the other parameters.

Theorem 1.3 is established via spectral analysis on the adjacency matrix of the Johnson graph. We show a relation between pseudo-randomness of a set and its correlation with eigenspaces whose eigenvalue is non-negligible (formally stated in Theorem 2.15). Namely, if a function is pseudo-random, then its weight on the low-level eigenspaces is small.

We remark that the conclusion of Theorem 1.3 is in fact slightly stronger. Namely, we may conclude that there is non-negligible fraction of zoom-ins that increases the density by at least ε. This stronger conclusion

[1]One typically considers the uniform distribution over the nodes, so mostly nodes are sets of size roughly $\ell = k/2$. One considers the following probability distribution over neighbors $y \in \{0,1\}^k$ of a vector $x \in \{0,1\}^k$: for each $1 \leqslant i \leqslant k$ re-sample x_i with probability δ to get y_i, and set $y_i = x_i$ with probability $1 - \delta$. Hence, edges mostly correspond to intersection of size roughly $t = (1 - \delta)\ell$.

follows from the current one: we find a zoom-in set that increases the density by more than ε iteratively. As long as the fraction of vertices we have covered with our zoom-in sets is $o(\delta)$, we may randomized the set S on the zoom-in sets so far. After this randomization, the density of each zoom-in set we have found so far is $\delta + o(\delta)$ and the expansion of S is bounded away from 1, and in particular we may find an additional zoom-in set.

1.2 Related Work

This paper is related to a recent line of work establishing the 2-to-2 Theorem in PCP [20, 6, 7, 21]. The heart of this line of work is a theorem similar to Theorem 1.3 for the Grassman graph rather than the Johnson graph, which was subsequently established in [21]. The nodes of the Grassman graphs are ℓ-dimensional linear subspaces rather than size-ℓ sets. Edges correspond to largely intersecting linear subspaces. The structure of the small non-expanding sets in the Grassman graph is more complicated than that of the Johnson graph, and our Theorem 1.3 was a crucial step towards the analogous theorem about the Grassman graph [21]. Our result was motivated, in part, by [7] which hypothesized and made partial progress on the expansion of Grassmann graphs. The current work builds on their technique and further develops it.

To prove Theorem 1.3, one has to analyze expectations of the form

$$\mathbb{E}\left[f(x_1,\ldots,x_r)f(y_1,\ldots,y_r)f(z_1,\ldots,z_r)f(w_1,\ldots,w_r)\right], \tag{1}$$

for a certain function of interest f emerging from the spectral decomposition on the Johnson Graph, where the expectation is taken uniformly over $x, y, z, w \in [k]^r$ that satisfy a predetermined set of equalities of the form $a = b$ for $a, b \in \{x_i, y_i, z_i, w_i \mid i = 1, \ldots, r\}$. We call such expectations four-wise correlations, since it is an expectation of the product of four values of f on correlated inputs.

For the analog in the Grassmann Graph, one has to study the four-wise correlations

$$\mathbb{E}\left[f(x_1,\ldots,x_r)f(y_1,\ldots,y_r)f(z_1,\ldots,z_r)f(w_1,\ldots,w_r)\right], \tag{2}$$

where now the expectation is uniform over $x_i, y_i, z_i, w_i \in \mathbb{F}_2^k$, that satisfy a predetermined set of linear equations in x_i, y_i, w_i, z_i. In particular, this set of equations could contain equalities as before, say $x_1 = y_2$.

In [7], the authors could analyze the expectations of the form (2) for $r = 1, 2$. For $r = 2$, they use brute force analysis to enumerate over all possible linear equation systems that determine the constraints among the variables, and upper bound each such configuration separately. The analysis uses a combination of fourier analysis on f and the Cauchy-Schwarz inequality, along with the (additional) notion of zoom-outs. As r increases, the number of configurations grows quickly and a case by case analysis becomes infeasible. It was very unclear if a more systematic approach is possible that is able to deal with larger r.

In this paper, we present a systematic approach for the Johnson Graph (wherein fourier analysis on f and the notion of zoom-outs is not needed).

It turns out that (as far as this analysis is concerned), the Johnson analysis is a special but crucial case of the Grassmann analysis. Indeed, in the Grassmann analysis, if the only linear dependencies are equalities, then the analysis, as far as its high level structure/strategy is concerned, reduces to the Johnson analysis. This turned out to be an insightful step in completing the Grassmann analysis in [21].

We also show how to analyze higher than four-wise correlations in the Johnson Graph. This leads to improved quantitative results in Theorem 1.3 and in the main technical result used to achieve it, Theorem 2.15 and its improved form Theorem 2.17.

2 Preliminaries

In this section we present the necessary background on the Johnson Graph. We shall use properties of it that are analogous to properties from [7] about the Grassmann Graph. We often omit proofs if they are nearly identical.

Definition 2.1. *Let $t < \ell < k$ be integers. The Generalized Johnson Graph $J(k, \ell, t)$ is defined as follows: the vertex set is $\binom{[k]}{\ell}$, the set of all subsets of $[k]$ of size ℓ. Two vertices $A, B \subseteq [k]$ are connected by an edge if and only if $|A \cap B| = t$.*

Abusing notation, we denote the normalized adjacency operator of the generalized Johnson Graph by $J(k, \ell, t)$. We think of it as operating on real valued functions $F\colon \binom{[k]}{\ell} \to \mathbb{R}$.

2.1 Fourier Analysis on the generalized Johnson Graph

Any graph, and in particular $J(k, \ell, t)$ induces a spectral decomposition of real-valued functions on it. In this section we introduce some properties of this decomposition. It has a similar structure to the spectral decomposition on the Grassmann Graph [7].

We endow the space of real-valued functions on $J(k, \ell, t)$ with the inner product

$$\langle F, G\rangle = \mathop{\mathbb{E}}_{A \in \binom{[k]}{\ell}} [F[A]G[A]]$$

for any $F, G\colon \binom{[k]}{\ell} \to \mathbb{R}$.

2.2 Level Functions

Definition 2.2. *Let $t < \ell < k$ be integers. For any $i = 0, ..., \ell$ we define the space spanned by the first i levels $J_{\leqslant i}$ as follows. $F \in J_{\leqslant i}$ if and only if there exists $f\colon \binom{[k]}{i} \to \mathbb{R}$ such that*

$$F[A] = \sum_{I \subseteq A, |I| = i} f(I).$$

One can easily verify that each $J_{\leqslant i}$ is a linear subspace, and that $J_{\leqslant \ell}$ contains all real-valued functions on $J(k, \ell, t)$. Furthermore, we have that $J_{\leqslant i} \subseteq J_{\leqslant i+1}$.

Definition 2.3. *We define the space of level i functions by $J_{=i} = J_{\leqslant i} \cap J^{\perp}_{\leqslant i-1}$. In words, it is the space of all functions from $J_{\leqslant i}$ perpendicular to $J_{\leqslant i-1}$.*

It follows by the definition that the space of real-valued functions on $J(k, \ell, t)$ can be written as $J_{=0} \oplus J_{=1} \oplus ... \oplus J_{=\ell}$.

Definition 2.4. *Let $t < \ell < k$ be integers. For $i = 0, ..., t$ define*

$$\lambda_i(k, \ell, t) = \frac{\binom{t}{i}}{\binom{\ell}{i}}.$$

For $i > t$, define

$$\lambda_i(k, \ell, t) = 0.$$

Definition 2.5. *Let $j < i < k$ be integers and $f\colon \binom{[k]}{i} \to \mathbb{R}$. For $J \subseteq [k]$, $|J| = j$ we denote*

$$\mu_J(f) = \mathop{\mathbb{E}}_{I \supseteq J}[f(I)].$$

Lemma 2.6. *Suppose $k \geqslant \ell^{8\ell}$, $\ell \geqslant 2$, and let $F \in J_{\leqslant i}$. Then $F \in J_{=i}$ iff for every $R \subseteq [k]$, $|R| < i$ we have that $\mu_R(f) = 0$.*

Proof. The proof is similar to the proof of [7, Lemma 2.19]. □

Theorem 2.7. *Let $t < \ell < k$ be integers. For any $i = 0, ..., \ell$, and $F \in J_{=i}$ given by*

$$F[A] = \sum_{I \subseteq A,\ |I|=i} f(I)$$

we have

$$\|J(k,\ell,t)F - \lambda_i(k,\ell,t)F\|_\infty \leqslant \frac{2^{4\ell}}{k}\|f\|_\infty.$$

In words, F is a near eigenvector of $J(k,\ell,t)$ of eigenvalue $\lambda_i(k,\ell,t)$.

Proof. Let F be such function and fix a vertex A in $J(k,\ell,t)$. Then

$$J(k,\ell,t)F[A] = \mathop{\mathbb{E}}_{B:|B\cap A|=t}[F[B]] = \mathop{\mathbb{E}}_{B:|B\cap A|=t}\left[\sum_{I\subseteq B,\ |I|=i} f(I)\right] = \mathop{\mathbb{E}}_{B:|B\cap A|=t}\left[\sum_{r=\max(0,t+i-\ell)}^{\min(t,i)} \sum_{I\subseteq B, |A\cap I|=r} f(I)\right].$$

We consider different values of i separately, starting with $i \leqslant t$. In this case, we separate the summand for $r = i$ to get

$$J(k,\ell,t)F[A] = \mathop{\mathbb{E}}_{B:|B\cap A|=t}\left[\sum_{I\subseteq A\cap B} f(I)\right] + \mathop{\mathbb{E}}_{B:|B\cap A|=t}\left[\sum_{r=\max(0,t+i-\ell)}^{i-1} \sum_{I\subseteq B, |A\cap I|=r} f(I)\right]. \qquad (3)$$

The first sum in (3) equals

$$\mathop{\mathbb{E}}_{B:|B\cap A|=t}\left[\binom{t}{i}\mathop{\mathbb{E}}_{I\subseteq A\cap B,\ |I|=i}[f(I)]\right] = \binom{t}{i}\mathop{\mathbb{E}}_{I\subseteq A}[f(I)] = \frac{\binom{t}{i}}{\binom{\ell}{i}}F[A].$$

The second sum in (3) is

$$\mathop{\mathbb{E}}_{B:|B\cap A|=t}\left[\sum_{r=\max(0,t+i-\ell)}^{i-1}\binom{t}{r}\binom{\ell-t}{i-r}\mathop{\mathbb{E}}_{I\subseteq B, |A\cap I|=r}[f(I)]\right] = \sum_{r=\max(0,t+i-\ell)}^{i-1}\binom{t}{r}\binom{\ell-t}{i-r}\mathop{\mathbb{E}}_{I:|I\cap A|=r}[f(I)].$$

We show that the last expectation is negligible. Fix $r \leqslant i-1$, then

$$\mathop{\mathbb{E}}_{I:|I\cap A|=r}[f(I)] = \mathop{\mathbb{E}}_{R\subseteq A,\ |R|=r}\left[\mathop{\mathbb{E}}_{I:I\cap A=R,\ |I|=i}[f(I)]\right]$$

For a fixed R, denote by $\mathcal{D}_1$ the uniform distribution over I where $I \cap A = R$, and by $\mathcal{D}_2$ uniform distribution over I where $I \supseteq R$. Note that these distributions are $\frac{i\ell}{k}$ close. Hence

$$\left| \mathop{\mathbb{E}}_{I:I\cap A=R,\ |I|=i} [f(I)] - \mu_R(f) \right| \leqslant \frac{i\ell}{k} \|f\|_\infty,$$

and since $\mu_R(f) = 0$ (Lemma 2.6) we conclude that the second sum in (3) is upper bounded by

$$\sum_{r=\max(0,t+i-\ell)}^{i-1} \binom{t}{r}\binom{\ell-t}{i-r} \left| \mathop{\mathbb{E}}_{I:|I\cap A|=r} [f(I)] \right| \leqslant i2^t 2^\ell \frac{i\ell}{k} \|f\|_\infty \leqslant \frac{2^{4\ell}}{k} \|f\|_\infty.$$

Next, we prove for $i > t$. In this case we have

$$J(k,\ell,t)F[A] = \mathop{\mathbb{E}}_{B:|B\cap A|=t} \left[\sum_{r=\max(0,t+i-\ell)}^{t} \sum_{I\subseteq B, |A\cap I|=r} f(I) \right].$$

The argument is identical to the analysis of the error term in the previous case, and is hence omitted. □

Remark 2.8. *With a more careful analysis, one can prove that $J_{=i}$ is actually an eigenspace of $J(k,\ell,t)$ with eigenvalue nearly λ_i.*

2.3 Decomposition

Recall that we have seen that the space of all real-valued functions on $\binom{[k]}{\ell}$ can be decomposed as $J_{=0} \oplus J_{=1} \oplus .... \oplus J_{=\ell}$. For any function $F\colon \binom{[k]}{\ell} \to \mathbb{R}$, we denote this decomposition by $F = F_{=0} + F_{=1} + ... + F_{=\ell}$ where $F_{=i} \in J_{=i}$. We also define $f_{=i}\colon \binom{[k]}{i} \to \mathbb{R}$ to be a function that satisfies

$$F_{=i}[A] = \sum_{I\subseteq A,\ |I|=i} f_{=i}(I).$$

It is easy to verify that $F_{=0} \equiv \mu(F)$. For large i, the functions $F_{=i}$ are harder to compute and we approximate them instead.

Definition 2.9. *Let $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ be a function, and let $i \in \{0, 1, ..., \ell\}$. We define the weight of F on level i to be*

$$W_{=i}[F] \stackrel{def}{=} \langle F_{=i}, F_{=i} \rangle.$$

2.4 Approximate Decomposition

As is the case in the Grassmann Graph, computing the exact decomposition is not easy, so we work with an approximate decomposition instead. Given a function $F\colon \binom{[k]}{\ell} \to \mathbb{R}$, define $f_{\approx 0} \equiv \mu(F)$. Inductively once $f_{\approx j}$ have been defined for all $j < i$, we define $f_{\approx i}\colon \binom{[k]}{i} \to \mathbb{R}$ by

$$f_{\approx i}(I) \stackrel{def}{=} \mu_I(F) - \sum_{J\subset I} f_{\approx |J|}(J)$$

for all $I \in \binom{[k]}{i}$.

Define

$$F_{\approx i}[A] \stackrel{def}{=} \sum_{I\subseteq A,\ |I|=i} f_{\approx i}(I).$$

Basic Facts

The following theorem is an analog of a theorem from the Grassmann Graph. The proof is an adaptation of the proof therein and is hence omitted.

Theorem 2.10. *[Analog of [7, Theorem 2.25]] Let k, ℓ be integers such that $k \geqslant 2^{200\ell^3}$, and let $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ be a function. Then for all $i = 0, ..., \ell$,*

$$\|F_{=i} - F_{\approx i}\|_2^2 \leqslant \frac{2^{100\ell^3}}{k} \|F\|_\infty.$$

The following fact will also be used. Its an analog of a fact from the Grassmann Graph, and the proofs are similar.

Fact 2.11. *[Analog of [7, Lemma A.2]] Let $\ell \leqslant k$ be integers, and let $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ be a function. Then for all $i = 1, ..., \ell$, $0 \leqslant j < i$ and $J \subseteq [k]$ of size j,*

$$\left| \mathop{\mathbb{E}}_{I \supseteq J,\ |I| = i} [f_{\approx i}(I)] \right| \leqslant \frac{2^{10i^2}}{k} \|F\|_\infty.$$

Fact 2.12. *[Analog of [7, Claim A.1]] Let $\ell \leqslant k$ be integers, and let $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ be a function. Then for all $i = 0, ..., \ell$, $\|f_{\approx i}\|_\infty \leqslant 2^{i^2} \|F\|_\infty$.*

Note that $f_{\approx i}$ depends of course on the underlying function F. Often times the function F will be clear from the context, however sometimes we will be dealing with more than one function simultaneously. In this case we shall denote the function $f_{\approx i}$ by $f_{\approx i, F}$.

2.5 Restrictions

Given a function $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ and $X \subseteq [k]$ of size at most $\ell - 1$, we define the restricted function $F|_X\colon \binom{[k]\setminus X}{\ell - |X|} \to \mathbb{R}$ by

$$F|_X[A] = F[X \cup A].$$

The following lemma expresses level $i+1$ components of a function F as a function of the level i components of F and its restrictions.

Lemma 2.13. *Let $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ be function, $0 \leqslant i \leqslant \ell - 1$ and $x \in [k]$. Then for any $I \subseteq [k] \setminus \{x\}$ of size i,*

$$f_{\approx i+1, F}(I \cup \{x\}) = f_{\approx i, F|_{\{x\}}}(I) - f_{\approx i, F}(I).$$

Proof. The proof is by induction on i. For $i = 0$ the claim is obvious, since we have $I = \emptyset$ and both sides are easily seen to be equal to $\mu_{\{x\}}(F) - \mu(F)$.

Let $i > 0$, assume the claim for every $j < i$ and prove for i. Fix I, then by definition

$$f_{\approx i+1, F}(I \cup \{x\}) = \mu_{I \cup \{x\}}(F) - \sum_{J \subseteq I} f_{\approx |J|, F}(J) - \sum_{J \subsetneq I} f_{\approx |J|+1, F}(J \cup \{x\}). \tag{4}$$

Consider the second sum. Since we sum only over J strictly contained in I, we may apply the induction hypothesis to get that

$$f_{\approx |J|+1, F}(J \cup \{x\}) = f_{\approx |J|, F|_{\{x\}}}(J) - f_{\approx |J|, F}(J).$$

Plug it into (4) to get

$$\begin{aligned} f_{\approx i+1,F}(I\cup\{x\}) &= \mu_{I\cup\{x\}}(F) - \sum_{J\subseteq I} f_{\approx|J|,F}(J) - \sum_{J\subsetneq I}\left(f_{\approx|J|,F|_{\{x\}}}(J) - f_{\approx|J|,F}(J)\right) \\ &= \mu_I(F|_{\{x\}}) - \sum_{J\subsetneq I} f_{\approx|J|,F|_{\{x\}}}(J) - f_{\approx i,F}(I) \\ &= f_{\approx i,F|_{\{x\}}}(I) - f_{\approx i,F}(I), \end{aligned}$$

in the last equality we used the definition of $f_{\approx i,F|_{\{x\}}}$. □

Corollary 2.14. *Let $F\colon \binom{[k]}{\ell} \to \mathbb{R}$ be function, $0 \leqslant i \leqslant \ell - 1$ and $X \subseteq [k]$, $|X| < i$. Then for any $I \subseteq [k]\setminus X$ of size $j = i - |X|$,*

$$f_{\approx i,F}(I\cup X) = \sum_{Y\subseteq X} (-1)^{|Y|} f_{\approx j,F|_{X\setminus Y}}(I).$$

Proof. By induction on $|X|$. For $|X| = 0$ this is obvious. Assuming the statement for X's of size at most j, let us prove for the statement for $|X| = j+1$. Write $X = \{x\}\cup X'$ for $|X'| = j$. Then by Lemma 2.13

$$f_{\approx i,F}(I\cup X) = f_{\approx i-1,F|_{\{x\}}}(I\cup X') - f_{\approx i-1,F}(I\cup X'). \tag{5}$$

Applying the induction hypothesis on each term, we get that

$$f_{\approx i-1,F|_{\{x\}}}(I\cup X') = \sum_{Y\subseteq X'} (-1)^{|Y|} f_{\approx j,F|_{\{x\}\cup X'\setminus Y}}(I) = \sum_{Y\subseteq X,\ Y\not\ni x} (-1)^{|Y|} f_{\approx j,F|_{X\setminus Y}}(I)$$

and

$$f_{\approx i-1,F}(I\cup X') = \sum_{Y\subseteq X'} (-1)^{|Y|} f_{\approx j,F|_{X'\setminus Y}}(I) = \sum_{Y\subseteq X,\ Y\ni x} (-1)^{|Y|+1} f_{\approx j,F|_{X\setminus Y}}(I).$$

Plugging the two into (5) completes the proof. □

2.6 Main Results

In this section, we state our main results. Below, we are given a (r,ε) pseudo-random set, whose density is δ. We suggest to think of r as a constant, and on ε as $c\delta$ for a large, constant c.

Theorem 2.15. *Suppose $k \geqslant 2^{200\ell^3}\delta^{-4}$. If S is (r,ε) pseudo-random, then for every $i = 0, ...r$,*

$$W_{=i}[S] \leqslant \mathsf{exp}(i)\delta\varepsilon^{1/4} + \mathsf{exp}(\ell^3)\frac{1}{k^{1/16}}.$$

We prove this theorem in Section 3.

Theorem 2.16. *Let $\alpha \in (0,1)$, and S be a subset of vertices in $J(k,\ell,\alpha\ell)$ of density δ. Suppose $k \geqslant 2^{200\ell^3}\delta^{-4}$, and let $\varepsilon > 0$, $r \in \mathbb{N}$.*

If S is (r,ε) pseudo-random, then

$$\Phi(S) \geqslant 1 - \alpha^{r+1} - \mathsf{exp}(r)\varepsilon^{1/4} - \frac{\mathsf{poly}(2^{\ell^3}, 1/\delta)}{k^{1/16}}.$$

The proof is given in Section 4. We remark that this theorem easily implies Theorem 1.3 from the introduction.

Finally, we prove the following quantitative improvement of Theorem 2.15 in Section 5.

Theorem 2.17. *There exists $c > 0$, such that the following holds. Let $r, m \in \mathbb{N}$, $\varepsilon \geqslant \delta > 0$. Suppose $k \geqslant 2^{2000m^2(m+1)\ell^3}\delta^{-4}$. If S is (r,ε) pseudo-random of density δ, then for every $i = 0, \ldots r$,*

$$W_{=i}[S] \leqslant (c \cdot m)^i \delta \varepsilon^{1-\frac{1}{2m}} + \frac{\mathsf{exp}(\ell^3)}{k^{1/(8m^2(m+1))}}.$$

In words, the above theorem asserts that for constant r, a (r,ε) pseudo-random set S can have at most $\delta\varepsilon^{1-o(1)}$ of its weight on the first r levels. This statement should be compared to the r-level inequalities on the Boolean hypercube, stating that if $f\colon \{-1,1\}^n \to \{0,1\}$ has density δ, then its fourier weight on level r is at most $\delta^2 \log^r(1/\delta)$.

3 Proof of Theorem 2.15

Notation. In this secton we sometimes use the notation $x = y \pm \varepsilon$ to say that $|x - y| \leqslant \varepsilon$.

Let S be a (r,ε) pseudo-random set of density δ. Let $F\colon J(k,\ell,t)$ be the indicator function of S, and let $i \in \{0, ..., r\}$. For $i = 0$ the claim is obvious, so consider $i > 0$. Write the function F according to its decomposition $F = F_{=0} + F_{=1} + \ldots + F_{=\ell}$, and denote $\eta = W_{=i}[F]$. Assume $\eta \geqslant \delta^2$ since otherwise we are done.

Then we have $\mathbb{E}_A\left[F_{=i}[A]^2\right] = \eta$, and by orthogonality $\mathbb{E}_A\left[(F - F_{=i})[A]^2\right] = \delta - \eta$.

3.1 Information about the second moments

Claim 3.1. $\mathbb{E}_A\left[F^2_{\approx i}[A]\right] = \eta \pm 3\frac{2^{50\ell^3}}{\sqrt{k}}$.

Proof. By the triangle inequality

$$|\|F_{=i}\|_2 - \|F_{\approx i}\|_2| \leqslant \|F_{=i} - F_{\approx i}\|_2 \leqslant \frac{2^{50\ell^3}}{\sqrt{k}}$$

and therefore

$$\|F_{\approx i}\|_2^2 = \|F_{=i}\|_2^2 \pm 3\frac{2^{50\ell^3}}{\sqrt{k}} = \eta \pm 3\frac{2^{50\ell^3}}{\sqrt{k}}.$$

□

Claim 3.2. $\mathbb{E}_I\left[f_{\approx i}(I)^2\right] = \frac{W^{=i}[F]}{\binom{\ell}{i}} \pm 4\frac{2^{50\ell^3}}{\sqrt{k}} = \frac{\eta}{\binom{\ell}{i}} \pm 4\frac{2^{50\ell^3}}{\sqrt{k}}$.

Proof. Expand out $\mathbb{E}_A\left[F^2_{\approx i}[A]\right]$. On the one hand it is equal $\eta \pm 3\frac{2^{50\ell^3}}{\sqrt{k}}$ by the previous claim. On the other hand, it is equal

$$\mathop{\mathbb{E}}_{A}\left[\left(\sum_{I\subseteq A,\ |I|=i} f_{\approx i}(I)\right)^2\right] = \mathop{\mathbb{E}}_{A}\left[\sum_{I\subseteq A,\ |I|=i} f_{\approx i}(I)^2\right] + \mathop{\mathbb{E}}_{A}\left[\sum_{I\neq I'\subseteq A,\ |I|=|I'|=i} f_{\approx i}(I) f_{\approx i}(I')\right]$$

$$= \binom{\ell}{i}\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right] + \mathop{\mathbb{E}}_{A}\left[\sum_{d=0}^{i-1}\sum_{\substack{I,I'\subseteq A\\ |I|=|I'|=i,\ |I\cap I'|=d}} f_{\approx i}(I) f_{\approx i}(I')\right]$$

$$= \binom{\ell}{i}\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right] + \sum_{d=0}^{i-1}\binom{\ell}{d}\binom{\ell-d}{i-d}\binom{\ell-i}{i-d}\mathop{\mathbb{E}}_{\substack{I,I',\ |I|=|I'|=i\\ |I\cap I'|=d}}\left[f_{\approx i}(I) f_{\approx i}(I')\right].$$

We shall show that the second sum is small. Fix $d \in \{0, ..., i-1\}$ and consider the last expectation. Then it is equal

$$\mathop{\mathbb{E}}_{I,\ |I|=i}\left[f_{\approx i}(I)\mathop{\mathbb{E}}_{\substack{I',\ |I'|=i\\ |I\cap I'|=d}}\left[f_{\approx i}(I')\right]\right].$$

Fix I in the outside expectation and consider the inner expectation is

$$\mathop{\mathbb{E}}_{D\subseteq I}\left[\mathop{\mathbb{E}}_{\substack{I',\ |I'|=i\\ I\cap I'=D}}\left[f_{\approx i}(I')\right]\right] = \mathop{\mathbb{E}}_{D\subseteq I}\left[\mu_D(f_{\approx i}) \pm \frac{i^2}{k}\|f_{\approx i}\|_\infty\right],$$

where we used the fact that the uniform distribution of I' over I' that contain D, and over I' such that $I'\cap I = D$, is i^2/k close. Applying Facts 2.11 and 2.12 we get that the overall expectation in absolute value is at most $\frac{2^{12i^2}}{k}$. Therefore for each d the corresponding expectation is at most

$$\mathop{\mathbb{E}}_{I,\ |I|=i}\left[\|f_{\approx i}\|_\infty \frac{2^{12i^2}}{k}\right] \leqslant \frac{2^{13i^2}}{k},$$

we used Fact 2.12.

Plugging it into the first equation in the proof, we see that

$$\left|\mathop{\mathbb{E}}_{A}\left[F^2_{\approx i}[A]\right] - \binom{\ell}{i}\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right]\right| \leqslant \sum_{d=0}^{i-1}\binom{\ell}{d}\binom{\ell-d}{i-d}\binom{\ell-i}{i-d}\frac{2^{13i^2}}{k} \leqslant \frac{2^{17\ell^2}}{k},$$

and therefore

$$\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right] = \frac{1}{\binom{\ell}{i}}\mathop{\mathbb{E}}_{A}\left[F^2_{\approx i}[A]\right] \pm \frac{1}{\binom{\ell}{i}}\frac{2^{17\ell^2}}{k} = \frac{\eta}{\binom{\ell}{i}} \pm 4\frac{2^{50\ell^3}}{\sqrt{k}}.$$

□

Corollary 3.3. *Let $A \subseteq [k]$, $|A| \leqslant a$. Then*

$$\mathop{\mathbb{E}}_{|B|=i-a}\left[f_{\approx i-a, F|_A}(B)^2\right] \leqslant \frac{2 i^{i-a}\varepsilon}{\ell^{i-a}} + 4\frac{2^{50\ell^3}}{\sqrt{k}}.$$

Proof. By Claim 3.2,

$$\mathop{\mathbb{E}}_{|B|=i-a}\left[f_{\approx i-a,F|_A}(B)^2\right] \leqslant \frac{W^{=i-a}[F|_A]}{\binom{\ell-a}{i-a}} \pm 4\frac{2^{50\ell^3}}{\sqrt{k}}.$$

Note that since F is (r,ε) pseudo-random, $W^{=i-a}[F|_A] \leqslant \mu(F|_A) \leqslant \mu(F)+\varepsilon \leqslant 2\varepsilon$. Additionally,

$$\binom{\ell-a}{i-a} \geqslant \left(\frac{\ell-a}{i-a}\right)^{i-a} = \left(\frac{\ell}{i}\right)^{i-a} \geqslant \ell^{i-a}i^{-(i-a)}.$$

Combining these two yields the result. □

Corollary 3.4. *Let $A \subseteq [k]$, $|A| = a$. Then*

$$\mathop{\mathbb{E}}_{\substack{B\subseteq[k]\setminus A\\ |B|=i-a}}\left[f_{\approx i,F}(A\cup B)^2\right] \leqslant \frac{2^{2i+1}i^{i-a}\varepsilon}{\ell^{i-a}} + 4\frac{2^{52\ell^3}}{\sqrt{k}}.$$

Proof. By Corollary 2.14

$$\begin{aligned} f_{\approx i,F}(A\cup B)^2 &= \left(\sum_{Y\subseteq A}(-1)^{|Y|}f_{\approx i-a,F|_{A\setminus Y}}(B)\right)^2 \\ &\leqslant 2^{|A|}\sum_{Y\subseteq A} f_{\approx i-a,F|_{A\setminus Y}}(B)^2, \end{aligned}$$

the last inequality is by Cauchy-Schwarz. Therefore by linearity of expectation and Corollary 3.3

$$\begin{aligned} \mathop{\mathbb{E}}_{\substack{B\subseteq[k]\setminus A\\ |B|=i-a}}\left[f_{\approx i,F}(A\cup B)^2\right] &\leqslant 2^{|A|}\sum_{Y\subseteq A}\mathop{\mathbb{E}}_{\substack{B\subseteq[k]\setminus A\\ |B|=i-a}}\left[f_{\approx i-a,F|_{A\setminus Y}}(B)^2\right] \\ &\leqslant 2^{|A|}\sum_{Y\subseteq A}\frac{2i^{i-a}\varepsilon}{\ell^{i-a}} + 4\frac{2^{50\ell^3}}{\sqrt{k}} \\ &\leqslant \frac{2^{2i+1}i^{i-a}\varepsilon}{\ell^{i-a}} + 4\frac{2^{52\ell^3}}{\sqrt{k}}. \end{aligned}$$

□

3.2 A lower bound on the fourth moment

Claim 3.5. $\Pr_A\left[|F_{=i}[A]| \geqslant \frac{\eta}{4\delta}\right] \geqslant \frac{\eta}{2}$.

Proof. By Markov's inequality

$$\Pr_A\left[|F[A]-F_{=i}[A]|^2 \geqslant 1-\frac{\eta}{2\delta}\right] \leqslant \frac{\delta-\eta}{1-\frac{\eta}{2\delta}} = \delta\frac{2\delta-2\eta}{2\delta-\eta} \leqslant \delta-\frac{\eta}{2}.$$

Since the probability $F[A]=1$ is δ, we conclude probability that $|F[A]-F_{=i}[A]|^2 \leqslant 1-\frac{\eta}{2\delta}$ and $F[A]=1$ is at least $\frac{\eta}{2}$. For any such A we have

$$|1-F_{=i}[A]|^2 \leqslant 1-\frac{\eta}{2\delta},$$

implying

$$F_{=i}[A] \geqslant \frac{\eta}{4\delta}.$$

□

Claim 3.6. $\Pr_A\left[F_{\approx i}[A] \geqslant \frac{\eta}{8\delta}\right] \geqslant \frac{\eta}{4}$.

Proof. Let $E_1 = \left\{ A \mid F_{=i}[A] \geqslant \frac{\eta}{4\delta} \right\}$, and $E_2 = \left\{ A \mid \ |F_{=i}[A] - F_{\approx i}[A]| \leqslant \frac{\eta}{8\delta} \right\}$.

By Theorem 2.10 we have $\|F_{=i} - F_{\approx i}\|_2^2 \leqslant \frac{2^{100\ell^3}}{k}$, and hence by Markov's inequality

$$\Pr\left[\overline{E_2}\right] = \Pr_A\left[|F_{=i}[A] - F_{\approx i}[A]| \geqslant \frac{\eta}{8\delta}\right] \leqslant \frac{2^{100\ell^3}}{k}\left(\frac{8\delta}{\eta}\right)^2 \leqslant \frac{64\cdot 2^{100\ell^3}}{k\delta^2} \leqslant \frac{\eta}{4},$$

where we used $\eta \geqslant \delta^2$ and the condition on k.

Noting that whenever E_1, E_2 occur we have $F_{\approx i}[A] \geqslant \frac{\eta}{8\delta}$, we conclude that

$$\Pr_A\left[F_{\approx i}[A] \geqslant \frac{\eta}{8\delta}\right] \geqslant \Pr\left[E_1\right] - \Pr\left[\overline{E_2}\right] \geqslant \frac{\eta}{4},$$

we used Claim 3.5 in the last inequality. □

The above claim immediately implies the following.

Claim 3.7. $\mathbb{E}_A\left[F_{\approx i}^4[A]\right] \geqslant \frac{\eta^5}{2^{14}\delta^4}$.

3.3 An upper bound on the fourth moment

The following lemma is the main technical lemma of this section.

Lemma 3.8. *Suppose $k \geqslant 2^{100\ell^3}$, $\ell \geqslant 2$. If F is (i,ε) pseudo-random, then*

$$\mathop{\mathbb{E}}_A\left[F_{\approx i}^4[A]\right] \leqslant \mathsf{exp}(i)\varepsilon\eta + \frac{2^{27\ell^3}}{k^{1/4}}.$$

Proof. By opening the brackets we have

$$\mathop{\mathbb{E}}_A\left[F_{\approx i}^4[A]\right] = \mathop{\mathbb{E}}_A\left[\sum_{I_1,I_2,I_3,I_4\subseteq A} f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)\right].$$

Denote

$$D(A,d) = \left\{ (I_1,I_2,I_3,I_4) \mid I_1,...,I_4 \subseteq A, |I_1\cup I_2\cup I_3\cup I_4| = d \right\},$$

and note that $|D(A,d)|$ depends only on i,d — denote it by $\beta_{i,d}$. Then

$$\mathop{\mathbb{E}}_A\left[F_{\approx i}^4[A]\right] = \sum_{d=i}^{4i}\beta_{i,d}\mathop{\mathbb{E}}_A\left[\mathop{\mathbb{E}}_{(I_1,I_2,I_3,I_4)\in D(A,d)}\left[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)\right]\right].$$

An intersection pattern $\vec{\sigma_d}$ of a tuple (I_1,I_2,I_3,I_4) (whose direct sum is d dimensional) is a vector indicating the size of the intersection of any pair, triple and quadruple of sets from the tuple. Note that when sampling the quadruple $(I_1,...,I_4)$, the probability of getting each quadruple depends only on its intersection pattern. Therefore there is a distribution $\gamma(\vec{\sigma}_d)$ over intersection patterns such that the previous expression equals

$$\mathop{\mathbb{E}}_{A}\left[F_{\approx i}^4[A]\right] = \sum_{d=i}^{4i}\sum_{\vec{\sigma}_d}\gamma(\vec{\sigma}_d)\beta_{i,d}\mathop{\mathbb{E}}_{A}\left[\mathop{\mathbb{E}}_{\substack{(I_1,I_2,I_3,I_4)\sim\gamma(\vec{\sigma}_d)\\ I_1,\dots,I_4\subseteq A}}\left[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)\right]\right]. \tag{6}$$

Note that now the distribution over $(I_1, ..., I_4)$ is uniform over all quadruples whose direct sum is of dimension d and is intersection pattern is $\vec{\sigma}_d$, and thus

$$\mathop{\mathbb{E}}_{A}\left[F_{\approx i}^4[A]\right] = \sum_{d=i}^{4i}\sum_{\vec{\sigma}_d}\gamma(\vec{\sigma})\beta_{i,d}\mathop{\mathbb{E}}_{(I_1,I_2,I_3,I_4)\sim\gamma(\vec{\sigma_d})}[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)].$$

The following lemma is the key in the previous lemma.

Lemma 3.9. *Suppose $k \geqslant 2^{100\ell^3}$. Let i, d be integers such that $i \leqslant d \leqslant 4i$ and let $\vec{\sigma}_d$ be any intersection pattern. Then*

$$\left|\mathop{\mathbb{E}}_{(I_1,I_2,I_3,I_4)\sim\gamma(\vec{\sigma}_d)}[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)]\right| \leqslant \frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + 30\frac{2^{25\ell^3}}{k^{1/4}}$$

We defer the proof of Lemma 3.9 to the next section and show how to complete the proof of Lemma 3.8 based on it. Using Lemma 3.9 we see that

$$(6) \leqslant \sum_{d=i}^{4i}\sum_{\vec{\sigma}}\gamma(\vec{\sigma})\beta_{i,d}\left(\frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + 30\frac{2^{25\ell^3}}{k^{1/4}}\right).$$

Note that

$$\beta_{i,d} \leqslant \binom{\ell}{d}\binom{d}{i}^4 \leqslant \left(\frac{\ell e}{d}\right)^d\left(\frac{d\cdot e}{i}\right)^{4i} \leqslant \ell^d d^{-d}(4e^2)^{4i},$$

and so by the previous inequality

$$\begin{aligned}(6) &\leqslant \sum_{d=i}^{4i}\sum_{\vec{\sigma}}\gamma(\vec{\sigma})\ell^d d^{-d}(4e^2)^{4i}\left(\frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + 30\frac{2^{25\ell^3}}{k^{1/4}}\right)\\ &\leqslant \left(\mathsf{exp}(i)\varepsilon\eta + 30\frac{2^{26\ell^3}}{k^{1/4}}\right)\sum_{d=i}^{4i}(i/d)^d\\ &\leqslant \mathsf{exp}(i)\varepsilon\eta + \frac{2^{27\ell^3}}{k^{1/4}}.\end{aligned}$$

$\square$

3.4 Wrapping things up

Proof of Theorem 2.15. Combining Claim 3.7 and Lemma 3.8 we see that

$$\frac{\eta^5}{2^{14}\delta^4} \leqslant \mathop{\mathbb{E}}_{A}\left[F_{\approx i}^4[A]\right] \leqslant \mathsf{exp}(i)\varepsilon\eta + \frac{2^{27\ell^3}}{k^{1/4}}.$$

Rearranging we see that

$$\begin{aligned}\eta &\leqslant \mathsf{exp}(i)\delta\varepsilon^{1/4} + 16\frac{2^{7\ell^3}}{k^{1/16}}\frac{\delta^2}{\sqrt{\eta}} \\ &\leqslant \mathsf{exp}(i)\cdot\delta\varepsilon^{1/4} + 16\frac{2^{7\ell^3}}{k^{1/16}},\end{aligned}$$

the last inequality is by $\eta \geqslant \delta^2$. □

3.5 Proof of Lemma 3.9

Proof. Let $x_1, ..., x_d$ be chosen uniformly from $[k]$, and let $P_1, ..., P_4$ be sets depending on $x_1, ..., x_d$ corresponding to the intersection pattern σ_d. Note that

$$\left|\mathop{\mathbb{E}}_{x_1,\dots,x_d}\left[f_{\approx i}(P_1)f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\right] - \mathop{\mathbb{E}}_{(I_1,I_2,I_3,I_4)\sim\gamma(\vec{\sigma_d})}\left[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)\right]\right| \leqslant \frac{d^2}{k}\|f_{\approx i}\|_\infty^4,$$

since the distributions of the I's and the P's are $\frac{d^2}{k}$ close to each other. Therefore, it suffices to show an upper bound of the expectation involving the P's.

Proposition 3.10. *If there is x_j that appears in only one of the P's, then*

$$\left|\mathop{\mathbb{E}}_{x_1,\dots,x_d}\left[f_{\approx i}(P_1)f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\right]\right| \leqslant \frac{2^{14i^2}d^2}{k}.$$

Proof. Assume without loss of generality $j = 1$ and that P_1 contains x_1. Then

$$\mathop{\mathbb{E}}_{x_1,\dots,x_d}\left[f_{\approx i}(P_1)f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\right] = \mathop{\mathbb{E}}_{x_2,\dots,x_d}\Big[f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\mathbb{E}_{x_1}\left[f_{\approx i}(P_1)\mid x_2,...,x_d\right]\Big]. \tag{7}$$

Fix $x_2, ..., x_d$ and denote $J = \{x_2, ..., x_d\} \cap P_1$. Since $x_1 \in P_1$, we have that $|J| \leqslant |P_1| - 1 = i - 1$. Additionally

$$\left|\mathbb{E}_{x_1}\left[f_{\approx i}(P_1)\mid x_2,...,x_d\right]\right| \leqslant \left|\mathop{\mathbb{E}}_{I\supseteq J}\left[f_{\approx i}(I)\right]\right| + \frac{d^2}{k}\|f_{\approx i}\|_\infty.$$

The latter expectation is upper bounded by $\frac{2^{10i^2}}{k}$ from Fact 2.11. Combining this with Fact 2.12 we conclude that

$$\left|\mathbb{E}_{x_1}\left[f_{\approx i}(P_1)\mid x_2,...,x_d\right]\right| \leqslant \frac{2^{11i^2}d^2}{k}.$$

Using this, the triangle inequality and Fact 2.12 on (7) completes the proof. □

Thus we may assume that every x_j appears at least in two sets. Next, denote by $H_4 = P_1 \cap P_2 \cap P_3 \cap P_4$ the set of x_j's appearing in all sets, $h_4 = |H_4|$, H_3 the set of x_j's appearing in three of the sets, $h_3 = |H_3|$, H_2 the set of x_j's appearing in precisely two sets, $h_2 = |H_2|$.

Claim 3.11.

$$\left| \mathop{\mathbb{E}}_{H_4,H_3,H_2} [f_{\approx i}(P_1) f_{\approx i}(P_2) f_{\approx i}(P_3) f_{\approx i}(P_4)] \right| \leqslant \mathop{\mathbb{E}}_{H_4,H_3} \left[\sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_1)\right]} \sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_2)\right]} \sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_3)\right]} \sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_4)\right]} \right]$$

Proof. Let $x \in H_2$, and suppose P_{j_1}, P_{j_2} contain it but not P_{j_3}, P_{j_4}. Then the left hand side is at most

$$\left| \mathop{\mathbb{E}}_{H_4,H_3,H_2 \setminus \{x\}} \left[|f_{\approx i}(P_{j_3})| \, |f_{\approx i}(P_{j_4})| \mathop{\mathbb{E}}_{x} [|f_{\approx i}(P_{j_1})| \, |f_{\approx i}(P_{j_2})|] \right] \right|.$$

Applying Cauchy-Schwarz inequality on the inner expectation, the above expression is upper-bounded by

$$\left| \mathop{\mathbb{E}}_{H_4,H_3,H_2 \setminus \{x\}} \left[|f_{\approx i}(P_{j_3})| \, |f_{\approx i}(P_{j_4})| \sqrt{\mathop{\mathbb{E}}_{x} \left[f^2_{\approx i}(P_{j_1})\right]} \sqrt{\mathop{\mathbb{E}}_{x} \left[f^2_{\approx i}(P_{j_2})\right]} \right] \right|.$$

Continuing in this manner - namely picking each time a new variable from H_2, isolating the two terms that depends on it and applying Cauchy-Schwarz on that expectation, yields the desired bound. □

We next upper bound

$$\mathop{\mathbb{E}}_{H_4,H_3} \left[\sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_1)\right]} \sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_2)\right]} \sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_3)\right]} \sqrt{\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_4)\right]} \right] \tag{8}$$

Using Cauchy-Schwarz inequality,

$$\begin{aligned}
(8) &\leqslant \sqrt{\mathop{\mathbb{E}}_{H_4,H_3} \left[\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_1)\right] \mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_2)\right] \right]} \sqrt{\mathop{\mathbb{E}}_{H_4,H_3} \left[\mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_3)\right] \mathop{\mathbb{E}}_{H_2} \left[f^2_{\approx i}(P_4)\right] \right]} \\
&\leqslant \max_{H_3 \cap P_2, H_4} \sqrt{\mathop{\mathbb{E}}_{H_3 \setminus P_2, H_2} \left[f^2_{\approx i}(P_1)\right]} \max_{H_3 \cap P_3, H_4} \sqrt{\mathop{\mathbb{E}}_{H_3 \setminus P_3, H_2} \left[f^2_{\approx i}(P_4)\right]} \\
&\quad \sqrt{\mathop{\mathbb{E}}_{H_4,H_3,H_2} \left[f^2_{\approx i}(P_2)\right]} \sqrt{\mathop{\mathbb{E}}_{H_4,H_3,H_2} \left[f^2_{\approx i}(P_3)\right]}
\end{aligned} \tag{9}$$

The product of the third and the fourth term is equal to

$$\mathop{\mathbb{E}}_{P_3} \left[f^2_{\approx i}(P_3)\right] \leqslant \frac{\eta \cdot i^i}{\ell^i} + 4\frac{2^{50\ell^3}}{\sqrt{k}}.$$

Where the last inequality is by Claim 3.2 and the estimate $\binom{\ell}{i} \geqslant \left(\frac{\ell}{i}\right)^i$.

Next, we estimate the maximums. For $j = 1, 2, 3, 4$ define $H_{3,j} = H_3 \cap P_j$. Apply Corollary 3.4 on the first maximum with $A = (H_4, H_{3,1} \cap P_2)$ and $B = (H_2 \cap P_1, H_{3,1} \setminus P_2)$, we get it is upper bounded by

$$\max_{A} \sqrt{\mathop{\mathbb{E}}_{B} \left[f^2_{\approx i}(A \cup B)\right]} \leqslant \sqrt{\frac{2^{2i+1} i^{|B|} \varepsilon}{\ell^{|B|}} + 4\frac{2^{52\ell^3}}{\sqrt{k}}} \leqslant \frac{2^{i+1} i^{\lambda_1/2} \sqrt{\varepsilon}}{\ell^{\lambda_1/2}} + 2\frac{2^{25\ell^3}}{k^{1/4}},$$

where $\lambda_1 = |H_2 \cap P_1| + |H_{3,1} \setminus P_2|$

Similarly, the second maximum is upper bounded by

$$\frac{2^{i+1} i^{\lambda_2/2}\sqrt{\varepsilon}}{\ell^{\frac{1}{2}\lambda_2}} + 2\frac{2^{25\ell^3}}{k^{1/4}},$$

where $\lambda_2 = |H_2 \cap P_4| + |H_{3,4} \setminus P_3|$

Combining everything, we see that

$$9 \leqslant \frac{2^{2i+2} i^{\frac{1}{2}(2i+\lambda_1+\lambda_2)}\varepsilon\eta}{\ell^{\frac{1}{2}(2i+\lambda_1+\lambda_2)}} + 29\frac{2^{25\ell^3}}{k^{1/4}}. \tag{10}$$

By counting occurrences of points in the P_2, we get that $i = h_4 + |H_2 \cap P_2| + |H_{3,2}|$ and by counting occurrences of points in P_3, $i = h_4 + |H_2 \cap P_3| + |H_{3,3}|$. Also, note that $H_{3,1} \setminus P_2 = H_3 \setminus H_{3,2}$ since any $x \in H_3 \setminus H_{3,2}$ is in 3 of the sets but not in P_2, and hence it is in P_1. Thus, $|H_{3,1} \setminus P_2| = |H_3 \setminus H_{3,2}| = h_3 - |H_{3,2}|$ and similarly $|H_{3,4} \setminus P_3| = |H_3 \setminus H_{3,3}| = h_3 - |H_{3,3}|$. Plugging everything into (11), we see that the exponent of ℓ (and also i) is equal to

$$h_4 + |H_2 \cap P_2| + |H_{3,2}| + h_4 + |H_2 \cap P_3| + |H_{3,3}| + |H_2 \cap P_1| + |H_2 \cap P_4| + h_3 - |H_{3,3}| + h_3 - |H_{3,2}|$$

which is $2(h_4 + h_3 + h_2) = 2d$. Thus,

$$9 \leqslant \frac{2^{2i+2} i^d \varepsilon\eta}{\ell^d} + 29\frac{2^{25\ell^3}}{k^{1/4}}. \tag{11}$$

□

4 Pseudorandomness implies Expansion

Proof of Theorem 2.16. Let S be a set as in the Theorem, and let F be its indicator functions in $J(k, \ell, \alpha\ell)$. Note that

$$\delta(1 - \Phi(S)) = \langle F, J(k, \ell, \alpha\ell)F\rangle. \tag{12}$$

Claim 4.1.

$$\langle F, J(k, \ell, \alpha\ell)F\rangle \leqslant \sum_{i=0}^{\ell} \lambda_i(k, \ell, \alpha\ell) W^{=i}[F] + \frac{2^{101\ell^3}}{k}.$$

Proof. Writing $F = F_{=0} + ... + F_{=\ell}$ we have

$$\begin{aligned}\langle F, J(k, \ell, \alpha\ell)F\rangle = \langle F, \sum_{i=0}^{\ell} J(k, \ell, \alpha\ell)F_{=i}\rangle &= \langle F, \sum_{i=0}^{\ell} \lambda_i(k, \ell, \alpha\ell)F_{=i}\rangle \\ &+ \langle F, \sum_{i=0}^{\ell} J(k, \ell, \alpha\ell)F_{=i} - \lambda_i(k, \ell, \alpha\ell)F_{=i}\rangle.\end{aligned}$$

The first term equals

$$\langle \sum_{i=0}^{\ell} F_{=i}, \sum_{i=0}^{\ell} \lambda_i(k, \ell, \alpha\ell)F_{=i}\rangle = \sum_{i=0}^{\ell} \lambda_i(k, \ell, \alpha\ell) W^{=i}[F]$$

by orthogonality. The second term can be upper bounded using Cauchy-Schwarz and the triangle inequality by

$$\left|\langle F, \sum_{i=0}^{\ell} J(k,\ell,\alpha\ell)F_{=i} - \lambda_i(k,\ell,\alpha\ell)F_{=i}\rangle\right| \leqslant \|F\|_2 \left(\sum_{i=0}^{\ell} \|J(k,\ell,\alpha\ell)F_{=i} - \lambda_i(k,\ell,\alpha\ell)F_{=i}\|_2\right).$$

Using Theorem 2.10 and $\|F\|_2 \leqslant 1$, the last expression is at most

$$\sum_{i=0}^{\ell} 2\frac{2^{100\ell^3}}{k} + \|J(k,\ell,\alpha\ell)F_{\approx i} - \lambda_i(k,\ell,\alpha\ell)F_{\approx i}\|_2 \leqslant \sum_{i=0}^{\ell} 2\frac{2^{100\ell^3}}{k} + \frac{2^{4\ell}}{k}\|f_{\approx i}\|_\infty.$$

the last inequality is by Theorem 2.7. Therefore using Fact 2.12 and simplification last expression is at most

$$(\ell+1)3\frac{2^{100\ell^3}}{k} \leqslant \frac{2^{101\ell^3}}{k}.$$

□

The following upper bound follows easily by the definition of λ_i.

Fact 4.2. $\lambda_i(k,\ell,\alpha\ell) \leqslant \alpha^i$.

Therefore we conclude that

$$\langle F, J(k,\ell,\alpha\ell)F\rangle \leqslant \sum_{i=0}^{r} \alpha^i W^{=i}[F] + \sum_{i=r+1}^{\alpha\ell} \alpha^i W^{=i}[F] + \frac{2^{101\ell^3}}{k}.$$

Let $i \leqslant r$. Observe that since S is (r,ε) pseudo-random it follows that S is (i,ε) pseudo-random. Therefore, applying Theorem 2.15 we see that for $i = 0, 1, ..., r$,

$$W^{=i}[F] \leqslant \mathsf{exp}(i)\delta\varepsilon^{1/4} + 16\frac{2^{7\ell^3}}{k^{1/16}}.$$

For the second sum, note that by Parseval the sum of all weights of F is at most δ. Hence, combining these two facts we conclude

$$\begin{aligned}\langle F, J(k,\ell,\alpha\ell)F\rangle &\leqslant \sum_{i=0}^{r} \alpha^i \mathsf{exp}(i)\delta\varepsilon^{1/4} + \alpha^{r+1}\sum_{i=r+1}^{\alpha\ell} W_{=i}[F] + 20\frac{2^{7\ell^3}}{k^{1/16}} \\ &\leqslant \mathsf{exp}(r)\delta\varepsilon^{1/4} + \delta\alpha^{r+1} + 20\frac{2^{7\ell^3}}{k^{1/16}}.\end{aligned}$$

Plugging this into (12) and simplifying yields

$$\Phi(S) \geqslant 1 - \alpha^{r+1} - \mathsf{exp}(r)\varepsilon^{1/4} - 20\frac{2^{7\ell^3}}{k^{1/16}\delta}.$$

□

5 Proof of Theorem 2.17

Let S be a (r, ε) pseudo-random set, and F be its indicator function. The proof of Theorem 2.17 follows the same outline as the proof of Theorem 2.15, except that we use higher moment.

Let $0 \leqslant i \leqslant r$.

Claim 5.1.

$$\mathop{\mathbb{E}}_{A}\left[F_{\approx i}^{2m}[A]\right] \geqslant \frac{\eta^{2m+1}}{2^{6m+2}\delta^{2m}}$$

Proof. Immediate from Claim 3.6. □

The intersection pattern of $P_1, \ldots, P_{2m}$ is a vector σ indicating the sizes of all all intersections of any collection of the sets.

Lemma 5.2. *Let σ_d be an intersection pattern for $2m$ sets, and let $P_1, \ldots, P_{2m}$ be sets that match this intersection pattern, that have formal elements $x_1, \ldots, x_d$. Then*

$$\left|\mathop{\mathbb{E}}_{x_1,\ldots,x_d}\left[\prod_{j=1}^{2m} f_{\approx i}(P_j)\right]\right| \leqslant 2^{4mi+2m}\frac{i^d\varepsilon^{2m-1}\eta}{\ell^d} + \frac{\mathsf{exp}(\ell^3)}{k^{1/(4m(2m+1))}}. \tag{13}$$

The rest of this section is devoted for the proof of this lemma.

Let σ_d be an intersection pattern. If among $x_1, \ldots, x_d$ there is a variable that appears only in one of the P's then lemma holds trivially:

Proposition 5.3. *If there is x_j that appears in only one of the P's, then*

$$\left|\mathop{\mathbb{E}}_{x_1,\ldots,x_d}\left[\prod_{j=1}^{2m} f_{\approx i}(P_j)\right]\right| \leqslant \frac{2^{(11+m)i^2d^2}}{k}.$$

Proof. The same proof as in Proposition 3.10. □

We thus assume from now on that each x_j appears in at least 2 of the P's. For $s = 2, 3 \ldots, 2m$, let H_s be the set of x_j's that appear in exactly s of the P's. For convenience, let us also denote $H_{\geqslant s} = H_s \cup H_{s+1} \cup \ldots \cup H_{2m}$ and $H_{\leqslant s} = H_2 \cup H_3 \cup \ldots \cup H_s$.

For each one of the P-sets, define $e_1(P) \overset{def}{=} |f_{\approx i}(P)|$ and for $s \geqslant 2$ define

$$e_s(P) \overset{def}{=} \mathop{\mathbb{E}}_{H_{\leqslant s}}\left[f_{\approx i}^2(P)\right].$$

For each $a = 2, 3, \ldots, 2m$, define an operator ρ_a on random variables defined by

$$\rho_a(Z) = \mathop{\mathbb{E}}_{H_a}\left[Z^{a/(a-1)}\right].$$

For $j = 1, 2, \ldots, 2m-1$, define $T_{j,j}(P) \overset{def}{=} e_j(P)$. Inductively for $a > j$, define

$$T_{j,a}(P) \overset{def}{=} \rho_a(T_{j,a-1}(P)).$$

Denote for each s,

$$Q_s[H_{\geqslant s+1}] \stackrel{def}{=} \mathop{\mathbb{E}}_{H_{\leqslant s}}\left[\prod_{j=1}^{2m} |f_{\approx i}(P_j)|\right].$$

Note that for $s = 2m$, this is the term we wish to bound.

Proposition 5.4. *For every* $2 \leqslant s \leqslant 2m$*,*

$$Q_s[H_{\geqslant s+1}] \leqslant \left(\prod_{j=1}^{2m} T_{1,s}(P_j)\right)^{1/s}.$$

We shall use the following fact in the proof, which is a direct corollary of Holder's inequality.

Fact 5.5. *Let* $h_1, \ldots, h_n \colon \binom{k}{m} \to \mathbb{R}$*. Then*

$$\|h_1 \ldots h_n\|_1 \leqslant \|h_1\|_n \cdots \|h_n\|_n.$$

Proof of Proposition 5.4. The proof is by induction on s. The base case $s = 1$ is trivial. Let $s \geqslant 1$, assume we have proven for s, and prove for $s+1$. Note that

$$Q_{s+1}[H_{\geqslant s+2}] = \mathop{\mathbb{E}}_{H_{s+1}}[Q_s[H_{\geqslant s+1}]].$$

Applying the induction hypothesis, we get that

$$Q_{s+1}[H_{\geqslant s+2}] \leqslant \mathop{\mathbb{E}}_{H_{s+1}}\left[\prod_{j=1}^{2m} (T_{1,s}(P_j))^{1/s}\right].$$

Write $H_{s+1} = \left\{x_{j_1}, \ldots, x_{j_q}\right\}$. Iteratively, for each $y \in H_{s+1}$ we consider the $s+1$ P's it appears in, and then apply Fact 5.5 on them. Thus, for instance suppose we have that y is in $P_1, \ldots, P_{s+1}$, then we would get

$$Q_{s+1}[H_{\geqslant s+2}] \leqslant \mathop{\mathbb{E}}_{H_{s+1}\setminus\{y\}}\left[\prod_{j=1}^{s+1}\left(\mathop{\mathbb{E}}_{y}\left[T_{1,s}(P_j)^{(s+1)/s}\right]\right)^{1/(s+1)} \prod_{j=s+2}^{2m} T_{1,s}(P_j)\right].$$

Repeating this process for every $y \in H_{s+1}$, one gets

$$Q_{s+1}[H_{\geqslant s+2}] \leqslant \prod_{j=1}^{2m}\left(\mathop{\mathbb{E}}_{H_{s+1}\cap P_j}\left[T_{1,s}(P_j)^{(s+1)/s}\right]\right)^{1/(s+1)} = \prod_{j=1}^{2m} T_{1,s+1}(P_j)^{1/(s+1)}.$$

$\square$

Thus, we have that

$$\mathsf{LHS}(13) \leqslant \left(\prod_{j=1}^{2m} T_{1,2m}(P_j)\right)^{1/2m}. \tag{14}$$

Proposition 5.6. *For any P-set P and $1 \leqslant j \leqslant 2m-1$*

$$T_{j,2m}(P) \leqslant T_{j+1,2m}(P) \cdot \max_{H_{\geqslant j}} e_j(P)^{\frac{2m}{j(j+1)}}.$$

Proof. The proof is by induction on j. The induction basis $j = 2m-1$ follows similar lines of the induction step and is hence omitted. Assume the claim for $j \leqslant 2m-1$ and prove for $j-1$.

$$\begin{aligned} T_{j-1,2m}(P) &= \rho_{2m} \circ \ldots \circ \rho_j \left(e_{j-1}(P)\right) \\ &= \rho_{2m} \circ \ldots \circ \rho_{j+1} \left(\mathop{\mathbb{E}}_{H_j} \left[e_{j-1}(P)^{j/(j-1)} \right] \right) \end{aligned}$$

Clearly,

$$\mathop{\mathbb{E}}_{H_j} \left[e_{j-1}(P)^{j/(j-1)} \right] \leqslant \mathop{\mathbb{E}}_{H_j} \left[e_{j-1}(P) \right] \max_{H_{\geqslant j}} e_{j-1}(P)^{1/(j-1)} = e_j(P) \max_{H_{\geqslant j}} e_{j-1}(P)^{1/(j-1)}.$$

Also, note that each operator ρ_a is monotone on non-negative random variables, and for a random variable Z and a constant $c \geqslant 0$ we have that $\rho_a(cZ) = \rho_a(c)\rho_a(Z)$. Thus, combining the above two we get that

$$\begin{aligned} T_{j-1,2m}(P) &\leqslant [\rho_{2m} \circ \ldots \circ \rho_{j+1}](e_j(P)) \cdot [\rho_{2m} \circ \ldots \circ \rho_{j+1}](\max_{H_{\geqslant j}} e_{j-1}(P)^{1/(j-1)}) \\ &= T_{j,2m}(P) \max_{H_{\geqslant j}} e_{j-1}(P)^{2m/j(j-1)}. \end{aligned}$$

□

Repeated application of the above proposition yields that for any P-set,

$$T_{1,2m}(P) \leqslant T_{2m,2m}(P) \prod_{a=1}^{2m-1} \max_{H_{\geqslant a+1}} e_a(P)^{\frac{2m}{a(a+1)}} = \prod_{a=1}^{2m} \max_{H_{\geqslant a+1}} e_a(P)^{2m \cdot g(a)}.$$

where $g(a) = \frac{1}{a(a+1)}$ for $a \leqslant 2m-1$ and $g(2m) = \frac{1}{2m}$. We used the fact that $T_{2m,2m}(P) = e_{2m}(P)$. Plugging this into (14) yields

$$\mathsf{LHS}(13) \leqslant \prod_{j=1}^{2m} \prod_{a=1}^{2m} \max_{H_{\geqslant a+1}} e_{a-1}(P_j)^{g(a)}.$$

Using Corollary 3.4, we see that that for $a < 2m$,

$$\max_{H_{\geqslant a+1}} e_a(P) \leqslant 2^{2i+1} \frac{i^{\left|P \cap H_{\leqslant a}\right|} \varepsilon}{\ell^{\left|P \cap H_{\leqslant a}\right|}} + \frac{\mathsf{exp}(\ell^3)}{\sqrt{k}}.$$

For $a = 2m$, by Claim 3.2

$$\max_{H_{\geqslant a+1}} e_a(P) \leqslant \frac{i^{\left|P \cap H_{\leqslant a}\right|} \eta}{\ell^{\left|P \cap H_{\leqslant a}\right|}} + \frac{\mathsf{exp}(\ell^3)}{\sqrt{k}}.$$

Therefore,

$$\begin{aligned}
\mathsf{LHS}(13) &\leqslant \prod_{j=1}^{2m}\left(\frac{i^{\left|P_j\cap H_{\leqslant 2m}\right|}\eta}{\ell^{\left|P_j\cap H_{\leqslant 2m}\right|}}+\frac{\exp(\ell^3)}{\sqrt{k}}\right)^{g(2m)}\cdot\prod_{j=1}^{2m}\prod_{a=1}^{2m-1}\left(2^{2i+1}\frac{i^{\left|P_j\cap H_{\leqslant a}\right|}\varepsilon}{\ell^{\left|P_j\cap H_{\leqslant a}\right|}}+\frac{\exp(\ell^3)}{\sqrt{k}}\right)^{g(a)}\\
&\leqslant \prod_{j=1}^{2m}\left(\frac{i^{\left|P_j\cap H_{\leqslant 2m}\right|}\eta}{\ell^{\left|P_j\cap H_{\leqslant 2m}\right|}}\right)^{g(2m)}\cdot\prod_{j=1}^{2m}\prod_{a=1}^{2m}\left(2^{2i+1}\frac{i^{\left|P_j\cap H_{\leqslant a}\right|}\varepsilon}{\ell^{\left|P_j\cap H_{\leqslant a}\right|}}\right)^{g(a)}+\frac{\exp(\ell^3)}{k^{1/(4m(2m+1))}}\\
&\leqslant 2^{4mi+2m}\eta\varepsilon^{2m-1}\prod_{j=1}^{2m}\prod_{a=1}^{2m}\frac{i^{g(a)\cdot\left|P_j\cap H_{\leqslant a}\right|}}{\ell^{g(a)\cdot\left|P_j\cap H_{\leqslant a}\right|}}+\frac{\exp(\ell^3)}{k^{1/(4m(2m+1))}}. \qquad (15)
\end{aligned}$$

where to compute the power of ε we used the fact that $\sum_{a=1}^{2m-1} g(a) = 1-\frac{1}{2m}$ (telescopic sum). Consider the last product. It is equal to

$$\begin{aligned}
\prod_{j=1}^{2m}\prod_{a=1}^{2m}\prod_{r=1}^{a}\frac{i^{g(a)\cdot|P_j\cap H_r|}}{\ell^{g(a)\cdot|P_j\cap H_r|}} &= \prod_{a=1}^{2m}\prod_{r=1}^{a}\frac{i^{g(a)\cdot\sum_{j=1}^{2m}|P_j\cap H_r|}}{\ell^{g(a)\cdot\sum_{j=1}^{2m}|P_j\cap H_r|}}\\
&= \prod_{a=1}^{2m}\prod_{r=1}^{a}\frac{i^{g(a)\cdot r|H_r|}}{\ell^{g(a)\cdot r|H_r|}}\\
&= \prod_{r=1}^{2m}\prod_{a=r}^{2m}\frac{i^{g(a)\cdot r|H_r|}}{\ell^{g(a)\cdot r|H_r|}}\\
&= \prod_{r=1}^{2m}\frac{i^{r|H_r|\sum_{a=r}^{2m}g(a)}}{\ell^{r|H_r|\sum_{a=r}^{2m}g(a)}}.
\end{aligned}$$

In the second equality, we used the fact that each element in H_r is counted r times by the P_j's. In the third equality we interchanged the order of multiplication. Note that $\sum_{a=r}^{2m} g(a) = \frac{1}{r}$, hence the last product is equal to

$$\prod_{r=1}^{2m}\frac{i^{|H_r|}}{\ell^{|H_r|}} = \frac{i^{\sum_{r=1}^{2m}|H_r|}}{\ell^{\sum_{r=1}^{2m}|H_r|}} = \frac{i^d}{\ell^d}.$$

Plugging this into 15 yields

$$\mathsf{LHS}(13) \leqslant 2^{4mi+2m}\frac{i^d\eta\varepsilon^{2m-1}}{\ell^d}+\frac{\exp(\ell^3)}{k^{1/(4m(2m+1))}},$$

concluding the proof.

The following corollary is proven using a similar argument to the one in Lemma 3.8. Since the proof is almost identical, we omit it.

Corollary 5.7. *Suppose $k \geqslant 2^{100\ell^3}$, $\ell \geqslant 2$. If F is (i,ε) pseudo-random, then*

$$\mathbb{E}_A\left[F_{\approx i}^{2m}[A]\right] \leqslant (10m)^{4mi}\eta\varepsilon^{2m-1}+\frac{\exp(\ell^3)}{k^{1/(4m(2m+1))}}.$$

Proof of Theorem 2.17. Combining Claim 5.1 and Corollary 5.7, we get that

$$\frac{\eta^{2m+1}}{2^{6m+2}\delta^{2m}} \leqslant (10m)^{4mi}\eta\varepsilon^{2m-1} + \frac{\mathsf{exp}(\ell^3)}{k^{1/(4m(2m+1))}}.$$

Rearranging (and using $\eta \geqslant \delta^2$) we get that

$$\eta^{2m} \leqslant 2^{6m+2}(10m)^{4mi}\delta^{2m}\varepsilon^{2m-1} + \frac{\mathsf{exp}(\ell^3)}{k^{1/(4m(2m+1))}},$$

taking $2m$-root yields

$$\eta \leqslant 16(10m)^{2i}\delta\varepsilon^{1-\frac{1}{2m}} + \frac{\mathsf{exp}(\ell^3)}{k^{1/(8m^2(2m+1))}}$$

finishing the proof.

6 Open Problems

Derandomizing the hypercube. Many results in the PCP literature [16], construction of integrality gaps and non-embeddability results [23] rely on the small-set expansion property of the (noisy) hypercube. This property states that for every noise rate $\rho > 0$, there exists $\varepsilon > 0$ such that a set of density δ in the noisy-hypercube has expansion at least $1 - O(\delta^{\varepsilon})$. The quantitative aspect of these results is often determined by the number of vertices in the noisy-hypercube, which is large. Improving these results can be achieved by constructing a hypercube like graphs (e.g. that have a "small-set expansion" type property) with significantly smaller number of vertices. This was done by Barak et al. [2] via the "Short Code", that in many ways behaves similarly to the noisy hypercube and thus may replace the noisy hypercube in the applications above. Can our results be used to get improved results? This task is not trivial since the Johnson Graph does not have a small-set expansion property. We do have however a characterization of sets for which the small-set expansion property fails that is sometimes sufficient. For instance, the outer PCP of [20, 7] deviates from classical PCP constructions by incorporating this characterization into the analysis.

Understanding slices of the hypercube. The Johnson Graph represents a slice of the Boolean hypercube, which is a well-studied object from a combinatorial point of view [11, 10, 12, 13]. It is related to the study of sharp threshold of Boolean functions and graph properties [14]. Can our result lead to new results about slices of the Boolean hypercube the way that small set expansion of the hypercube and hypercontractivity in general led to results about the hypercube? Subsequent to our work, Lifshitz and the second author have made progress in this direction [24].

Direct product testing. For the Grassmann Graph, the characterization of small-sets that have near-perfect expansion implies a "direct product testing" result, i.e., an encoding and a (2-to-2) test on words, such that any word that passes the test with noticeable probability has global structure (inside some subgraph of Grassmann Graph) [3]. Our results are related, in the same sense, to the problem of direct product testing on the hypercube [18, 8, 4]. Can one obtain new results about direct product testing via our theorem? It is interesting to note that the notion of "zoom-ins" appears in many of these works, and often as an intermediate step (if there are more than 2 queries).

References

[1] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. *J. ACM*, 62(5):42, 2015.

[2] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. *SIAM J. Comput.*, 44(5):1287–1324, 2015.

[3] Boaz Barak, Pravesh Kothari, and David Steurer. Small-set expansion in shortcode graph and the 2-to-1 conjecture. *manuscript*, 2018.

[4] Roee David, Irit Dinur, Elazar Goldenberg, Guy Kindler, and Igor Shinkar. Direct sum testing. *SIAM J. Comput.*, 46(4):1336–1369, 2017.

[5] Irit Dinur and Elazar Goldenberg. Locally testing direct product in the low error range. pages 613–622, 2008.

[6] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? *Electronic Colloquium on Computational Complexity (ECCC)*, 23:198, 2016.

[7] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. On non-optimally expanding sets in grassmann graphs. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:94, 2017.

[8] Irit Dinur and Inbal Livni Navon. Exponentially small soundness for the direct product z-test. In Ryan O'Donnell, editor, *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*, volume 79 of *LIPIcs*, pages 29:1–29:50. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

[9] Irit Dinur and David Steurer. Direct product testing. In *Computational Complexity Conference*, 2014.

[10] Yuval Filmus. Friedgut-Kalai-Naor theorem for slices of the boolean cube. *Chicago J. Theor. Comput. Sci.*, 2016, 2016.

[11] Yuval Filmus. An orthogonal basis for functions over a slice of the boolean hypercube. *Electr. J. Comb.*, 23(1):P1.23, 2016.

[12] Yuval Filmus, Guy Kindler, Elchanan Mossel, and Karl Wimmer. Invariance principle on the slice. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 15:1–15:10, 2016.

[13] Yuval Filmus and Elchanan Mossel. Harmonicity and invariance on slices of the boolean cube. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 16:1–16:13, 2016.

[14] Ehud Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *J. Amer. Math. Soc*, 12:1017–1054, 1998.

[15] Oded Goldreich and Shmuel Safra. A combinatorial consistency lemma with application to proving the PCP theorem. *SIAM J. Comput.*, 29(4):1132–1154, 2000.

[16] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[17] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(04):439–562, 2006.

[18] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query pcps. *SIAM J. Comput.*, 41(6):1722–1768, 2012.

[19] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query PCPs. *SIAM Journal on Computing*, 41(6):1722–1768, 2012.

[20] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games and grassmann graphs. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:124, 2016.

[21] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in grassmann graph have near-perfect expansion. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:6, 2018.

[22] Subhash Khot and Dana Moshkovitz. Candidate hard unique game. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 63–76, 2016.

[23] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative-type metrics into l_1. *J. ACM*, 62(1):8:1–8:39, 2015.

[24] Noam Lifshitz and Dor Minzer, 2018. in preparation.

[25] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[26] Prasad Raghavendra and David Steurer. Integrality gaps for strong SDP relaxations of UNIQUE GAMES. In *50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA*, pages 575–585, 2009.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il