THEORY OF COMPUTING, Volume 21 (2), 2025, pp. 1–43
www.theoryofcomputing.org

Small-Set Expansion in the Johnson Graph

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

Received June 5, 2019; Revised February 20, 2024; Published May 26, 2025

Abstract. An n-vertex graph is called a *small-set expander* if any set of size $o(n)$ contains at most a $o(1)$ fraction of the edges that touch it. The goal of this paper is to investigate small-set expansion properties of the Johnson graph, which is not a small-set expander.

We obtain a qualitative descriptions of all small sets that violate the small-set expansion property in the Johnson graph: we show that any such set is correlated with some union of small intersections of "basic sets," where each basic set is a dictatorship—i. e., belonging to it depends only on containing a single coordinate i. This condition is necessary and sufficient, since any such set violates the small-set expansion property.

The statement and its proof are inspired by recent analogous questions on the Grassmann graph (Dinur et al., *STOC'18* and *Israel J. Math., 2021*) and their

*Research supported by NSF CCF-1422159, Simons Collaboration on Algorithms and Geometry, and Simons Investigator Award.

†This work was completed while the author was in the School of Computer Science, Tel Aviv University and was supported by Clore scholarship.

‡Supported by the National Science Foundation under grants number 1648712, 2200956 and 1705028.

§Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 835152), by an ISF grant 2017/13 and by a BSF grant 2016414.

ACM Classification: G.2.2

AMS Classification: 06E30, 05C48

Key words and phrases: Johnson graph, small-set expansion, hypercontractivity

 DOI: 10.4086/toc.2025.v021a002

application to the 2-to-1 Games Conjecture. To prove our results, we build on and extend the techniques of Dinur et al., *Israel J. Math., 2021*. Subsequently to our work, the full expansion hypothesis for the Grassmann graph was proved in Khot et al., *Ann. Math. 2023.*

1 Introduction

1.1 Graph expansion

For a regular graph $G = (V, E)$ and a set of vertices $S \subseteq V$, the edge expansion of S, denoted by $\Phi_G(S)$, is the probability of escaping the set S in a single step. That is, $\Phi_G(S)$ is the probability of landing outside S after picking a vertex u from S randomly, and walking along a randomly chosen edge (u, v).[1] Unless specifically stated otherwise, when we say "randomly" we mean uniformly. Expander graphs, i. e., graphs in which any set of size at most half the vertices has constant edge expansion, are widely used in Theoretical Computer Science in pseudorandomness, probabilistically checkable proofs, and more (see [27]). One particularly appealing aspect of expanders is that they have numerous, seemingly different definitions that yet turn out to be equivalent: a graph is an expander if and only if the second eigenvalue of its normalized adjacency matrix is bounded away from 1.

Small-set expansion is an incomparable notion of edge expansion. A graph G is called an (η, δ) small-set expander if any set S of vertices containing at most a δ fraction of the vertices, has $\Phi_G(S) \geqslant 1 - \eta$. In words, the probability to start from a random vertex in S and stay inside S after a random step, is at most η. While this notion is very natural, it is much less understood: it is often difficult to check if a graph is a small-set expander, as there is no equivalent eigenvalue-based definition; this is also the reason it is harder to work with. Another stark contrast is that while, given an input graph, the task of checking whether it is an expander graph can be done efficiently (as it amounts to computing eigenvalues of the adjacency matrix), no such results are known for small-set expansion. In fact, it is conjectured in [39, 40] that for all $\eta > 0$ there exists $\delta > 0$, such that given a graph $G = (V, E)$ it is NP-hard to distinguish between the case that G is an (η, δ) small-set expander, and the case that G contains a subset $S \subseteq V$ with $|S| \leqslant \delta\,|V|$ and $\Phi_G(S) \leqslant \eta$,

1.2 Noisy hypercube

For a parameter $\varepsilon > 0$, the noisy hypercube $\mathcal{H}_\varepsilon[n]$, defined below, is a widely used weighted undirected graph in Theoretical Computer Science and in the Analysis of Boolean Functions. The vertex set V is $\{0, 1\}^n$, and the edges are weighted according to the following randomized rule. To sample an edge, pick a vertex $u = (u_1, \ldots, u_n) \in_R V$ uniformly at random, and choose its neighbour $v = (v_1, \ldots, v_n)$ according to the distribution $T_{1-\varepsilon}(u)$ defined as follows: for each coordinate $i \in [n]$ independently, set $v_i = u_i$ with probability $1 - \varepsilon$ and otherwise resample

[1]Note that this definition also makes sense for weighted graphs, i. e., graphs in which the edges are assigned weights.

$v_i \in_R \{0,1\}$ (thus, the expected value of the Hamming distance between u and v is $\varepsilon n/2$). The edge (u,v) is the output of the procedure, and the weight of an edge (x,y) is the probability that (x,y) is selected.

The noisy hypercube is well known to be a small-set expander for any constant $\varepsilon > 0$: for every $\varepsilon > 0$, there exists $c > 0$, such that for small enough δ, any set S containing at most a δ fraction of the vertices has edge expansion at least $1 - \delta^c$. This fact follows from Bonami's Hypercontractive Inequality [6, 5, 24], and for Boolean-valued functions, they are equivalent.[2] This is helpful in deducing structural results on Boolean functions $f : \{0,1\}^n \to \{0,1\}$ that are the indicator functions of small sets, such as the KKL Theorem [29], that are then helpful for constructing PCPs [26, 15, 32, 13], SDP integrality gaps [16, 38], metric embedding lower bounds [36, 35] and much more.

1.2.1 Level inequalities on the hypercube

The operator $T_{1-\varepsilon}$ that defines the edges of $H_\varepsilon[n]$ can be viewed as an operators on function $f : \{0,1\}^n \to \mathbb{R}$. The function $T_{1-\varepsilon} f : \{0,1\}^n \to \mathbb{R}$ is defined by

$$(T_{1-\varepsilon} f)(x) = \mathop{\mathbb{E}}_{y \sim T_{1-\varepsilon}(x)} [f(y)].$$

The Fourier–Walsh decomposition of a function allows one to write any function $f : \{0,1\}^n \to \mathbb{R}$ as $F_0 + F_1 + \ldots + F_n$, where each component F_i is an eigenvector of $T_{1-\varepsilon}$ with eigenvalue $(1-\varepsilon)^i$, that are orthogonal to each other. Using the small-set expansion property (or the hypercontractive inequality), one can show that if $f : \{0,1\}^n \to \{0,1\}$ is the indicator function of a set of fractional size δ, then almost all of the mass in the above decomposition lies on F_i for $i \geqslant \Omega(\log(1/\delta))$ (and this in turn can be used to prove many classical results such as KKL and Friedgut's Junta theorem [29, 21]).

1.2.2 Applications of the hypercube

As discussed earlier, the noisy hypercube is a useful building block in various constructions in Theoretical Computer Science. A typical application of it is to construct PCPs, where a problem P is reduced to a problem P' using the hypercube as a gadget. A "good assignment" in P' naturally corresponds to a Boolean function $f : \{0,1\}^n \to \{0,1\}$ with local properties (such as satisfying $f(x) + f(y) = f(x+y)$ for at least $\frac{1}{2} + \eta$ fraction of the pairs $x, y \in \{0,1\}^n$, or being noise stable, i. e., satisfying $f(x) = f(y)$ with close to 1 probability when x is sampled uniformly and $y \sim T_{1-\varepsilon}(x)$), and the goal is to extract global/structural information about it. The latter corresponds to a "good assignment" to the first problem P.

Another application is to Metric Embedding, where the hypercube is folded under some group of symmetries G, so that a function $f : \{0,1\}^n/G \to \{0,1\}$ can be naturally viewed as a G-invariant function $f : \{0,1\}^n \to \{0,1\}$, from which some properties can be deduced.

[2] The Hypercontractive Inequality states that the normalized adjacency operator of $H_\varepsilon[n]$ is a contraction from $L_{p(\varepsilon)}$ to L_2 for some $p(\varepsilon) > 2$, and the small-set expansion property of the noisy hypercube makes the same assertion but only with regards to Boolean-valued functions.

In both applications, the fact that the hypercube has exponential size in n leads to deterioration of parameters (be it (1) the size of the instance in PCP constructions, or (2) the number of points in the metric for nonembeddability results), thus it could be useful to find alternatives to the hypercube that contain significantly fewer vertices. There is a known way [4] to "fold" the hypercube under some group of symmetries, so that it is significantly smaller and still useful for (some) applications [30, 10].

1.3 The Johnson graph

The Johnson graph $J(n, \ell, t)$ is a natural potential substitute to the hypercube that has significantly fewer vertices, defined for integer parameters $0 < t < \ell < n$ as follows. The vertex set of the graph is $\binom{[n]}{\ell}$, the collection of all size-ℓ subsets of $[n]$, and two vertices are adjacent if they intersect in size t:

$$E = \{(A, B) \mid |A \cap B| = t\}.$$

Denote $\ell = pn$, for some $0 < p < 1$. Works concerning the Johnson graph [41, 18, 17, 19, 20] mainly dealt with the case p is constant bounded away from 0 and 1 and $t = \ell - 1$, which is closely related to the Boolean hypercube $\{0, 1\}^n$ with the p-biased measure. Indeed, for such p analogs of the KKL Theorem, Friedgut's Junta theorem and the Majority is Stablest theorem are known. When $t = \alpha\ell$ for α close to 1 (say, $\alpha = 0.99$), the graph is closely related to the noisy hypercube $H_{(1-\alpha)}[n]$, and in particular it is a small-set expander [37].

In this paper we will only be concerned with the case $p = o(1)$ and $t = \alpha\ell$ where α is bounded away from 0 and 1. In this regime of parameters, the graph is no longer a small-set expander. For each $i \in [n]$ define $S_i = \{A \subseteq [n] \mid |A| = \ell, i \in A\}$, and note that S_i is a small set with edge expansion bounded away from 1. Indeed, it contains $\binom{n-1}{\ell-1}/\binom{n}{\ell} = p = o(1)$ fraction of the vertices in the graph, and its edge expansion is $1 - \binom{\ell-1}{t-1}/\binom{\ell}{t} = 1 - t/\ell = 1 - \alpha$.

This example (which is also known as the dictatorship function) serves as a basic building block for a wider collection of small sets with edge expansion bounded away from 1: several basic sets can be combined by unions of intersections such as $(S_1 \cap S_2) \cup (S_1 \cap S_3 \cap S_4) \cup S_5$. One can show that a union of intersections of constant width, namely in which each intersection involves only constantly many sets, has far from 1 edge expansion. Also, whenever the width of union is not too large, the set would be small. The question we address in this paper, is whether these are essentially the only small sets with exhibiting such behaviour. More precisely:

Question 1.1. Are all small sets S with expansion $\Phi(S) \leqslant 1 - \eta$ correlated with constant-width unions of intersections of basic sets?

Our results give a positive answer to this question in the case that ℓ is significantly smaller than n.

1.4 Main results

A set of vertices in $J(n, \ell, t)$ of the form $S = \{A \mid |A| = \ell, i \in A\}$ for some $i \in [n]$ is called a basic set.

Definition 1.2. Let $r \in \mathbb{N}$, $\varepsilon > 0$. A set S is called (r, ε)-*pseudorandom* if for any r basic sets $S_1, \ldots, S_r$,

$$|S \cap (S_1 \cap \ldots \cap S_r)| \leqslant \varepsilon \, |S_1 \cap \ldots \cap S_r| \, .$$

In words, an (r, ε)-pseudorandom set S contains at most ε fraction of the vertices from any width r intersection of basic subsets. We remark that any (r, ε)-pseudorandom set is also automatically a small set. With this definition, we can now state our main result.

Theorem 1.3. *For every $\alpha \in (0, 1)$, and $\eta > 0$, there are $r \in \mathbb{N}$, $\varepsilon > 0$ such that the following holds. Take large enough $\ell \geqslant \ell_0(\alpha, \eta)$, and take large enough $n \geqslant n_0(\ell)$. Then any (r, ε)-pseudorandom set S in $J(n, \ell, \alpha\ell)$ has close to 1 expansion:*

$$\Phi(S) \geqslant 1 - \eta.$$

Equivalently, this theorem asserts that a set that has expansion bounded away from 1 cannot be pseudorandom, thus there are basic sets $S_1, \ldots, S_r$ such that

$$|S \cap (S_1 \cap \ldots \cap S_r)| \geqslant \varepsilon \, |S_1 \cap \ldots \cap S_r| \, . \tag{1.1}$$

Using this, one can prove a characterization of small sets S whose expansion is bounded away from 1 as follows. Apply Theorem 1.3 to find a tuple $S_1, \ldots, S_r$ of basic sets, for which equation (1.1) holds. Defining $S' = S \setminus (S_1 \cap \ldots \cap S_r)$, one can show that the expansion of S' is not much more than the expansion of S, and thus theorem can be applied again to S' to find another tuple satisfying equation (1.1). Iterating this argument, one can extract many tuples $(S_1, \ldots, S_r)$, until a significant correlation is found between S and a width r combination of basic sets.

A technical challenge that arises when working with the Johnson graph is that it is not a product graph, and thus its eigenfunction decomposition (the analog of Fourier-Walsh decomposition) is more complicated to study. Still, any function $F \colon J(n, \ell, \alpha\ell) \to \mathbb{R}$ can be written as

$$F = F_0 + F_1 + \ldots + F_\ell, \tag{1.2}$$

where F_i is an eigenvector of the normalized adjacency operator of $J(n, \ell, \alpha\ell)$ with eigenvalue $\approx \alpha^i$. En route to proving Theorem 1.3, we prove an analog of the level inequalities of the hypercube:

Theorem 1.4 (Informal)**.** *Fix $\alpha \in (0, 1)$, let $\ell \geqslant \ell_0(\alpha)$ be large enough, and take large enough $n \geqslant n_0(\ell)$. For every $\varepsilon > 0$ there exists $r \in \mathbb{N}$, such that any set S that is (r, ε)-pseudorandom, has all but an $o(1)$ fraction of its mass on levels $i \geqslant \Omega(\log(1/\varepsilon))$.*

1.5 Related work

This paper is related to a recent line of work establishing the 2-to-2 Theorem in PCP [33, 11, 12, 34]. A key component in this line of work is a theorem similar to Theorem 1.3 for the Grassmann graph rather than the Johnson graph, which was subsequently established in [34]. The nodes of the

Grassmann graphs are ℓ-dimensional linear subspaces rather than size-ℓ sets. Edges correspond to pairs of subspaces with large intersection. The structure of the small non-expanding sets in the Grassmann graph is more complicated than in the Johnson graph, and our Theorem 1.3 was an important step towards the analogous theorem about the Grassmann graph [34].

In this article, we work directly with the Johnson graph (which is not a product graph) and this makes the spectral decomposition (1.4) somewhat complicated. To overcome this issue, we develop an explicit, approximate decomposition that is more convenient to work with. An alternative approach would have been to replace the Johnson graph with a Cayley, product graph that is closely related in terms of expansion; indeed, such approach was taken in [34] for example. We believe however, that the explicit approximate decomposition approach may be of interest in domains that have no clear product analogs, such as high-dimensional expanders [2, 25] and may have other merits down the line.

1.6 Proof overview

To prove Theorem 1.3, we use the spectral decomposition in (1.2), and as each F_i is an eigenvector of $J(n, \ell, \alpha\ell)$ of eigenvalues (roughly) α^i, one has that (after using orthonormality)

$$1 - \Phi(S) = \frac{1}{\mu(S)} \sum_{i=0}^{\ell} \alpha^i \|F_i\|_2^2.$$

It can be shown, using Parseval's identity that the contribution of $i > r$ in the above sum is at most $\alpha^{r+1}\mu(S)$, so $1 - \Phi(S) \leqslant \frac{1}{\mu(S)} \sum\limits_{i=0}^{r} \alpha^i \|F_i\|_2^2 + \alpha^{r+1}$. Taking large enough r, α^{r+1} is close to 0, and therefore to complete the proof we must show that the first sum is also small. Here, we crudely bound

$$\frac{1}{\mu(S)} \sum_{i=0}^{r} \alpha^i \|F_i\|_2^2 \leqslant r \max_{i=0,\dots,r} \frac{\|F_i\|_2^2}{\mu(S)}.$$

By Parseval, the sum over i of $\|F_i\|_2^2$ is $\mu(S)$, hence we refer to the quantity $\frac{\|F_i\|_2^2}{\mu(S)}$ as the relative spectral weight of S on level i. Thus, the proof of Theorem 1.3 boils down to showing that pseudorandom functions cannot have large weight on small levels i.

To gain some intuition, assume this is not the case, i. e., the quantity $\frac{\|F_i\|_2^2}{\mu(S)}$ is large. Then F_i is correlated with a Boolean function F; in the extreme case where instead of correlation we would have closeness, this would allow us to say that the fourth moment of F_i is roughly the same as $\|F_i\|_2^{1/2}$. Something similar (though weaker) could be said in the case of correlations, and indeed we are able to establish a lower bound on the fourth moment of F_i. To get a contradiction, we show that due to other considerations, an upper bound on the fourth moment of F_i may be established, and combining these two results gives a contradiction (or more precisely, shows that $\frac{\|F_i\|_2^2}{\mu(S)}$ must be small if S is (i, ε)-pseudorandom).

For the proof of the upper bound, we need more information regarding the structure of the level i component of S, i. e., of F_i, and as we shall see we may write this function as

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

for some function $f \colon \binom{[n]}{i} \to \mathbb{R}$ satisfying several orthogonality conditions. Thinking of $A \in \binom{[n]}{\ell}$ as being randomly chosen and then of the $f(I)$ above as random variables, if they were uncorrelated, we would expect $F_i[A]$ to be well behaved, and in particular we would expect some concentration around the mean to occur. In particular, its fourth moment would not behave like that of a Boolean function with the same ℓ_2-norm, and formalizing this shows that this would indeed contradict the lower bound we proved earlier on $\|F_i\|_4$. Thus, at least qualitatively, the task of upper bounding $\|F_i\|_4$ reduces to upper bounding correlations between the $f(I)$, and this is the heart of the matter of our argument that we explain next.

More precisely, it turns out that one has to analyze expectations of the form

$$\mathbb{E}\,[f(x_1,\ldots,x_r)f(y_1,\ldots,y_r)f(z_1,\ldots,z_r)f(w_1,\ldots,w_r)], \tag{1.3}$$

for a certain function of interest f emerging from the spectral decomposition on the Johnson graph, where the expectation is taken uniformly over $x, y, z, w \in [n]^r$ 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

$$\mathbb{E}\,[f(x_1,\ldots,x_r)f(y_1,\ldots,y_r)f(z_1,\ldots,z_r)f(w_1,\ldots,w_r)], \tag{1.4}$$

where now the expectation is uniform over $x_i, y_i, z_i, w_i \in \mathbb{F}_2^n$, 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 [12], the authors could analyze the expectations of the form (1.4) 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 prove an upper bound on 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 (where Fourier analysis on f and the notion of zoom-outs are 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 [34].

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 2.18 and its improved form Theorem 2.20.

1.7 Subsequent and future work

We end this introductory section by subsequent developments and future directions.

1. Following our work, the analogous characterization of small sets that have bounded away from 1 expansion in the Grassmann graph was proven [34].
2. The vertices of the Johnson graph form the Boolean slice of the hypercube, which is a well-studied object from a combinatorial point of view [18, 17, 19, 20]. It is related to the study of sharp thresholds of Boolean functions and graph properties [22]. Subsequent to our work, some progress in this direction [31] was made, including an improved quantitative version of Bourgain's Sharp Threshold Theorem [22].
3. Our work has also inspired the study of the complexity of Unique Games over graphs that exhibit characterizations of small sets that violate the small expansion property [1, 3]. This class extends the class of certifiable small-set expanders, on which Unique Games are known to be polynomial time solvable.

We next move on to a few open directions that stem from our work.

1. Many results in the PCP literature [26], construction of integrality gaps and non-embeddability results [36, 35] rely on the small-set expansion property of the (noisy) hypercube. 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 hypercube like graphs (e. g., that have a "small-set expansion" type property) with significantly smaller number of vertices. We believe that our results can be used to get improved constructions to the above problems. This task is not trivial since the Johnson graph is not a small-set expander, but one could hope to instead use the characterization of sets violating the small-set expansion property given herein. The outer PCP of [33, 12] is an example in which a characterization of this sort is used.
2. For the Grassmann graph, the characterization of small sets that have non-perfect expansion implies a "direct product testing" result, i. e., an encoding scheme and a set of (2-to-1) tests on words, such that any word that passes a notable fraction of the tests must have a global structure. Our results are related, in the same sense, to the problem of direct product testing on the hypercube [28, 14, 7]. It is interesting to note that the notion of "basic sets" appears in many of these works, and often as an intermediate step (if there are more than 2 queries).

2 Preliminaries

In this section we present the necessary background on the Johnson graph.

2.1 Notation.

We shall use the big O notation: for nonnegative-valued functions $f, g : \mathbb{N} \to \mathbb{R}$, by $f = O(g)$ we mean that $f(n) \leqslant C \cdot g(n)$ for some constant $C > 0$ and all sufficiently large n, and by $f = \Omega(g)$ we mean that $f(n) \geqslant c \cdot g(n)$ for some constant $c > 0$ and all sufficiently large n. When we write expectations such as $\mathbb{E}_A[F[A]]$, unless specified otherwise, we mean that A is sampled uniformly from the domain of F.

Definition 2.1. Let $t < \ell < n$ be integers. The *generalized Johnson graph* $J(n, \ell, t)$ is defined as follows: the vertex set is $\binom{[n]}{\ell}$, the set of all subsets of $[n]$ of size ℓ. Two vertices $A, B \subseteq [n]$ are adjacent if $|A \cap B| = t$.

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

2.2 Fourier analysis on the generalized Johnson graph

Any undirected regular graph, and in particular $J(n, \ell, t)$, induces a spectral decomposition of real-valued functions on it. Spectral decompositions on the Johnson graph, and more generally in association schemes have been studied in the work of Delsarte [9, 8]. In this section, we briefly present such decomposition (similar to the one used in the context of the Grassmann graph [12]); we refer the reader to [23] for a more thorough study.

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

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

for any $F, G\colon \binom{[n]}{\ell} \to \mathbb{R}$. We denote the average of a function $F\colon \binom{[n]}{\ell} \to \mathbb{R}$ as $\mu(F) = \mathbb{E}_{A \in \binom{[n]}{\ell}} [F[A]]$.

2.3 Level functions

Definition 2.2. Let $t < \ell < n$ be integers. For any $i = 0, \ldots, \ell$ 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 $f\colon \binom{[n]}{i} \to \mathbb{R}$ such that for all $A \in \binom{[n]}{\ell}$

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

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

Definition 2.3. We define the space of level i functions by $J_{=i} = J_{\leqslant i} \cap J_{\leqslant i-1}^{\perp}$. 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 $J(n, \ell, t)$ can be written as $J_{=0} \oplus J_{=1} \oplus \ldots \oplus J_{=\ell}$.

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

$$\lambda_i(n, \ell, t) = \frac{\sum\limits_{r=\ell-t}^{\ell} (-1)^{r-(\ell-t)+i} \binom{r}{\ell-t} \binom{n-2r}{\ell-r} \binom{n-r-i}{r-i}}{\binom{\ell}{t} \binom{n-\ell}{\ell-t}}.$$

A standard fact (see for example [23, Theorem 6.5.2]) asserts that the $\lambda_i(n, \ell, t)$ are the eigenvalues of $J(n, \ell, t)$, and furthermore one has:

Theorem 2.5. *Let $t < \ell < n$ be integers. For any $i = 0, \ldots, \ell$, and $F \in J_{=i}$, we have $J(n, \ell, t)F = \lambda_i(n, \ell, t)F$. Moreover, $\mathsf{dim}(J_{=i}) = \binom{n}{i} - \binom{n}{i-1}$.*

The following notation will be convenient.

Definition 2.6. Let $j < i < n$ be integers and $f \colon \binom{[n]}{i} \to \mathbb{R}$. For $J \subseteq [n]$, $|J| = j$ we denote

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

The following simple claim will be used extensively:

Claim 2.7. *Let $J' \subseteq J$. Then $1_{B \cap J = J'} = \sum\limits_{J'' : J' \subseteq J'' \subseteq J} (-1)^{|J'' \setminus J'|} 1_{B \supseteq J''}$.*

Proof. Note that as the right hand side is equal to $1_{B \supseteq J'} \sum\limits_{J'' \subseteq J \setminus J'} (-1)^{|J''|} 1_{B \supseteq J''}$, it suffices to prove the statement for $J' = \emptyset$. For this, we note that $1_{B \cap J = \emptyset} = \prod\limits_{j \in J} \left(1 - 1_{B \ni j}\right)$, and the result follows by expanding this out. □

We have the following lemma:

Lemma 2.8. *Suppose $n \geqslant 10\ell^2$, and let $F \in J_{\leqslant i}$ be given by $F[A] = \sum\limits_{I \subseteq A} f(I)$. Then $F \in J_{=i}$ if and only if for every $R \subseteq [n]$, $|R| < i$ we have that $\mu_R(f) = 0$.*

Proof. For the $\Leftarrow$ direction, let $r \leqslant i - 1$ and let R be of size r. Then

$$\langle F, 1_{A \supseteq R} \rangle = \sum_I f(I) \mathop{\mathbb{E}}_{A} [1_{A \supseteq R \cup I}] = \sum_I f(I) p_{|R \cup I|},$$

where $p_m = \frac{\binom{n-m}{\ell-m}}{\binom{n}{\ell}}$ for $m \leqslant \ell$ and 0 otherwise. Thus,

$$\langle F, 1_{A \supseteq R} \rangle = \sum_{r'=0}^{r} p_{r+i-r'} \sum_{\substack{R' \subseteq R \\ |R'| = r'}} \sum_I f(I) 1_{I \cap R = R'}.$$

Using Claim 2.7, we plug in $1_{I\cap R=R'} = \sum_{R'\subseteq R''\subseteq R}(-1)^{|R''\setminus R'|}1_{I\supseteq R''}$ and get that

$$\langle F, 1_{A\supseteq R}\rangle = \sum_{r'=0}^{r} p_{r+i-r'} \sum_{\substack{R'\subseteq R\\ |R'|=r'}} \sum_{R'\subseteq R''\subseteq R} (-1)^{|R''\setminus R'|} \mathop{\mathbb{E}}_{I}[f(I)1_{I\supseteq R''}].$$

As $\mathbb{E}_I[f(I)1_{I\supseteq R''}]$ is proportional to $\mu_{R''}(f)$ and $|R''| \leqslant |R| \leqslant r \leqslant i-1$, all of the above expectations are 0.

The $\Rightarrow$ direction. We prove by induction on $r = 0, 1, \ldots, i-1$ that if $F \in J_{=i}$ is given as in the lemma, and $R \subseteq [n]$ has size r, then $\mu_R(f) = 0$. The base case, $r = 0$, follows since $\mu(f) = \frac{\mu(F)}{\binom{\ell}{i}} = 0$, as $\mu(F) = 0$.

Let $r \leqslant i-1$, assume the statement for $r' \leqslant r-1$, and prove for r. As in the previous direction

$$0 = \langle F, 1_{A\supseteq R}\rangle = \sum_{r'=0}^{r} p_{r+i-r'} \sum_{\substack{R'\subseteq R\\ |R'|=r'}} \sum_{R'\subseteq R''\subseteq R} (-1)^{|R''\setminus R'|} \mathop{\mathbb{E}}_{I}[f(I)1_{I\supseteq R''}].$$

As $\mathbb{E}_I[f(I)1_{I\supseteq R''}]$ is proportional to $\mu_{R''}(f)$, by induction hypothesis it is 0 if $|R''| < r$, so we get that only the case that $R'' = R$ contributes to the above sum, so

$$0 = \mathop{\mathbb{E}}_{I}[f(I)1_{I\supseteq R}] \sum_{r'=0}^{r} p_{r+i-r'} \sum_{\substack{R'\subseteq R\\ |R'|=r'}} (-1)^{|R\setminus R'|} = \mathop{\mathbb{E}}_{I}[f(I)1_{I\supseteq R}] \sum_{r'=0}^{r} p_{r+i-r'} \binom{r}{r'} (-1)^{r-r'},$$

Thus, since $\mathbb{E}_I[f(I)1_{I\supseteq R}]$ is proportional to $\mu_R(f)$, to show that $\mu_R(f) = 0$ it suffices to argue that the above sum is non-zero. This is true since $p_{r+i-r'}\binom{r}{r'}$ is an increasing function of r': we have $p_{r+i-r'-1} \geqslant \frac{n}{\ell} p_{r+i-r'}$, $\binom{r}{r'+1} \geqslant \binom{r}{r'}\frac{1}{r}$ and $\frac{n}{r\ell} \geqslant \frac{n}{\ell^2} \geqslant 10$, so

$$\left|\sum_{r'=0}^{r} p_{r+i-r'} \binom{r}{r'} (-1)^{r'}\right| \geqslant p_i - \sum_{0\leqslant r'<r} \binom{r}{r'} p_{r+i-r'} \geqslant p_i - \sum_{0\leqslant r'<r} \frac{1}{10^{r-r'}} p_i \geqslant \frac{p_i}{2} > 0. \qquad \square$$

2.4 Approximate eigenvalues

The formula for $\lambda_i(n, \ell, t)$ is quite complex, and it is not even clear from it what is the order of magnitude of $\lambda_i(n, \ell, t)$. Thus, we shall use the following approximations of $\lambda_i(n, \ell, t)$:

Definition 2.9. Let $t < \ell < n$ be integers. For $i = 0, \ldots, t$ define $\lambda_{\approx i}(n, \ell, t) = \frac{\binom{t}{j}}{\binom{\ell}{j}}$, and for $t < i \leqslant \ell$ define $\lambda_{\approx i}(n, \ell, t) = 0$.

Lemma 2.10. *For all $0 \leqslant j \leqslant \ell$, $\left|\lambda_j(n, \ell, t) - \lambda_{\approx j}(n, \ell, t)\right| \leqslant \frac{\ell}{n-\ell}$.*

Proof. For $j = 0$, a direct computation shows that $\lambda_{\approx 0}(n,\ell,t) = \lambda_0(n,\ell,t) = 1$.

For $0 < j \leqslant \ell$, set $J = \{1,\dots,j\}$, define $G[A] = 1_{A\supseteq J}$ and let F be the projection of G onto $V_{=j}$. Then by Theorem 2.5

$$\lambda_j(n,\ell,t)\mu_J(F) = \mathop{\mathbb{E}}_{A\supseteq J}\left[\lambda_j(n,\ell,t)F[A]\right] = \mathop{\mathbb{E}}_{A\supseteq J}\left[J(n,\ell,t)F[A]\right] = \mathop{\mathbb{E}}_{A\supseteq J}\left[\mathop{\mathbb{E}}_{|B\cap A|=t}[F[B]]\right]$$
$$= \sum_{J'\subseteq J}\mathop{\mathbb{E}}_{\substack{A\supseteq J,B\\|B\cap A|=t}}\left[F[B]1_{B\cap J=J'}\right].$$

Note that conditioned on $B\cap J = J'$, the distribution over B is uniform over all ℓ-size subsets intersecting J in J', so the last sum is equal to

$$\sum_{J'\subseteq J} p_{J'}\mathop{\mathbb{E}}_{B\cap J=J'}[F[B]],$$

where $p_{J'}$ is the probability that $B\cap J = J'$. We note that $p_J = \lambda_{\approx j}$, and we next bound the expectation for other subsets J'. Let $q_{J'} = \Pr_A[A\cap J = J']$. Then

$$\mathop{\mathbb{E}}_{A\cap J=J'}[F[A]] = \frac{\mathbb{E}_A\left[F[A]1_{A\cap J=J'}\right]}{q_{J'}} = \sum_{J'\subseteq J''\subseteq J}(-1)^{|J''\setminus J'|}\frac{\mathbb{E}_A\left[F[A]1_{A\supseteq J''}\right]}{q_{J'}},$$

where we used Claim 2.7. Since $F\in V_{=j}$ and $1_{A\supseteq J''}\in V_{\leqslant |J''|}$, the expectation in the numerator is 0 unless $J'' = J$, so

$$\mathop{\mathbb{E}}_{A\cap J=J'}[F[A]] = \frac{\mathbb{E}_A\left[F[A]1_{A\supseteq J}\right]}{q_{J'}} = \frac{q_J}{q_{J'}}\mu_J(F).$$

Plugging this above, we conclude that

$$\lambda_j(n,\ell,t)\mu_J(F) = \sum_{J'\subseteq J} p_{J'}\frac{q_J}{q_{J'}}\mu_J(F),$$

so $\lambda_j(n,\ell,t) = \sum_{J'\subseteq J} p_{J'}\frac{q_J}{q_{J'}}$. Thus,

$$\left|\lambda_j(n,\ell,t) - \lambda_{\approx j}(n,\ell,t)\right| = \sum_{J'\subsetneq J} p_{J'}\frac{q_J}{q_{J'}} \leqslant \max_{J'\subsetneq J}\frac{q_J}{q_{J'}} = \max_{r\leqslant j-1}\frac{\binom{n-j}{\ell-j}}{\binom{n-j}{\ell-r}}.$$

The maximum is attained at $r = j-1$, and by direct computation is equal to $\frac{\ell-j+1}{n-\ell}$. $\square$

2.5 Decomposition and approximate decomposition

2.5.1 Decomposition

Recall that we have seen that the space of all real-valued functions on $\binom{[n]}{\ell}$ can be decomposed as $J_{=0}\oplus J_{=1}\oplus\ldots\oplus J_{=\ell}$. For any function $F\colon\binom{[n]}{\ell}\to\mathbb{R}$, we denote this decomposition by

$F = F_{=0} + F_{=1} + \ldots + F_{=\ell}$ where $F_{=i} \in J_{=i}$. We also define $f_{=i} \colon \binom{[n]}{i} \to \mathbb{R}$ to be a function that satisfies

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

We remark that by dimension considerations, the choice of the function $f_{=i}$ is unique. Indeed, any function in $J_{\leqslant i}$ admits such representation, and as the dimension of $J_{\leqslant i}$ is $\binom{n}{i}$, which is the same as the dimension of the space of functions G that can be written as $\sum_{I \subseteq A,\ |I|=i} g(I)$ for some $g \colon \binom{[n]}{i} \to \mathbb{R}$. We also remark that by Lemma 2.8, the function $f_{=i}$ automatically satisfies several orthogonality conditions.

The notion of level i weight of F will be important for us:

Definition 2.11. Let $F \colon \binom{[n]}{\ell} \to \mathbb{R}$ be a function, and let $i \in \{0, 1, \ldots, \ell\}$. We define the weight of F on level i to be

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

Computing the spectral decomposition. It is easy to verify that $F_{=0} \equiv \mu(F)$, and it is not very hard to get an explicit formula for $F_{=1}$. As i gets larger and larger though, getting a convenient, precise formula for $F_{=i}$ is more challenging. Thus, in the next section we define approximate versions of these functions.

2.5.2 Approximate decomposition

Next, we introduce the approximate decomposition and approximate eigenvalues that will be easier for us to work with. Given a function $F \colon \binom{[n]}{\ell} \to \mathbb{R}$, define $f_{\approx 0} \equiv \mu(F)$. Inductively once $f_{\approx j}$ have been defined for all $j < i$, we define $f_{\approx i} \colon \binom{[n]}{i} \to \mathbb{R}$ by

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

for all $I \in \binom{[n]}{i}$. We then define $F_{\approx i} \colon \binom{[n]}{\ell} \to \mathbb{R}$ by

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

To work with $f_{\approx i}, F_{\approx i}$ instead of $f_{=i}$ and $F_{=i}$, we will show that they are close in ℓ_2 distance.

Theorem 2.12. *There is an absolute constant $C > 0$ such that the following holds. Let $n \geqslant \ell \geqslant i$ be integers such that $n \geqslant 2^{3i^2+iC}\ell^{2i+1}$, and let $F \colon \binom{[n]}{\ell} \to \mathbb{R}$ be a function. Then*

$$\|F_{=i} - F_{\approx i}\|_2^2 \leqslant \frac{2^{4i^2+iC}\ell^{2i}}{n}\|F\|_\infty^2.$$

Proof. Deferred to Appendix A.2. □

For our argument to go through, we will need a few more facts about $f_{\approx i}$, $F_{\approx i}$. Note that by Claim 2.8, we have that $\mu_R(f_{=i}) = 0$ for all $|R| < i$. The following fact asserts that this is also approximately the case for $f_{\approx i}(I)$.

Fact 2.13. *Let ℓ, n be integers such that $n \geqslant 2\ell^2$, and let $F\colon \binom{[n]}{\ell} \to \mathbb{R}$ be a function. Then for all $i = 1, \ldots, \ell$, $0 \leqslant j < i$ and $J \subseteq [n]$ of size j,*

$$\left|\mu_J(f_{\approx i})\right| \leqslant 2^{3i^2+Ci}\frac{\ell}{n}\|F\|_\infty,$$

where $C > 0$ is some absolute constant.

Proof. Deferred to Appendix A.1. □

The following fact asserts that if F is bounded, then so is $f_{\approx i}$.

Fact 2.14. *Let $\ell \leqslant n$ be integers, and let $F\colon \binom{[n]}{\ell} \to \mathbb{R}$ be a function. Then for all $i = 0, \ldots, \ell$, $\|f_{\approx i}\|_\infty \leqslant 2^{i^2}\|F\|_\infty$.*

Proof. We prove this by induction on i. Note that by definition of $f_{\approx i}$, we have that

$$\|f_{\approx i}\|_\infty \leqslant \|F\|_\infty + \sum_{j=0}^{i-1}\binom{i}{j}\|f_{\approx j}\|_\infty.$$

Thus, defining $a_i = \max(\|F\|_\infty, \max_{0\leqslant j\leqslant i}\|f_{\approx j}\|_\infty)$, the above inequality gives the recurrence $a_i \leqslant a_{i-1} + (2^i - 1)a_{i-1} = 2^i a_{i-1}$. Hence $a_i \leqslant 2^{\sum_{j=1}^{i} j} a_0 \leqslant 2^{i^2}\|F\|_\infty$ as $a_0 \leqslant \|F\|_\infty$. □

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 $f_{\approx i,F}$.

2.6 Restrictions

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

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

The following definition extends the notion of pseudorandom sets from the introduction to functions.

Definition 2.15. Let $r \in \mathbb{N}$ and $\varepsilon > 0$. For $\ell > r$, a function $F\colon \binom{[n]}{\ell} \to \mathbb{R}$ is called (r, ε)-*pseudorandom* if for any set $X \subseteq [n]$ of size at most r, $\|F|_X\|_2^2 \leqslant \varepsilon$.

In this language, the definition of pseudorandom sets is equivalent to saying that the indicator function of the set is pseudorandom as per the definition above.

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

$$f_{\approx i+1,F}(I \cup \{x\}) = \mu_{I\cup\{x\}}(F) - \sum_{J\subseteq I} f_{\approx |J|,F}(J) - \sum_{J\subsetneq I} f_{\approx |J|+1,F}(J \cup \{x\}). \tag{2.1}$$

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 (2.1) 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 2.17. *Let $F\colon \binom{[n]}{\ell} \to \mathbb{R}$ be function, $0 \leqslant i \leqslant \ell - 1$ and $X \subseteq [n]$, $|X| < i$. Then for any $I \subseteq [n] \setminus X$ 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 sets X 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 2.16

$$f_{\approx i,F}(I \cup X) = f_{\approx i-1,F|_{\{x\}}}(I \cup X') - f_{\approx i-1,F}(I \cup X'). \tag{2.2}$$

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 (2.2) completes the proof. □

2.7 Main results

In this section we state our main results. Our first result is Theorem 1.4 from the introduction, which in a more precise form states:

Theorem 2.18. *There exists $C > 0$ such that the following holds. Suppose $n \geqslant 2^{8i^2+iC}\ell^{\frac{20i}{3}}\frac{1}{\delta^4}$. If S has density δ and is (r, ε)-pseudorandom, then for every $i = 0, \ldots r$,*

$$W^{=i}[1_S] \leqslant e^{O(i)}\delta\varepsilon^{1/3} + \frac{1}{n^{1/24}}.$$

We prove this theorem in Section 3. A quick corollary of the above result is Theorem 1.3 from the introduction:

Theorem 2.19. *Let $\alpha \in (0, 1)$, and S be a subset of vertices in $J(n, \ell, \alpha\ell)$ of density δ. Let $r \in \mathbb{N}$, $\varepsilon > 0$, and suppose $n \geqslant 2^{8r^2+rC}\ell^{\frac{20r}{3}}\frac{1}{\delta^4}$ and that S is (r, ε)-pseudorandom. Then*

$$\Phi(S) \geqslant 1 - \alpha^{r+1} - e^{O(r)}\varepsilon^{1/4} - \frac{r}{n^{1/24}\delta}.$$

The proof is given in Section 4. Finally, we prove the following quantitative improvement of Theorem 2.18 in Section 5.

Theorem 2.20. *There exists $C > 0$, such that the following holds for $r, m \in \mathbb{N}$ and $\delta > 0$ satisfying $n \geqslant (2^{3r^2+rC}\ell^{2r})^{4m(2m-1)^2}\delta^{-4}$. If S is (r, ε)-pseudorandom of density δ, then for every $i = 0, \ldots, r$,*

$$W^{=i}[1_S] \leqslant (Cm)^{2i}\delta\varepsilon^{1-\frac{1}{2m-1}} + \frac{2^{3i^2+iC}\ell^{2i}}{n^{1/(4m(2m-1)^2)}}.$$

In words, the above theorem asserts that for constant r, an (r, ε)-pseudorandom set S can have at most $\delta\varepsilon^{1-o(1)}$ of its weight on the first r levels. Ignoring the error terms, the best pseudorandomness one may hope for, in terms of ε, is that $\varepsilon = O(\delta)$, namely that no small restrictions increase the density of S by more than a constant factor. In that case, while Theorem 2.18 asserts that the weight of S on level i is small compared to the measure of S (more precisely, it bounds it by $O_i(\delta^{4/3})$), Theorem 2.20 shows a much stronger bound, and in fact that the weight on level i is not much larger than the weight of level 0. While the latter quantity is clearly equal to δ^2, Theorem 2.20 gives a bound of $\delta^2 \log(1/\delta)^{O(i)}$ on the former quantity in the case that $\varepsilon = O(\delta)$ (by picking $m \approx \log(1/\delta)$). This statement should be compared to the r-level inequalities on the Boolean hypercube, stating that if $f: \{-1, 1\}^n \to \{0, 1\}$ has density δ, then its Fourier weight on level r is at most $\delta^2 O(\log(1/\delta))^r$.

3 Proof of Theorem 2.18

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

Let S be an (r, ε)-pseudorandom set of density δ. Let $F\colon J(n, \ell) \to \{0, 1\}$ be the indicator function of S, and let $i \in \{0, \ldots, r\}$. 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; indeed, then $W_{=i}[F] = \eta \leqslant \delta^2 \leqslant \delta\varepsilon^{1/3}$, as $\varepsilon \geqslant \mu(F) = \delta$. Thus, we have $\mathbb{E}_A\left[F_{=i}[A]^2\right] = \eta$, and by orthogonality $\mathbb{E}_A\left[(F - F_{=i})[A]^2\right] = \delta - \eta$.

Throughout, $C > 0$ is an absolute constant (that may not be the same in different occurrences), and we assume that $n \geqslant 2^{8i^2+iC}\ell^{\frac{20i}{3}}\frac{1}{\delta^4}$.

3.1 Information about the second moment

Claim 3.1. $\mathbb{E}_A\left[F_{\approx i}^2[A]\right] = \eta \pm \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}}$.

Proof. By the triangle inequality and Theorem 2.12, $|\|F_{=i}\|_2 - \|F_{\approx i}\|_2| \leqslant \|F_{=i} - F_{\approx i}\|_2 \leqslant \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}}$. Multiplying the inequality by $|\|F_{=i}\|_2 + \|F_{\approx i}\|_2|$, we get that

$$\left|\|F_{=i}\|_2^2 - \|F_{\approx i}\|_2^2\right| \leqslant \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}}\,|\|F_{=i}\|_2 + \|F_{\approx i}\|_2| \leqslant \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}}\left(2\|F_{=i}\|_2 + \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}}\right) \leqslant \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}},$$

the last inequality follows as $\|F_{=i}\|_2 \leqslant \|F\|_2 \leqslant 1$ by Parseval. □

Claim 3.2. *We have*

$$\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right] = \frac{W^{=i}[F]}{\binom{\ell}{i}} \pm \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}} = \frac{\eta}{\binom{\ell}{i}} \pm \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}.$$

Proof. Expand out $\mathbb{E}_A\left[F_{\approx i}^2[A]\right]$. On the one hand it is equal to $\eta \pm \frac{2^{2i^2+iC}\ell^i}{\sqrt{n}}$ by the previous claim and the condition on n. On the other hand, it is equal to

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

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

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

Fix I in the outside expectation, and consider the inner expectation. We may write it as

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

Fix I and D, consider the distribution over I' where $I' \cap I = D$; we compare this distribution to the uniform distribution over all I' containing D. We note that conditioned on I' containing D, the probability that $I' \cap I \neq D$ is at most $\frac{i^2}{n}$; moreover, conditioned on this event the distribution over I' is the same as in the above expectation. Therefore, the distribution over I' in the above expectation is $\frac{i^2}{n}$ close in statistical distance to the uniform distribution over I' containing D, and so we have that

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

Applying Facts 2.13 and 2.14 we get that the expectation in absolute value is at most $\frac{2^{5i^2+iC}}{n}$, for some absolute constant $C > 0$. Plugging it into the first equation in the proof, we see that

$$\left|\mathop{\mathbb{E}}_{A}\left[F_{\approx i}^2[A]\right] - \binom{\ell}{i}\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right]\right| \leqslant \sum_{d=0}^{i-1}\binom{\ell}{d}\binom{\ell-d}{i-d}\binom{\ell-i}{i-d}\frac{2^{5i^2+iC}}{n} \leqslant \frac{\ell^{3i}2^{5i^2+iC}}{n},$$

and therefore

$$\mathop{\mathbb{E}}_{I}\left[f_{\approx i}(I)^2\right] = \frac{1}{\binom{\ell}{i}}\mathop{\mathbb{E}}_{A}\left[F_{\approx i}^2[A]\right] \pm \frac{1}{\binom{\ell}{i}}\frac{\ell^{3i}2^{5i^2+iC}}{n} \pm \frac{1}{\binom{\ell}{i}}\frac{2^{2i^2+iC}\ell^i}{\sqrt{n}} = \frac{\eta}{\binom{\ell}{i}} \pm \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}.$$ □

We end this section with the following two corollaries, that establish upper bounds on the second moments of the level components of restrictions of F. We remark that this is the only form in which the pseudorandomness of F is used in subsequent proofs.

Corollary 3.3. *Let $0 \leqslant b \leqslant i$ be integers, and let $B \subseteq [n]$ be of size at most b. Then*

$$\mathop{\mathbb{E}}_{|D|=i-b}\left[f_{\approx i-b, F|_B}(D)^2\right] \leqslant \frac{i^{i-b}\varepsilon}{\ell^{i-b}} + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}.$$

Proof. By Claim 3.2,

$$\mathop{\mathbb{E}}_{|D|=i-b}\left[f_{\approx i-b,F|_B}(D)^2\right] \leqslant \frac{W^{=i-b}[F|_B]}{\binom{\ell-b}{i-b}} \pm \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}.$$

Note that since F is (r,ε)-pseudorandom, $W^{=i-b}[F|_B] \leqslant \mu(F|_B) \leqslant \varepsilon$. Using the bound $\binom{\ell-b}{i-b} \geqslant (\ell/i)^{i-b}$ completes the proof. □

Corollary 3.4. *Let $0 \leqslant b \leqslant i$ be integers, and let $B \subseteq [n]$ be of size b. Then*

$$\mathop{\mathbb{E}}_{\substack{D\subseteq[n]\setminus B\\|D|=i-b}}\left[f_{\approx i,F}(B\cup D)^2\right] \leqslant \frac{2^{2i+1}i^{i-b}\cdot\varepsilon}{\ell^{i-b}} + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}},$$

where $C > 0$ is some absolute constant.

Proof. By Corollary 2.17

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

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

$$\mathop{\mathbb{E}}_{\substack{D\subseteq[n]\setminus B\\|D|=i-b}}\left[f_{\approx i,F}(B\cup D)^2\right] \leqslant 2^{|B|}\sum_{Y\subseteq B}\mathop{\mathbb{E}}_{\substack{D\subseteq[n]\setminus B\\|D|=i-b}}\left[f_{\approx i-b,F|_{B\setminus Y}}(D)^2\right].$$

Fix $Y \subseteq B$, and note that sampling $D \subseteq [n]\setminus Y$ yields a D that is disjoint from B with probability at least $1 - i^2/n \geqslant 1/2$; conditioned on that, the distribution of D is uniform among the subsets of $[n]\setminus Y$ of size $i-b$. Thus,

$$\mathop{\mathbb{E}}_{\substack{D\subseteq[n]\setminus B\\|D|=i-b}}\left[f_{\approx i-b,F|_{B\setminus Y}}(D)^2\right] \leqslant 2\mathop{\mathbb{E}}_{\substack{D\subseteq[n]\setminus Y\\|D|=i-b}}\left[f_{\approx i-b,F|_{B\setminus Y}}(D)^2\right],$$

and plugging this bound yields that

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

□

3.2 A lower bound on the fourth moment

Claim 3.5. *If $n \geqslant \ell^{2i} 2^{4i^2+iC} \delta^{-4}$, then $\mathbb{E}_A \left[F^4_{\approx i}[A]\right] \geqslant \frac{\eta^4}{16\delta^3}$.*

Proof. Note that

$$\eta = \langle F, F_{=i} \rangle = \langle F, F_{\approx i} \rangle + \langle F, F_{=i} - F_{\approx i} \rangle.$$

By Cauchy–Schwarz and Theorem 2.12 we have

$$\langle F, F_{=i} - F_{\approx i} \rangle \leqslant \|F\|_2 \|F_{=i} - F_{\approx i}\|_2 \leqslant \sqrt{\delta} \frac{2^{2i^2+iC} \ell^i}{\sqrt{n}} \leqslant \frac{\eta}{2}$$

by assumption on n, and $\eta \geqslant \delta^2$. Thus, $\langle F, F_{\approx i} \rangle \geqslant \eta/2$. Using Hölder's inequality we get that

$$\langle F, F_{=i} \rangle \leqslant \|F\|_{4/3} \|F_{\approx i}\|_4 \leqslant \delta^{3/4} \|F_{\approx i}\|_4,$$

and combining we get that $\|F_{\approx i}\|_4^4 \geqslant \frac{\eta^4}{16\delta^3}$. □

3.3 An upper bound on the fourth moment

The following lemma is the main technical result of this section. For the proof, we first define the notion of intersection patterns. For $A = \{x_1, \dots, x_\ell\}$ thought of as a set of ℓ symbols, let

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

and note that $|D(A,d)|$ depends only on ℓ, i, d — denote it by $\beta_{i,d,\ell}$. Fix d and a tuple $(I_1, \dots, I_4) \in D(A,d)$. Consider a 15-dimensional vector $\vec{\sigma}_d$, that has a coordinate for any non-empty $T \subseteq \{1,2,3,4\}$, whose value on coordinate T is equal to the number of elements in $\cap_{i \in T} I_i$. We refer to this vector $\vec{\sigma}_d$ as the *intersection pattern* of (I_1, I_2, I_3, I_4).

Lemma 3.6. *Suppose $n \geqslant 2^{4i^2+iC} \ell^{2i+1} \eta^{-2}$; if F is (i, ε)-pseudorandom, then*

$$\mathop{\mathbb{E}}_{A} \left[F^4_{\approx i}[A]\right] \leqslant \mathrm{e}^{O(i)} \varepsilon \eta + \frac{2^{2i^2+iC} \ell^{5i}}{n^{1/4}}.$$

Proof. By opening the brackets we have

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

Thus, using our notation we get that

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

where by the expectation on A we mean that the symbols $x_1, \dots, x_\ell$ are sampled uniformly so that A is uniformly chosen from $\binom{[n]}{\ell}$.

We note that by symmetry, the probability a tuple $(I_1, \dots, I_4)$ is sampled depends only on its intersection pattern. We denote by $\beta(\vec{\sigma}_d)$ the probability that a specific intersection pattern $\vec{\sigma}_d$ is chosen, and by $\gamma(\vec{\sigma}_d)$ the distribution over the tuples $(I_1, \dots, I_4)$ that have this intersection pattern. Then

$$\mathop{\mathbb{E}}_{A}\left[F^4_{\approx i}[A]\right] = \sum_{d=i}^{4i}\sum_{\vec{\sigma}_d}\beta(\vec{\sigma}_d)\beta_{i,d,\ell}\mathop{\mathbb{E}}_{A}\left[\mathop{\mathbb{E}}_{\substack{(I_1,I_2,I_3,I_4)\sim\gamma(\vec{\sigma}_d)\\ I_1,\dots,I_4\subseteq A}}\left[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)\right]\right]. \tag{3.1}$$

Note that now the distribution over $(I_1, \dots, I_4)$ is uniform over all quadruples whose union has size d and its intersection pattern is $\vec{\sigma}_d$, and thus

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

The following lemma is key in completing the proof of Lemma 3.6.

Lemma 3.7. *There is $C > 0$ such that the following holds. Suppose $n \geqslant 2^{4i^2+iC}\ell^{2i+1}\eta^{-2}$. Let i, d be integers such that $i \leqslant d \leqslant 4i$ and let $\vec{\sigma}_d$ be any intersection pattern. Then*

$$\left|\mathop{\mathbb{E}}_{(I_1,I_2,I_3,I_4)\sim\gamma(\vec{\sigma}_d)}\left[f_{\approx i}(I_1)f_{\approx i}(I_2)f_{\approx i}(I_3)f_{\approx i}(I_4)\right]\right| \leqslant \frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}}.$$

We defer the proof of Lemma 3.7 to the next section and show how to complete the proof of Lemma 3.6 based on it. Using Lemma 3.7 we see that

$$(3.1) \leqslant \sum_{d=i}^{4i}\sum_{\vec{\sigma}_d}\beta(\vec{\sigma}_d)\beta_{i,d,\ell}\left(\frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}}\right).$$

Note that

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

and so by the previous inequality

$$(3.1) \leqslant \sum_{d=i}^{4i}\sum_{\vec{\sigma}_d}\beta(\vec{\sigma}_d)\ell^d d^{-d}(4\mathrm{e}^2)^{4i}\left(\frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}}\right) \leqslant \mathrm{e}^{O(i)}\varepsilon\eta + \frac{2^{2i^2+iC}\ell^{5i}}{n^{1/4}}.$$

□

3.4 Wrapping things up

Proof of Theorem 2.18. Combining Claim 3.5 and Lemma 3.6 we see that

$$\frac{\eta^4}{16\delta^3} \leqslant \mathop{\mathbb{E}}_{A}\left[F^4_{\approx i}[A]\right] \leqslant \mathrm{e}^{O(i)}\varepsilon\eta + \frac{2^{2i^2+iC}\ell^{5i}}{n^{1/4}}.$$

Rearranging we see that

$$\eta \leqslant \mathrm{e}^{O(i)}\delta\varepsilon^{1/3} + \frac{2^{\frac{2i^2}{3}+iC}\ell^{\frac{5i}{3}}}{n^{\frac{1}{12}}}\frac{\delta}{\eta^{1/3}} \leqslant \mathrm{e}^{O(i)}(i)\delta\varepsilon^{1/3} + \frac{1}{n^{1/24}},$$

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

3.5 Proof of Lemma 3.7

Proof. Let $x_1, ..., x_d$ be chosen independently and uniformly from $[n]$. Thinking of $x_1, \dots, x_d$ as symbols, we let $P_1, ..., P_4 \subseteq \{x_1, \dots, x_d\}$ be fixed sets whose intersection pattern is $\vec{\sigma}_d$. Namely, each one of P_1, P_2, P_3 and P_4 is as a set containing i of the symbols $x_1, \dots, x_d$, and the intersection pattern of these sets is exactly $\vec{\sigma}_d$. We would like to consider the value of $f_{\approx i}$ on the P_i however for a specific choice of x there is some small probability that P_i is a set of smaller size. We extend the definition of $f_{\approx i}(I)$ to be 0 if $|I| < i$. Thus, we note that

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

since the distributions of the sets I and the sets P are $\frac{d^2}{n}$-close to each other in statistical distance (as they are only different if some repetition occurs in one of the P_i). Therefore, it suffices to show an upper bound of the expectation involving the P_i.

Proposition 3.8. *If $\vec{\sigma}_d$ is such that there is an x_j that appears in only one of the P_i, then*

$$\left|\mathop{\mathbb{E}}_{x_1,\dots,x_d}\left[f_{\approx i}(P_1)f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\right]\right| \leqslant \frac{2^{3i^2+Ci}d^2}{n}.$$

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

$$\mathop{\mathbb{E}}_{x_1,\dots,x_d}\left[f_{\approx i}(P_1)f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\right] = \mathop{\mathbb{E}}_{x_2,\dots,x_d}\left[f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\mathbb{E}_{x_1}\left[f_{\approx i}(P_1) \mid x_2, ..., x_d\right]\right]. \tag{3.2}$$

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

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

The latter expectation is upper bounded by $\frac{2^{3i^2+Ci}}{n}$ from Fact 2.13. Combining this with Fact 2.14 we conclude that

$$|\mathbb{E}_{x_1}\left[f_{\approx i}(P_1) \mid x_2, \ldots, x_d\right]| \leqslant \frac{2^{3i^2+Ci}d^2}{n}.$$

Using this, the triangle inequality and Fact 2.14 on (3.2) 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 those x_j that appear in all sets, $h_4 = |H_4|$, H_3 the set of those x_j that appear in exactly three of the sets, $h_3 = |H_3|$, H_2 the set of those x_j that appear in precisely two sets, $h_2 = |H_2|$.

Claim 3.9.

$$\left|\mathop{\mathbb{E}}_{H_4,H_3,H_2}\left[f_{\approx i}(P_1)f_{\approx i}(P_2)f_{\approx i}(P_3)f_{\approx i}(P_4)\right]\right| \leqslant \mathop{\mathbb{E}}_{H_4,H_3}\left[\prod_{j=1}^{4}\sqrt{\mathop{\mathbb{E}}_{H_2}\left[f_{\approx i}^2(P_j)\right]}\right].$$

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

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

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

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

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. For example, continuing another step, say we have $y \in H_2$ that appears in P_{j_3}, P_{j_2}, we write the expectation as

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

□

We next give an upper bound on the quantity

$$\mathop{\mathbb{E}}_{H_4,H_3}\left[\prod_{j=1}^{4}\sqrt{\mathop{\mathbb{E}}_{H_2}\left[f_{\approx i}^2(P_j)\right]}\right]. \tag{3.3}$$

Using the Cauchy–Schwarz inequality,

$$
\begin{aligned}
(3.3) &\leqslant \sqrt{\mathop{\mathbb{E}}_{H_4,H_3}\left[\mathop{\mathbb{E}}_{H_2}\left[f_{\approx i}^2(P_1)\right]\mathop{\mathbb{E}}_{H_2}\left[f_{\approx i}^2(P_2)\right]\right]}\sqrt{\mathop{\mathbb{E}}_{H_4,H_3}\left[\mathop{\mathbb{E}}_{H_2}\left[f_{\approx i}^2(P_3)\right]\mathop{\mathbb{E}}_{H_2}\left[f_{\approx i}^2(P_4)\right]\right]} \\
&\leqslant \max_{H_3\cap P_2,H_4}\sqrt{\mathop{\mathbb{E}}_{H_3\setminus P_2,H_2}\left[f_{\approx i}^2(P_1)\right]}\max_{H_3\cap P_3,H_4}\sqrt{\mathop{\mathbb{E}}_{H_3\setminus P_3,H_2}\left[f_{\approx i}^2(P_4)\right]} \\
&\quad \sqrt{\mathop{\mathbb{E}}_{H_4,H_3,H_2}\left[f_{\approx i}^2(P_2)\right]}\sqrt{\mathop{\mathbb{E}}_{H_4,H_3,H_2}\left[f_{\approx i}^2(P_3)\right]}.
\end{aligned}
\tag{3.4}
$$

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

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

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 at most

$$
\max_A \sqrt{\mathop{\mathbb{E}}_{B}\left[f_{\approx i}^2(A\cup B)\right]} \leqslant \sqrt{\frac{2^{2i+1}i^{|B|}\varepsilon}{\ell^{|B|}} + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}} \leqslant \frac{2^{i+1}i^{\lambda_1/2}\sqrt{\varepsilon}}{\ell^{\lambda_1/2}} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}},
$$

where $\lambda_1 = |H_2\cap P_1| + |H_{3,1}\setminus P_2|$ (we remark that this is the place in the proof where the pseudo-randomness of F is used). Similarly, the second maximum is upper bounded by

$$
\frac{2^{i+1}i^{\lambda_2/2}\sqrt{\varepsilon}}{\ell^{\lambda_2/2}} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}},
$$

where $\lambda_2 = |H_2\cap P_4| + |H_{3,4}\setminus P_3|$. Combining everything, we see that

$$
(3.4) \leqslant \frac{2^{2i+2}i^{\frac{1}{2}(2i+\lambda_1+\lambda_2)}\varepsilon\eta}{\ell^{\frac{1}{2}(2i+\lambda_1+\lambda_2)}} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}}. \tag{3.5}
$$

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 (3.5), we see that twice 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,

$$
(3.4) \leqslant \frac{2^{2i+2}i^d\varepsilon\eta}{\ell^d} + \frac{2^{1.5i^2+iC}\ell^{0.5i}}{n^{1/4}}. \qquad \square
$$

4 Pseudorandomness implies expansion

Proof of Theorem 2.19. Let S be a set as in the Theorem, and let F be its indicator function in $J(n, \ell, \alpha\ell)$. Note that

$$\delta(1 - \Phi(S)) = \langle F, J(n, \ell, \alpha\ell)F \rangle. \tag{4.1}$$

Claim 4.1.

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

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

$$\langle F, J(n, \ell, \alpha\ell)F \rangle = \langle F, \sum_{i=0}^{\ell} J(n, \ell, \alpha\ell)F_{=i} \rangle = \langle F, \sum_{i=0}^{\ell} \lambda_i(n, \ell, \alpha\ell)F_{=i} \rangle = \sum_{i=0}^{\ell} \lambda_i(n, \ell, \alpha\ell)\langle F_{=i}, F_{=i} \rangle.$$

□

To prove Theorem 2.19, we bound the right hand side of Claim 4.1; contribution of small values i is bounded by appealing to the pseudorandomness properties of F, and contribution of large values of i is bounded by using the fact the corresponding eigenvalues are small. Below are the details.

By Lemma 2.10 we have that

$$\lambda_i(n, \ell, \alpha\ell) \leqslant \frac{\binom{t}{i}}{\binom{\ell}{i}} + \frac{\ell}{n - \ell} \leqslant \left(\frac{t}{\ell}\right)^i + \frac{\ell}{n - \ell} = \alpha^i + \frac{\ell}{n - \ell},$$

so we conclude that

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

Let $i \leqslant r$. Observe that since S is (r, ε)-pseudorandom it follows that S is (i, ε)-pseudorandom. Therefore, applying Theorem 2.18 we see that for $i = 0, 1, \ldots, r$,

$$W^{=i}[F] \leqslant \mathrm{e}^{O(i)} \delta \varepsilon^{1/4} + \frac{1}{n^{1/24}}.$$

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

$$\begin{aligned}\langle F, J(n, \ell, \alpha\ell)F \rangle &\leqslant \sum_{i=0}^{r} \alpha^i \mathrm{e}^{O(i)} \delta \varepsilon^{1/4} + \alpha^{r+1} \sum_{i=r+1}^{\ell} W_{=i}[F] + \frac{r}{n^{1/24}} \\ &\leqslant \mathrm{e}^{O(r)} \delta \varepsilon^{1/4} + \delta \alpha^{r+1} + \frac{r}{n^{1/24}}.\end{aligned}$$

Plugging this into (4.1) and simplifying yields $\Phi(S) \geqslant 1 - \alpha^{r+1} - \mathrm{e}^{O(r)} \varepsilon^{1/4} - \frac{1}{n^{1/24}\delta}$. □

5 Proof of Theorem 2.20

Let S be an (r, ε)-pseudorandom set, and F be its indicator function. The proof of Theorem 2.20 follows the same outline as the proof of Theorem 2.18, except that we use higher moments. Let $0 \leqslant i \leqslant r$.

Throughout this section, we denote by C an absolute constant (that may not be the same in different occurrences), and assume that $n \geqslant (2^{3r^2+rC}\ell^{2r})^{4m(2m-1)^2}\delta^{-4}$.

Claim 5.1. $\mathbb{E}_A\left[F_{\approx i}^{2m}[A]\right] \geqslant \frac{\eta^{2m}}{2^{2m}\delta^{2m-1}}$.

Proof. As in Claim 3.5, we have that $\langle F, F_{\approx i}\rangle \geqslant \eta/2$. On the other hand, by Hölder's inequality we have

$$\langle F, F_{\approx i}\rangle \leqslant \|F\|_{2m/(2m-1)}\|F_{\approx i}\|_{2m} = \delta^{(2m-1)/2m}\|F_{\approx i}\|_{2m}.$$

Rearranging yields the result. □

The intersection pattern of $P_1, \ldots, P_{2m}$ is a vector σ indicating the sizes of 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 in the symbols $x_1, \ldots, x_d$. Then*

$$\left|\mathop{\mathbb{E}}_{x_1,\ldots,x_d}\left[\prod_{j=1}^{2m} f_{\approx i}(P_j)\right]\right| \leqslant 2^{4mi+2m}\frac{i^d\varepsilon^{2m-1}\eta}{\ell^d} + \frac{2^{8m}2^{(3i^2+iC)/2m(2m-1)}\ell^{i/2m(2m-1)}}{n^{1/4m(2m-1)}}. \tag{5.1}$$

The next section is devoted to the proof of this lemma.

5.1 Proof of Lemma 5.2

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_j then the lemma holds trivially:

Proposition 5.3. *There is an absolute constant $C > 0$ such that if there is an x_j that appears in only one of the P_i then*

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

Proof. The proof is the same as the proof of Proposition 3.8. □

We thus assume from now on that each x_j appears in at least 2 of the P_i. For $s = 2, 3 \ldots, 2m$, let H_s be the set of those x_j that appear in exactly s of the P_i. 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 $e_1(P) \stackrel{def}{=} |f_{\approx i}(P)|$ and for $s \geqslant 2$ define

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

(we note that $e_s(P)$ is only a function of $P \cap H_{>s}$).

For each $a = 2, 3, \ldots, 2m$, define an operator ρ_a on random variables by $\rho_a(Z) = \mathbb{E}_{H_a}\left[Z^{a/(a-1)}\right]$. For $j = 1, 2, \ldots, 2m-1$, define $T_{j,j}(P) \stackrel{def}{=} e_j(P)$. Inductively for $a > j$, define

$$T_{j,a}(P) \stackrel{def}{=} \rho_a(T_{j,a-1}(P)) = \mathop{\mathbb{E}}_{H_a}\left[T_{j,a-1}(P)^{\frac{a}{a-1}}\right].$$

Denote $Q_s[H_{\geqslant 2}] \stackrel{def}{=} \prod_{j=1}^{2m} \left|f_{\approx i}(P_j)\right|$, and for each $s \geqslant 2$ let

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

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

Proposition 5.4. *For every $1 \leqslant s \leqslant 2m$, and any setting of the variables in $H_{\geqslant s+1}$, we have*

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

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

Fact 5.5. *Let $h_1, \ldots, h_q \colon \binom{[n]}{m} \to \mathbb{R}$. Then*

$$\|h_1 \ldots h_q\|_1 \leqslant \|h_1\|_q \cdots \|h_q\|_q.$$

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

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

Iteratively, for each $y \in H_{s+1}$ we consider the $s+1$ sets P 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

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

Thus, we have that

$$\mathsf{LHS}(5.1) \leqslant \left(\prod_{j=1}^{2m} T_{1,2m}(P_j)\right)^{1/2m}. \tag{5.2}$$

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

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

Proof. Fix $j \leqslant 2m-1$; then

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

Clearly, for all settings of $H_{\geqslant j+1}$ we have

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

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

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

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 (5.2) yields

$$\mathsf{LHS}(5.1) \leqslant \prod_{j=1}^{2m}\prod_{a=1}^{2m} \max_{H_{\geqslant a+1}} e_a(P_j)^{g(a)}. \tag{5.3}$$

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

$$\max_{H_{\geqslant a+1}} e_a(P) \leqslant 2^{2i+1}\frac{i^{|P\cap H_{\leqslant a}|}\varepsilon}{\ell^{|P\cap H_{\leqslant a}|}} + + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}.$$

For $a = 2m$, by Claim 3.2

$$\max_{H_{\geqslant a+1}} e_a(P) \leqslant \frac{i^{|P\cap H_{\leqslant a}|}\eta}{\ell^{|P\cap H_{\leqslant a}|}} + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}.$$

Therefore, plugging into (5.3) we get

$$\begin{aligned}
\mathsf{LHS}(5.1) &\leqslant \prod_{j=1}^{2m}\left(\frac{i^{\left|P_j\cap H_{\leqslant 2m}\right|}\eta}{\ell^{\left|P_j\cap H_{\leqslant 2m}\right|}} + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}\right)^{g(2m)} \cdot \prod_{j=1}^{2m}\prod_{a=1}^{2m-1}\left(2^{2i+1}\frac{i^{\left|P_j\cap H_{\leqslant a}\right|}\varepsilon}{\ell^{\left|P_j\cap H_{\leqslant a}\right|}} + \frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}\right)^{g(a)} \\
&\leqslant \prod_{j=1}^{2m}\left(\frac{i^{\left|P_j\cap H_{\leqslant 2m}\right|}\eta}{\ell^{\left|P_j\cap H_{\leqslant 2m}\right|}}\right)^{g(2m)} \cdot \prod_{j=1}^{2m}\prod_{a=1}^{2m}\left(2^{2i+1}\frac{i^{\left|P_j\cap H_{\leqslant a}\right|}\varepsilon}{\ell^{\left|P_j\cap H_{\leqslant a}\right|}}\right)^{g(a)} + 2^{8m}\left(\frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}\right)^{\min(g(2m-1),g(2m))} \\
&\leqslant 2^{4mi+2m}\eta\varepsilon^{2m-1}\prod_{j=1}^{2m}\prod_{a=1}^{2m}\frac{i^{g(a)\cdot\left|P_j\cap H_{\leqslant a}\right|}}{\ell^{g(a)\cdot\left|P_j\cap H_{\leqslant a}\right|}} + 2^{8m}\left(\frac{2^{3i^2+iC}\ell^i}{\sqrt{n}}\right)^{\min(g(2m-1),g(2m))}.
\end{aligned} \tag{5.4}$$

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

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

In the second equality, we used the fact that each element in H_r is counted r times by the P_j. 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 (5.4) yields

$$\mathsf{LHS}(5.1) \leqslant 2^{4mi+2m} \frac{i^d \eta \varepsilon^{2m-1}}{\ell^d} + \frac{2^{8m} 2^{(3i^2+iC)/2m(2m-1)} \ell^{i/2m(2m-1)}}{n^{1/4m(2m-1)}},$$

concluding the proof of Lemma 5.2.

5.2 Concluding Theorem 2.20

In this section, we finish the proof of Theorem 2.20. The following corollary is proven using a similar argument to the one in Lemma 3.6. Since the proof is almost identical, we give a rough outline of it.

Corollary 5.7. *If F is (i, ε)-pseudorandom, then*

$$\mathop{\mathbb{E}}_{A} \left[F_{\approx i}^{2m}[A] \right] \leqslant (10m)^{4mi} \eta \varepsilon^{2m-2} + \frac{2^{8m} 2^{3i^2 m + imC} \ell^{(2m+1)i}}{n^{1/(4m(2m-1))}}.$$

Proof. As in (3.1), we expand out

$$\mathop{\mathbb{E}}_{A} \left[F_{\approx i}^{2m}[A] \right] = \sum_{d=i}^{2mi} \sum_{\vec{\sigma}_d} \beta(\vec{\sigma}_d) \beta_{i,d,\ell,m} \mathop{\mathbb{E}}_{A} \left[\mathop{\mathbb{E}}_{\substack{(I_1,\ldots,I_{2m}) \sim \gamma(\vec{\sigma}_d) \\ I_1,\ldots,I_{2m} \subseteq A}} \left[f_{\approx i}(I_1) \cdots f_{\approx i}(I_{2m}) \right] \right], \tag{5.5}$$

where $\vec{\sigma}_d$ now ranges over all intersection patterns of $2m$ sets of size i, $\beta(\vec{\gamma}_d)$ is the distribution over the intersection patterns and $\beta_{i,d,\ell,m}$ counts the number of $I_1, \ldots, I_{2m} \subseteq A$ of size i whose union has size d. We then appeal to Lemma 5.2 (instead of Lemma 3.7) to bound the absolute value of the expectation inside, and use the crude bound $\beta_{i,d,\ell,m} \leqslant \binom{\ell}{d} \binom{d}{i}^{2m} \leqslant \ell^d d^{2mi}$ to finish the proof. □

Combining Claim 5.1 and Corollary 5.7, we get that

$$\frac{\eta^{2m}}{2^{2m} \delta^{2m-1}} \leqslant (10m)^{4mi} \eta \varepsilon^{2m-2} + \frac{2^{8m} 2^{3i^2 m + imC} \ell^{(2m+1)i}}{n^{1/(4m(2m-1))}}.$$

Rearranging we get that

$$\eta^{2m-1} \leqslant 2^{2m}(10m)^{4mi}\delta^{2m-1}\varepsilon^{2m-2} + \frac{2^{10m}2^{3i^2m+imC}\ell^{(2m+1)i}}{n^{1/(4m(2m-1))}}\frac{\delta^{2m-1}}{\eta},$$

taking $(2m-1)$-root yields

$$\eta \leqslant 16(10m)^{2i}\delta\varepsilon^{1-\frac{1}{2m-1}} + \frac{2^{3i^2+iC}\ell^{2i}}{n^{1/(4m(2m-1)^2)}}\frac{\delta}{\eta^{1/(2m-1)}} \leqslant 16(10m)^{2i}\delta\varepsilon^{1-\frac{1}{2m-1}} + \frac{2^{3i^2+iC}\ell^{2i}}{n^{1/(4m(2m-1)^2)}}.$$

finishing the proof.

A Missing proofs

A.1 Proof of Fact 2.13

Let $a_i = \max_{|R|<i} |\mu_R(f_{\approx i})|$. For a_1, we note that $\mu(f_{\approx 1}) = 0$, so $a_1 = 0$. Next, we bound a_i in terms of $a_1, \ldots, a_{i-1}$, and thereby prove Fact 2.13 by induction.

Let J be such that $a_i = \left|\mu_J(f_{\approx i})\right|$, and note that we may take such J of size $i-1$. Then

$$\begin{aligned}\mu_J(f_{\approx i}) = \mathop{\mathbb{E}}_{I\supseteq J}[f_{\approx i}(I)] &= \mathop{\mathbb{E}}_{I\supseteq J}\left[\mu_I(F) - \sum_{i'=0}^{i-1}\sum_{|I'|=i', I'\subseteq I} f_{\approx i'}(I')\right] \\ &= \mu_J(F) - \sum_{i'=0}^{i-1}\sum_{|I'|=i', I'\subseteq J} f_{\approx i'}(I') - \mathop{\mathbb{E}}_{I\supseteq J}\left[\sum_{i'=0}^{i-1}\sum_{\substack{|I'|=i'\\ I'\subseteq I\\ I'\nsubseteq J}} f_{\approx i'}(I')\right] \\ &= -\mathop{\mathbb{E}}_{I\supseteq J}\left[\sum_{i'=0}^{i-1}\sum_{\substack{|I'|=i'\\ I'\subseteq I\\ I'\nsubseteq J}} f_{\approx i'}(I')\right],\end{aligned}$$

where the last equality is by definition of $f_{\approx j}(J)$. We may write

$$\sum_{|I'|=i', I'\subseteq I, I'\nsubseteq J} f_{\approx i'}(I') = \sum_{\substack{R\subsetneq J\\ |R|<i'}}\sum_{\substack{|I'|=i'\\ I'\subseteq I\\ I'\cap J=R}} f_{\approx i'}(I'),$$

and plugging that in above yields

$$\mu_J(f_{\approx i}(I)) = -\sum_{i'=0}^{i-1}\sum_{R\subsetneq J, |R|<i'} a_{i,i',j,R}\mathop{\mathbb{E}}_{I\supseteq J}\left[\mathop{\mathbb{E}}_{\substack{I'\subseteq I\\ I'\cap J=R}}[f_{\approx i'}(I')]\right],$$

where $a_{i,i',j,R}$ counts the number of $I' \subseteq I$ such that $I' \cap J = R$; we will only use the fact that $a_{i,i',j,R} \leqslant 2^i$, and get that

$$\mu_J(f_{\approx i}) \leqslant i2^i \max_{i',R,|R'|<i'} \left| \mathop{\mathbb{E}}_{I \supseteq J, I' \subseteq I : I' \cap J = R} [f_{\approx i'}(I')] \right|.$$

Fix i' and R that maximize this. Note that the distribution over I' is uniform among these that satisfy $I' \cap J = R$, so the expectation above is equal to

$$\begin{aligned}\frac{1}{\Pr_{I'}[I' \cap J = R]} \mathop{\mathbb{E}}_{I'}\left[f_{\approx i'}(I') 1_{I' \cap J = R}\right] &= \frac{1}{\Pr_{I'}[I' \cap J = R]} \sum_{R \subseteq R' \subseteq J} (-1)^{|R'' \setminus R'|} \mathop{\mathbb{E}}_{I'}[f_{\approx i'}(I') 1_{I' \supseteq R'}] \\ &= \sum_{R \subseteq R' \subseteq J} \frac{\Pr_{I'}[I' \supseteq R']}{\Pr_{I'}[I' \cap J = R]} (-1)^{|R'' \setminus R'|} \mu_{R'}(f_{\approx i'}),\end{aligned}$$

where in the second transition we used Claim 2.7. Combining everything, we get that

$$a_i = \left|\mu_J(f_{\approx i})\right| \leqslant \sum_{i'=0}^{i-1} \sum_{R \subseteq R' \subseteq J} \frac{\Pr_{I'}[I' \supseteq R']}{\Pr_{I'}[I' \cap J = R]} \left|\mu_{R'}(f_{\approx i'})\right|.$$

For R' such that $|R'| \leqslant i'-1$, we get that $|\mu_{R'}(f_{\approx i'})| \leqslant a_{i'}$. Otherwise, $|R'| \geqslant i'$, and thus $|R| < |R'|$, and we have $|\mu_{R'}(f_{\approx i'})| \leqslant 2^{i^2} \|F\|_\infty$ by Fact 2.14. Lastly, we have

$$\frac{\Pr_{I'}[I' \supseteq R']}{\Pr_{I'}[I' \cap J = R]} \leqslant \frac{\frac{\ell(\ell-1)\cdots(\ell-|R'|)}{n(n-1)\cdots(n-|R'|)}}{\frac{\ell(\ell-1)\cdots(\ell-|R|)}{n^{|R|}} \cdot \Omega(1)} \leqslant O\left(\frac{\ell^{|R'|-|R|}}{n^{|R'|-|R|}}\right),$$

which is at most $O(1)$ for all $|R'| \geqslant |R|$ and at most $O\left(\frac{\ell}{n}\right)$ if $|R'| > |R|$. Together, we conclude that

$$a_i \leqslant 2^{i+C} a_{i-1} + 2^{2i^2+C} \frac{\ell}{n} \|F\|_\infty$$

for some absolute constant $C > 0$. Thus, looking at the sequence b_i where $b_0 = 0$ and $b_i = 2^{i+C} b_{i-1} + 2^{2i^2+C} \frac{\ell}{n} \|F\|_\infty$, we have that $a_i \leqslant b_i$, and solving the recurrence gives that

$$b_i \leqslant \left(2^{2i^2+C} + 2^{2i^2+C+i+C} + 2^{2i^2+C+i+(i-1)+2C} + \ldots + 2^{2i^2+C+i+(i-1)+\ldots+1+iC}\right) \frac{\ell}{n} \|F\|_\infty,$$

giving the bound $b_i \leqslant 2^{3i^2+iC'} \frac{\ell}{n} \|F\|_\infty$, for some absolute constant $C' > 0$. □

Corollary A.1. *Let $1 \leqslant i \leqslant \ell$, and $F \colon \binom{[n]}{i} \to \mathbb{R}$ and assume that $n \geqslant 2\ell^2$. Then for all R of size at most $i-1$ we have*

$$|\mu_R(F_{\approx i})| \leqslant 2^{3i^2+Ci} \frac{\ell^{i+1}}{n} \|F\|_\infty$$

for some absolute constant $C > 0$.

Proof. By definition,

$$\mu_R(F_{\approx i}) = \mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{I\subseteq A} f_{\approx i}(I)\right] = \sum_{j=0}^{i-1}\sum_{J\subseteq R,|J|=j}\mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{\substack{I:I\cap R=J\\ I\subseteq A}} f_{\approx i}(I)\right] = \sum_{j=0}^{i-1}\sum_{J\subseteq R,|J|=j} a_{i,j,\ell}\mathop{\mathbb{E}}_{I\cap R=J}[f_{\approx i}(I)].$$

In the last transition, we turned the sum into expectation by dividing and multiplying by $a_{i,j,\ell}$, that counts the number of $I\subset A$ such that $I\cap R=J$; we will only use the trivial bound $a_{i,j,\ell}\leqslant \ell^i$, hence

$$|\mu_R(F_{\approx i})| \leqslant i2^i\ell^i \max_{j\leqslant i-1,|J|=j}\left|\mathop{\mathbb{E}}_{I\cap R=J}[f_{\approx i}(I)]\right|.$$

To bound this, we fix j and J that achieve this maximum, and write

$$\mathop{\mathbb{E}}_{I\cap R=J}[f_{\approx i}(I)] = \frac{1}{\Pr_I[I\cap R=J]}\mathop{\mathbb{E}}_I\left[f_{\approx i}(I)1_{I\cap R=J}\right] = \frac{1}{\Pr_I[I\cap R=J]}\sum_{J\subseteq J'\subseteq R}(-1)^{|J'\setminus J|}\mathop{\mathbb{E}}_I\left[f_{\approx i}(I)1_{I\supseteq J'}\right],$$

where we used Claim 2.7. Thus, we get

$$\left|\mathop{\mathbb{E}}_{I\cap R=J}[f_{\approx i}(I)]\right| \leqslant \sum_{J\subseteq J'\subseteq R}\frac{\Pr_I[I\supseteq J']}{\Pr_I[I\cap R=J]}\mu_{J'}(f_{\approx i}).$$

As before, we have that

$$\frac{\Pr_I[I\supseteq J']}{\Pr_I[I\cap R=J]} \leqslant O\left(\frac{\ell^{|J'|-|J|}}{n^{|J'|-|J|}}\right),$$

hence this is at most $O(1)$ for all J', J (as $|J'|\geqslant |J|$ always), and at most $O(\ell/n)$ if $|J'|>|J|$. Hence, we get that

$$\sum_{J\subseteq J'\subseteq R}\frac{\Pr_I[I\supseteq J']}{\Pr_I[I\cap R=J]}\mu_{J'}(f_{\approx i}) \leqslant 2^i O\left(\frac{\ell}{n}\right)\|f_{\approx i}\|_\infty + 2^i O(1)\max_{|J'|=j}\left|\mu_{J'}(f_{\approx i})\right|.$$

Using Facts 2.14 and 2.13, this is at most $2^{3i^2+Ci}\frac{\ell}{n}\|F\|_\infty$ for some absolute constant $C>0$, and we are done. □

Claim A.2. *Assume $n\geqslant 2\ell^2$ and let $1\leqslant i\leqslant \ell$, $H\in J_{\leqslant i-1}$ be given as $H[A]=\sum_{I'\subseteq A} h(I')$, and $G\colon \binom{[n]}{\ell}\to\mathbb{R}$ be some function. If $\mathbb{E}_{I'}\left[h(I')^2\right]\leqslant M$, and $\mathop{\mathbb{E}}_{\substack{I'\subseteq A\\ |I'|=i-1}}\left[\mu_{I'}(G)^2\right]\leqslant \eta$, then*

$$|\langle F_{\approx i}, H\rangle| \leqslant \ell^{i-1}\sqrt{\eta M}.$$

Proof. By definition,

$$\langle F_{\approx i}, H\rangle = \mathop{\mathbb{E}}_A[F_{\approx i}[A]H[A]] = \binom{\ell}{i-1}\mathop{\mathbb{E}}_A\left[\mathop{\mathbb{E}}_{\substack{I'\subseteq A\\ |I'|=i-1}}[F_{\approx i}[A]h(I')]\right] = \binom{\ell}{i-1}\mathop{\mathbb{E}}_{\substack{I'\subseteq A\\ |I'|=i-1}}[\mu_{I'}(F_{\approx i})h(I')].$$

Using Cauchy–Schwarz, we may upper bound the absolute value of this by

$$\binom{\ell}{i-1}\sqrt{\mathop{\mathbb{E}}_{\substack{I'\subseteq A\\|I'|=i-1}}[\mu_{I'}(F_{\approx i})^2]}\sqrt{\mathop{\mathbb{E}}_{\substack{I'\subseteq A\\|I'|=i-1}}[h(I')^2]} \leqslant \ell^{i-1}\sqrt{\eta M}. \qquad \square$$

A.2 Proof of Theorem 2.12

Normalizing F, we shall assume that $\|F\|_\infty = 1$.

Define $G_i = F - \sum_{j=0}^{i} F_{\approx j}$. We show that G_i is nearly perpendicular to $J_{\leqslant i}$. Intuitively, this will allow us to conclude that $\sum_{j=0}^{i} F_{\approx j}$ is close to $\sum_{j=0}^{i} F_{=j}$ for all i, from which Theorem 2.12 follows via an easy induction.

Claim A.3. *For all $R \subseteq [n]$ of size at most i, we have that $|\mu_R(G_i)| \leqslant \frac{2^{3i^2+iC}\cdot \ell^{i+1}}{n}$ for some absolute constant $C > 0$.*

Proof. It suffices to prove the statement for R of size i. By definition,

$$\mu_R(G_i) = \mu_R(F) - \sum_{j=0}^{i}\mu_R(F_{\approx j}).$$

Expanding, we have

$$\mu_R(F_{\approx i}) = \mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{J\subseteq A} f_{\approx j}(J)\right] = \sum_{J\subseteq R} f_{\approx j}(J) + \mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{J\subseteq A, J\nsubseteq R} f_{\approx j}(J)\right],$$

so we get that

$$\begin{aligned}\mu_R(G_i) &= \mu_R(F) - \sum_{j=0}^{i}\sum_{J\subseteq R} f_{\approx j}(J) - \sum_{j=0}^{i}\mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{J\subseteq A, J\nsubseteq R} f_{\approx j}(J)\right]\\ &= \mu_R(F) - f_{\approx i}(R) - \sum_{j=0}^{i-1}\sum_{J\subseteq R} f_{\approx j}(J) - \sum_{j=0}^{i}\mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{J\subseteq A, J\nsubseteq R} f_{\approx j}(J)\right]\\ &= -\sum_{j=0}^{i}\mathop{\mathbb{E}}_{A\supseteq R}\left[\sum_{J\subseteq A, J\nsubseteq R} f_{\approx j}(J)\right],\end{aligned}$$

where we used the definition of $f_{\approx i}$; it remains to bound the last sum. We partition the sum according to $J \cap R$:

$$\sum_{j=0}^{i} \mathop{\mathbb{E}}_{A \supseteq R}\left[\sum_{\substack{R' \subsetneq R \\ |R'| \leqslant j-1}} \sum_{J \subseteq A, J \cap R = R'} f_{\approx j}(J)\right] = \sum_{j=0}^{i} \sum_{\substack{R' \subsetneq R \\ |R'| \leqslant j-1}} \mathop{\mathbb{E}}_{A \supseteq R}\left[a_{|R'|,j,i,\ell} \mathop{\mathbb{E}}_{J \subseteq A, J \cap R = R'}\left[f_{\approx j}(J)\right]\right],$$

where $a_{|R'|,j,i,\ell}$ is the number of sets J of size j such that $J \subseteq A$ and $J \cap R = R'$; we will only use the fact that it is at most ℓ^j, so that in absolute value the above sum is at most

$$2^i \ell^j \max_{\substack{j, R' \subsetneq R \\ |R'| \leqslant j-1}} \left|\mathop{\mathbb{E}}_{J \cap R = R'}\left[f_{\approx j}(J)\right]\right|.$$

Fix j and R'. By Claim 2.7 we have

$$\mathop{\mathbb{E}}_{J}\left[f_{\approx j}(J) 1_{J \cap R = R'}\right] = \sum_{R'' : R' \subseteq R'' \subseteq R} (-1)^{|R'' \setminus R'|} 1_{J \supseteq R''},$$

so

$$\mathop{\mathbb{E}}_{J \cap R = R'}\left[f_{\approx j}(J)\right] = \sum_{R'' : R' \subseteq R'' \subseteq R} (-1)^{|R'' \setminus R'|} \frac{\Pr_J\left[J \supseteq R''\right]}{\Pr_J\left[J \cap R = R'\right]} \mu_{R''}(f_{\approx j}).$$

Taking absolute value and using the triangle inequality, we get that

$$\left|\mathop{\mathbb{E}}_{J \cap R = R'}\left[f_{\approx j}(J)\right]\right| \leqslant \sum_{R'' : R' \subseteq R'' \subseteq R} \frac{\Pr_J\left[J \supseteq R''\right]}{\Pr_J\left[J \cap R = R'\right]} \left|\mu_{R''}(f_{\approx j})\right|.$$

Note that

$$\frac{\Pr_J\left[J \supseteq R''\right]}{\Pr_J\left[J \cap R = R'\right]} \leqslant \frac{\frac{\ell(\ell-1)\cdots(\ell-|R''|)}{n(n-1)\cdots(n-|R''|)}}{\frac{\ell(\ell-1)\cdots(\ell-|R'|)}{n^{|R'|}} \cdot \Omega(1)} \leqslant O\left(\frac{\ell^{|R''|-|R'|}}{n^{|R''|-|R'|}}\right).$$

This is at most $O(1)$ for all $R' \subseteq R''$, and as $|R'| \leqslant j-1$, this is $O(\ell/n)$ if $|R''| \geqslant j$. We thus get from Fact 2.13 that

$$\left|\mathop{\mathbb{E}}_{J \cap R = R'}\left[f_{\approx j}(J)\right]\right| \leqslant 2^j O\left(\frac{\ell}{n}\right) + 2^j \cdot O(1) \cdot 2^{3j^2+Cj} \frac{\ell}{n} \leqslant 2^{3j^2+C'j} \frac{\ell}{n}.$$

$\square$

Claim A.4. *Let $j \leqslant i$. Then $\left|\langle G_i, F_{=j}\rangle\right| \leqslant \frac{2^{3i^2+iC} \ell^{1.5i+1}}{n}$.*

Proof. Note that

$$\left|\langle G_i, F_{=j}\rangle\right| = \left|\mathop{\mathbb{E}}_{A}\left[\sum_{J \subseteq A} f_{=j}(J) G_i[A]\right]\right| = \binom{\ell}{j}\left|\mathop{\mathbb{E}}_{I}\left[f_{=j}(J)\mu_I(G_i)\right]\right| \leqslant \binom{\ell}{j}\sqrt{\mathop{\mathbb{E}}_{J}\left[f_{=j}(J)^2\right]}\sqrt{\mathop{\mathbb{E}}_{J}\left[\mu_J(G_i)^2\right]}.$$

Using Claims A.3 and A.7 we get that $|\langle G_i, F_{=i}\rangle| \leqslant \frac{2^{3i^2+iC} \ell^{i+0.5j+1}}{n}$. $\square$

Note that
$$\langle G_i, F_{=i}\rangle = \langle F, F_{=i}\rangle - \langle F_{\approx i}, F_{=i}\rangle = \|F_{=i}\|_2^2 - \langle F_{\approx i}, F_{=i}\rangle,$$
so we get that
$$\left|\|F_{=i}\|_2^2 - \langle F_{\approx i}, F_{=i}\rangle\right| \leqslant \frac{2^{3i^2+iC}\ell^{1.5i+1}}{n}. \tag{A.1}$$

Claim A.5. *Let $j \leqslant i$. If $n \geqslant 2^{3i^2+iC}\ell^{2i+1}$, then $\left|\langle G_i, F_{\approx j}\rangle\right| \leqslant \frac{2^{3i^2+iC}\ell^{1.5i+1}}{n}$.*

Proof. Note that
$$\langle G_i, F_{\approx j}\rangle = \left|\mathop{\mathbb{E}}_{A}\left[\sum_{J\subseteq A} f_{\approx j}(J) G_i[A]\right]\right| = \binom{\ell}{j}\left|\mathop{\mathbb{E}}_{I}\left[f_{\approx j}(J)\mu_I(G_i)\right]\right| \leqslant \binom{\ell}{j}\sqrt{\mathop{\mathbb{E}}_{J}\left[f_{\approx j}(J)^2\right]}\sqrt{\mathop{\mathbb{E}}_{J}\left[\mu_J(G_i)^2\right]}.$$

Using Fact 2.14 and Claim A.7 we get that $|\langle G_i, F_{=i}\rangle| \leqslant \frac{2^{4i^2+iC}\ell^{i+j+1}}{n}$. □

Claim A.6. *There is an absolute constant $C > 0$ such that if If $n \geqslant 2^{3i^2+iC}\ell^{2i+1}$, then*
$$\left|\langle F_{=i}, F_{\approx i}\rangle - \|F_{\approx i}\|_2^2\right| \leqslant \frac{2^{4i^2+iC}\ell^{2i}}{n}.$$

Proof. Note that
$$\langle G_i, F_{\approx i}\rangle = \langle F, F_{\approx i}\rangle - \langle F_{\approx i}, F_{\approx i}\rangle - \sum_{j<i}\langle F_{\approx j}, F_{\approx i}\rangle = \langle F_{=i}, F_{\approx i}\rangle - \|F_{\approx i}\|_2^2 + \sum_{j<i}\langle F_{=j}, F_{\approx i}\rangle - \langle F_{\approx j}, F_{\approx i}\rangle.$$

so
$$\left|\langle F_{=i}, F_{\approx i}\rangle - \|F_{\approx i}\|_2^2\right| \leqslant \underbrace{\left|\langle G_i, F_{\approx i}\rangle\right|}_{(I)} + \sum_{j<i}\underbrace{\left|\langle F_{\approx j}, F_{\approx i}\rangle\right|}_{(II)} + \underbrace{\left|\langle F_{=j}, F_{\approx i}\rangle\right|}_{(III)} \leqslant \frac{2^{4i^2+iC}\ell^{2i}}{n},$$

for some absolute constant $C > 0$. To justify the last inequality, we note that (I) may be bounded using Claim A.2 by appealing to Claim A.3 to bound η and Fact 2.14 to bound M. Similarly, (II) may be bounded using Claim A.2 by appealing to Corollary A.1 and Fact 2.14 to bound M and η. Lastly, (III) may be bounded using Claim A.2 by appealing to Corollary A.1 and Claim A.7 to bound M and η. □

We are now ready to prove Theorem 2.12.

Proof of Theorem 2.12. We have
$$\|F_{=i} - F_{\approx i}\|_2^2 = \|F_{=i}\|_2^2 - \langle F_{=i}, F_{\approx i}\rangle + \|F_{\approx i}\|_2^2 - \langle F_{=i}, F_{\approx i}\rangle.$$

Using (A.1) and Claim A.6 gives the desired bound. □

A.3 Auxiliary claims

Claim A.7. *There is an absolute constant $C > 0$ such that for all $F\colon \binom{[n]}{\ell} \to \mathbb{R}$, if $n \geqslant 2^{i+C}\ell^{2i+1}$, then*

$$\mathop{\mathbb{E}}_{I}\left[f_{=i}(I)^2\right] \leqslant \frac{2\|F\|_2^2}{\sqrt{\binom{\ell}{i}}}.$$

Proof. By orthogonality of $F_{=i}$ we have

$$\|F\|_2^2 \geqslant \|F_{=i}\|_2^2 = \mathop{\mathbb{E}}_{A}\left[\left(\sum_{I\subseteq A} f_{=i}(I)\right)^2\right] = \mathop{\mathbb{E}}_{A}\left[\sum_{I\subseteq A} f_{=i}(I)^2 + \sum_{\substack{I,I'\subseteq A\\ I\neq I'}} f_{=i}(I)f_{=i}(I')\right].$$

Thus,

$$\|F\|_2^2 \geqslant \binom{\ell}{i}\mathop{\mathbb{E}}_{I}\left[f_{=i}(I)^2\right] + \sum_{r=0}^{i-1} a_{r,i,\ell}\mathop{\mathbb{E}}_{|R|=r}\left[\mathop{\mathbb{E}}_{\substack{I,I'\\ I\cap I'=R}}\left[f_{=i}(I)f_{=i}(I')\right]\right],$$

where $a_{r,i,\ell}$ counts the number of $I, I' \subseteq A$ that intersect in a set of size r; we will only use the obvious bound $a_{r,i,\ell} \leqslant \ell^{2i}$. Thus,

$$\|f_{=i}\|_2^2 \leqslant \frac{\|F\|_2^2}{\binom{\ell}{i}} + i\ell^{2i}\max_{r\leqslant i-1, |R|=r}\left|\mathop{\mathbb{E}}_{\substack{I,I'\\ I\cap I'=R}}\left[f_{=i}(I)f_{=i}(I')\right]\right|. \tag{A.2}$$

Fix r and R that maximize this. Then

$$\begin{aligned}\mathop{\mathbb{E}}_{\substack{I,I'\\ I\cap I'=R}}\left[f_{=i}(I)f_{=i}(I')\right] &= \frac{1}{\Pr_{I,I'}\left[I\cap I'=R\right]}\mathop{\mathbb{E}}_{I'}\left[\mathop{\mathbb{E}}_{I}\left[f_{=i}(I)f_{=i}(I')1_{I\cap I'=R}\right]\right]\\ &= \frac{1}{\Pr_{I,I'}\left[I\cap I'=R\right]}\mathop{\mathbb{E}}_{I'}\left[f_{=i}(I')\sum_{R\subseteq R'\subseteq I'}(-1)^{|R'\setminus R|}\mathop{\mathbb{E}}_{I}\left[f_{=i}(I)1_{I\supseteq R'}\right]\right],\end{aligned}$$

where we used Claim 2.7. Thus, in absolute value this is at most

$$\frac{1}{\Pr_{I,I'}\left[I\cap I'=R\right]}\mathop{\mathbb{E}}_{I'}\left[|f_{=i}(I')|\sum_{R\subseteq R'\subseteq I'}\Pr_{I}\left[I\supseteq R'\right]|\mu_{R'}(f_{=i})|\right],$$

and using Cauchy–Schwarz this is at most

$$\begin{aligned}&\frac{1}{\Pr_{I,I'}\left[I\cap I'=R\right]}\sqrt{\mathop{\mathbb{E}}_{I'}\left[f_{=i}(I')^2\right]}\sqrt{\mathop{\mathbb{E}}_{I'}\left[2^i\sum_{R\subseteq R'\subseteq I'}\Pr_{I}\left[I\supseteq R'\right]^2|\mu_{R'}(f_{=i})|^2\right]}\\ &= \sqrt{\mathop{\mathbb{E}}_{I'}\left[f_{=i}(I')^2\right]}\sqrt{\underbrace{\mathop{\mathbb{E}}_{I'}\left[2^i\sum_{R\subseteq R'\subseteq I'}\frac{\Pr_{I}\left[I\supseteq R'\right]^2}{\Pr_{I,I'}\left[I\cap I'=R\right]^2}|\mu_{R'}(f_{=i})|^2\right]}_{(I)}}.\end{aligned}$$

We inspect (I). As before, we have

$$\frac{\Pr_I[I \supseteq R']^2}{\Pr_{I,I'}[I \cap I' = R]^2} = O\left(\left(\frac{\ell}{n}\right)^{2(|R'|-|R|)}\right).$$

The contribution from a single R' such that $|R'| < i$ is 0 by Lemma 2.8. Otherwise, $R' = I'$ and as $|R| < i$ the contribution is at most

$$\frac{O(2^i)\ell^2}{n^2} \mathop{\mathbb{E}}_{I'}\left[f_{=i}(I')^2\right].$$

Overall, we get that

$$\sqrt{(I)} \leqslant O\left(\frac{2^i \ell}{n}\right)\|f_{=i}\|_2.$$

Plugging everything into (A.2) gives that

$$\|f_{=i}\|_2^2 \leqslant \frac{\|F\|_2^2}{\binom{\ell}{i}} + \frac{2^{i+C}\ell^{2i+1}}{n}\|f_{=i}\|_2^2,$$

and as $\frac{2^{i+C}\ell^{2i+1}}{n} \leqslant 1/2$, we get that

$$\|f_{=i}\|_2^2 \leqslant \frac{2\|F\|_2^2}{\binom{\ell}{i}}. \qquad \square$$

References

[1] MITALI BAFNA, BOAZ BARAK, PRAVESH K. KOTHARI, TSELIL SCHRAMM, AND DAVID STEURER: Playing unique games on certified small-set expanders. In *Proc. 53rd STOC*, pp. 1629–1642. ACM Press, 2021. [doi:10.1145/3406325.3451099] 8

[2] MITALI BAFNA, MAX HOPKINS, TALI KAUFMAN, AND SHACHAR LOVETT: Hypercontractivity on high dimensional expanders. In *Proc. 54th STOC*, pp. 185–194. ACM Press, 2022. [doi:10.1145/3519935.3520040] 6

[3] MITALI BAFNA AND DOR MINZER: Solving unique games over globally hypercontractive graphs. In *Proc. 39th Comput. Complexity Conf. (CCC'24)*, pp. 3:1–15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2024. [doi:10.4230/LIPIcs.CCC.2024.3, arXiv:2304.07284] 8

[4] 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. [doi:10.1137/130929394] 4

[5] WILLIAM BECKNER: Inequalities in Fourier analysis. *Ann. Math.*, 102(1):159–182, 1975. [doi:10.2307/1970980] 3

[6] ALINE BONAMI: Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier*, 20(2):335–402, 1970. Also accessible via Numdam. [doi:10.5802/aif.357] 3

[7] ROEE DAVID, IRIT DINUR, ELAZAR GOLDENBERG, GUY KINDLER, AND IGOR SHINKAR: Direct sum testing. *SIAM J. Comput.*, 46(4):1336–1369, 2017. [doi:10.1137/16M1061655] 8

[8] PHILIPPE DELSARTE: The association schemes of coding theory. In MARSHALL HALL JR. AND JACOBUS H. VAN LINT, editors, *Combinatorics*, pp. 143–161. Springer, 1975. [doi:10.1007/978-94-010-1826-5_7] 9

[9] PHILIPPE DELSARTE: Association schemes and t-designs in regular semilattices. *J. Combin. Theory–A*, 20(2):230–243, 1976. [doi:10.1016/0097-3165(76)90017-0] 9

[10] IRIT DINUR AND VENKATESAN GURUSWAMI: PCPs via the low-degree long code and hardness for constrained hypergraph coloring. *Israel J. Math.*, 209(2):611–649, 2015. [doi:10.1007/s11856-015-1231-3] 4

[11] IRIT DINUR, SUBHASH KHOT, GUY KINDLER, DOR MINZER, AND MULI SAFRA: Towards a proof of the 2-to-1 games conjecture? In *Proc. 50th STOC*, pp. 376–389. ACM Press, 2018. [doi:10.1145/3188745.3188804] 5

[12] IRIT DINUR, SUBHASH KHOT, GUY KINDLER, DOR MINZER, AND MULI SAFRA: On non-optimally expanding sets in Grassmann graphs. *Israel J. Math.*, 243(1):377–420, 2021. Preliminary version in STOC'18. [doi:10.1007/s11856-021-2164-7] 5, 7, 8, 9

[13] IRIT DINUR, ELCHANAN MOSSEL, AND ODED REGEV: Conditional hardness for approximate coloring. *SIAM J. Comput.*, 39(3):843–873, 2009. [doi:10.1137/07068062X] 3

[14] IRIT DINUR AND INBAL LIVNI NAVON: Exponentially small soundness for the direct product Z-test. *Theory of Computing*, 19(3):1–56, 2023. Preliminary version in CCC'17. [doi:10.4086/toc.2023.v019a003] 8

[15] IRIT DINUR AND SAMUEL SAFRA: On the hardness of approximating Minimum Vertex Cover. *Ann. Math.*, 162(1):439–485, 2005. [doi:10.4007/annals.2005.162.439] 3

[16] URIEL FEIGE AND GIDEON SCHECHTMAN: On the optimality of the random hyperplane rounding technique for MAX CUT. *Random Struct. Algor.*, 20(3):403–440, 2002. [doi:10.1002/rsa.10036] 3

[17] YUVAL FILMUS: Friedgut-Kalai-Naor theorem for slices of the Boolean cube. *Chicago J. Theoret. Comp. Sci.*, 2016(14):1–17, 2016. [doi:10.4086/cjtcs.2016.014] 4, 8

[18] YUVAL FILMUS: An orthogonal basis for functions over a slice of the Boolean hypercube. *Electr. J. Combinat.*, 23(1):P1.23, 2016. [doi:10.37236/4567] 4, 8

[19] YUVAL FILMUS, GUY KINDLER, ELCHANAN MOSSEL, AND KARL WIMMER: Invariance principle on the slice. *ACM Trans. Comput. Theory*, 10(3):1–37, 2018. Preliminary version in CCC'16. [doi:10.1145/3186590] 4, 8

[20] YUVAL FILMUS AND ELCHANAN MOSSEL: Harmonicity and invariance on slices of the Boolean cube. In *Proc. 31st Comput. Complexity Conf. (CCC'16)*, pp. 16:1–16:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016. [doi:10.4230/LIPIcs.CCC.2016.16] 4, 8

[21] EHUD FRIEDGUT: Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998. [doi:10.1007/PL00009809] 3

[22] EHUD FRIEDGUT: Sharp thresholds of graph properties, and the k-SAT problem. *J. AMS*, 12(4):1017–1054, 1999. With an appendix by JEAN BOURGAIN. [doi:10.1090/S0894-0347-99-00305-7] 8

[23] CHRISTOPHER GODSIL AND KAREN MEAGHER: *Erdős-Ko-Rado theorems: Algebraic approaches*. Cambridge Univ. Press, 2016. [doi:10.1017/CBO9781316414958] 9, 10

[24] LEONARD GROSS: Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975. [doi:10.2307/2373688] 3

[25] TOM GUR, NOAM LIFSHITZ, AND SIQI LIU: Hypercontractivity on high dimensional expanders. In *Proc. 54th STOC*, pp. 176–184. ACM Press, 2022. [doi:10.1145/3519935.3520004] 6

[26] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. [doi:10.1145/502090.502098] 3, 8

[27] SHLOMO HOORY, NATHAN LINIAL, AND AVI WIGDERSON: Expander graphs and their applications. *Bull. AMS*, 43(04):439–562, 2006. [doi:10.1090/s0273-0979-06-01126-8] 2

[28] RUSSELL IMPAGLIAZZO, VALENTINE KABANETS, AND AVI WIGDERSON: New direct-product testers and 2-query PCPs. *SIAM J. Comput.*, 41(6):1722–1768, 2012. [doi:10.1137/09077299X] 8

[29] JEFF KAHN, GIL KALAI, AND NATHAN LINIAL: The influence of variables on Boolean functions (extended abstract). In *Proc. 29th FOCS*, pp. 68–80. IEEE Comp. Soc., 1988. [doi:10.1109/SFCS.1988.21923] 3

[30] DANIEL M. KANE AND RAGHU MEKA: A PRG for Lipschitz functions of polynomials with applications to sparsest cut. In *Proc. 45th STOC*, pp. 1–10. ACM Press, 2013. [doi:10.1145/2488608.2488610] 4

[31] PETER KEEVASH, NOAM LIFSHITZ, EOIN LONG, AND DOR MINZER: Hypercontractivity for global functions and sharp thresholds. *J. AMS*, 37(1):245–279, 2024. [doi:10.1090/jams/1027] 8

[32] SUBHASH KHOT, GUY KINDLER, ELCHANAN MOSSEL, AND RYAN O'DONNELL: Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007. [doi:10.1137/S0097539705447372] 3

[33] SUBHASH KHOT, DOR MINZER, AND MULI SAFRA: On independent sets, 2-to-2 games, and Grassmann graphs. In *Proc. 49th STOC*, pp. 576–589. ACM Press, 2017. [doi:10.1145/3055399.3055432] 5, 8

[34] SUBHASH KHOT, DOR MINZER, AND MULI SAFRA: Pseudorandom sets in Grassmann graph have near-perfect expansion. *Ann. Math.*, 198(1):1–92, 2023. [doi:10.4007/annals.2023.198.1.1] 5, 6, 7, 8

[35] SUBHASH KHOT AND ASSAF NAOR: Nonembeddability theorems via Fourier analysis. *Mathematische Annalen*, 334:821–852, 2006. Preliminary version in FOCS'05. [doi:10.1007/s00208-005-0745-0] 3, 8

[36] 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):8:1–39, 2015. [doi:10.1145/2629614] 3, 8

[37] TZONG-YOW LEE AND HORNG-TZER YAU: Logarithmic Sobolev inequality for some models of random walks. *Ann. Probab.*, 26(4):1855–1873, 1998. [doi:10.1214/aop/1022855885] 4

[38] PRASAD RAGHAVENDRA: Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th STOC*, pp. 245–254. ACM Press, 2008. [doi:10.1145/1374376.1374414] 3

[39] PRASAD RAGHAVENDRA AND DAVID STEURER: Graph expansion and the Unique Games Conjecture. In *Proc. 42nd STOC*, pp. 755–764. ACM Press, 2010. [doi:10.1145/1806689.1806792] 2

[40] PRASAD RAGHAVENDRA, DAVID STEURER, AND MADHUR TULSIANI: Reductions between expansion problems. In *Proc. 27th IEEE Conf. on Comput. Complexity (CCC'12)*, pp. 64–73. IEEE Comp. Soc., 2012. [doi:10.1109/CCC.2012.43] 2

[41] KARL WIMMER: Low influence functions over slices of the Boolean hypercube depend on few coordinates. In *Proc. 29th IEEE Conf. on Comput. Complexity (CCC'14)*, pp. 120–131. IEEE Comp. Soc., 2014. [doi:10.1109/CCC.2014.20] 4

AUTHORS

Subhash Khot
Professor
Courant Institute of Mathematical Sciences
New York University
NY, USA
khot@cs.nyu.edu
https://cs.nyu.edu/~khot/

Dor Minzer
Assistant Professor
Department of Mathematics
Massachusetts Institute of Technology
MA, USA
dminzer@mit.edu
https://sites.google.com/view/dorminzer/home

Dana Moshkovitz
Associate Professor
Department of Computer Science
University of Austin Texas
TX, USA
dana.moshkovitz@gmail.com
https://www.cs.utexas.edu/~danama/

Muli Safra
Professor
Department of Computer Science
Tel-Aviv University
Tel-Aviv, Israel
safra@mail.tau.ac.il
https://sites.google.com/site/mulisafra/home

ABOUT THE AUTHORS

Subhash Khot is a Professor in the Computer Science Department at New York University. He has a Ph. D. from Princeton University, completed in 2003 under the supervision of Sanjeev Arora. His teacher and mentor at Vyankatrao Highschool, Mr. Vaman G. Gogate played a decisive role in directing his attention to mathematics. If not for Mr. Gogate's guidance and some fortuitous turn of events, the chance of someone from the remote town of Ichalkaranji pursuing mathematical research was strictly nil. Subsequently, Mr. Gogate also mentored Amit Deshpande and Raghav Kulkarni (TCS), Abhijit Gadde (Physics), and Dinesh Bharadia (EE). He kept in touch with all his past students and continued to provide guidance on all aspects of life. He passed away in 2020 on Sept 5, which coincidentlly is celebrated as the Teacher's Day in India.

Dor Minzer is an Assistant Professor in the Mathematics Department at Massachusetts Institute of Technology. He has a Ph. D. from Tel Aviv University, completed in 2018 under the guidance of Muli Safra, after which he spent two years as a postdoctoral fellow at the Institute for Advanced Study, Princeton. Dor is interested in Hardness of Approximation, Analysis of Boolean functions and relations to adjacent areas such as extremal combinatorics and error correcting codes. He started his academic studies at the Open University of Israel, where he met Professor Vadim Greenstein who has greatly influenced Dor and his passion for math.

Dana Moshkovitz is an Associate Professor of Computer Science at UT Austin. Her research is in Theoretical Computer Science. Much of it focuses on the limitations of approximation algorithms and probabilistic checking of proofs. Dana did her Ph. D. at the Weizmann Institute in Israel. Her thesis co-won the Nessyahu Prize for best math Ph. D. thesis in Israel in 2009, and part of this work was recognized by the FOCS 2008 Best Paper Award. Dana went on to spend two years at Princeton University and the Institute of Advanced Study before joining MIT as an assistant professor. Dana is the recipient of the Jerome Saltzer teaching award of MIT EECS.

Muli Safra is a Professor of Computer Science at Tel Aviv University. He has a Ph. D. from the Weizmann Institute, completed in 1990 under the supervision of Amir Pnueli. He is one of the recipients of the 2001 Gödel Prize..