Electronic Colloquium on Computational Complexity, Report No. 94 (2017)

On Non-Optimally Expanding Sets in Grassmann Graphs

Irit Dinur * Subhash Khot † Guy Kindler ‡ Dor Minzer § Muli Safra ¶

Abstract

The paper investigates expansion properties of the Grassmann graph, motivated by recent results of [KMS16, DKK+16] concerning hardness of the Vertex-Cover and of the 2-to-1 Games problems. Proving the hypotheses put forward by these papers seems to first require a better understanding of these expansion properties.

We consider the edge expansion of small sets, which is the probability of choosing a random vertex in the set and traversing a random edge touching it, and landing outside the set.

A random small set of vertices has edge expansion nearly 1 with high probability. However, some sets in the Grassmann graph have strictly smaller edge expansion. We present a hypothesis that proposes a characterization of such sets: any such set must be denser inside subgraphs that are by themselves (isomorphic to) smaller Grassmann graphs. We say that such a set is *non-pseudorandom*. We achieve partial progress towards this hypothesis, proving that sets whose expansion is strictly smaller than $7/8$ are non-pseudorandom.

This is achieved through a spectral approach, showing that Boolean valued functions over the Grassmann graph that have significant correlation with eigenspaces corresponding to the top two non-trivial eigenvalues (that are approximately $1/2$ and $1/4$) must be non-pseudorandom.

Contents

*Department of Computer Science and Applied Mathematics, Weizmann Institute. Supported by an ISF-UGC grant 1399/14 and BSF grant 2014371.

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

‡School of Computer Science and Engineering, Hebrew University. Supported by ISF grant no. 1692/13.

§School of Computer Science, Tel Aviv University. Supported by Len Blavatnik and the Blavatnik Family foundation.

¶School of Computer Science, Tel Aviv University. Supported by Len Blavatnik and the Blavatnik Family foundation.

ISSN 1433-8092

1 Introduction

The PCP Theorem [AS98, ALM+98] is a widely applicable, fundamental result in theoretical computer science. One area in particular in which it is vital, is hardness of approximation of optimization problems [FGL+96]. Several useful parameters in PCP constructions were highlighted by the work of Dinur and Safra [DS05], however, the existence of optimal PCPs with respect to these parameters is not known. The Unique-Games Conjecture of Khot [Kho02], which asserts the existence of such PCPs, is a prominent open question in theoretical computer science and in PCP theory in particular. This conjecture has far-reaching consequences — it was shown to imply optimal hardness results for almost all optimization problems [KKMO07, KR08, Rag08]. Research aimed at the design of algorithms for Unique-Games has led to some interesting algorithmic ideas [CMM06, Kol11, ABS15, Tre08]. In contrast, progress towards a proof has been slow – a first candidate for hard instance of Unique-Games has only recently been suggested [KM16].

An independent line of research towards Khot's 2-to-1 Conjecture, a weaker variant of the Unique-Games Conjecture (also from [Kho02]), was initiated recently: First in [KMS16], which focused on the NP-hardness of approximating Vertex-Cover, and then in [DKK+16], where it was extended to 2-to-1 games. A key mathematical object in both constructions is the *Grassmann Graph*, and an associated agreement test for linear functions. The results in [KMS16, DKK+16] relied on conjectured properties of this Grassmann agreement test, specifically they conjectured that when a certain assignment passes the test with non-negligible probability, that assignment must have a large-scale structure consistent with some global linear function.

The motivation of this work is to better understand the structure of the Grassmann graph, with the hope of approaching a proof for the conjectured properties of the associated consistency test. Towards this end, we study properties of sets that do not expand well; such sets play crucial role in all currently known challenging examples for the Grassmann consistency test and their understanding seems to be a prerequisite for proper understanding of the test.

1.1 The Grassmann Graph and the linearity agreement test

Let V be a k dimensional linear space over $\mathbb{F}_2$. The ℓ-dimensional *Grassmann encoding* of a linear function $\mathcal{H}\colon V \to \mathbb{F}_2$ is a table $\mathcal{F}$ of values, assigning to each ℓ-dimensional subspace $L \subseteq V$ the restriction of $\mathcal{H}$ to

it, namely $\mathcal{F}(L) = \mathcal{H}|_L$. For any ℓ-dimensional subspace $L, L' \subseteq V$, $\mathcal{F}(L)$ and $\mathcal{F}(L')$ obviously agree on $L \cap L'$, as they are both consistent with the global $\mathcal{H}$.

Suppose, on the other hand, that we are given a table $\mathcal{F}$ specifying a linear function to every ℓ-dimensional subspace of V and we would like to test whether the table is consistent with some global linear function over V. The *Grassmann consistency test*, used in [KMS16, DKK$^+$16], is as follows: Pick two random ℓ-dimensional subspace $L, L' \subseteq V$ under the constraint that $dim(L \cap L') = \ell - 1$, and verify that $\mathcal{F}(L)$ and $\mathcal{F}(L')$ agree on $L \cap L'$ (the *Grassmann graph* $G(V, \ell)$ is obtained by connecting such L and L' by an edge). The consistency of an assignment $\mathcal{F}$ is the probability it passes the Grassmann test.

As noted above, a Grassmann encoding of a linear function has consistency 1. What can we say if a table $\mathcal{F}$ passes the test with probability δ? It is easy to show that when $\delta = 1$ the table must be an encoding of a global linear function over V. But the case $\delta < 1$ is not nearly as trivial, and when δ is a small positive constant this becomes quite subtle.

Before we continue the discussion about the relation between δ and a global structure, we remark about the relation between the Grassmann test and the 2-to-1 conjecture.

From 2-to-2 to 2-to-1. Note that the Grassmann test is 2-to-2 in the following sense: for every pair L, L' of subspaces that are considered by the test, every linear function on their intersection $L \cap L'$ has exactly two extensions to L that are consistent with it, and also two consistent extensions to L'. One can therefore partition the possible labels of both L and L' into pairs, so the test on L, L' defines a matching between these sets of pairs. A simple technique can transform this 2-to-2 test into a 2-to-1 test as required so as to prove the 2-to-1 conjecture – this is done in [DKK$^+$16].

From passing the test to global structure. Let $\mathcal{F}$ be a table assigning a linear function $\mathcal{F}[L]$ to every ℓ dimensional subspace $L \subseteq V$, and suppose that $\mathcal{F}$ has non-negligible consistency δ, namely it passes the Grassmann test with probability δ. Must $\mathcal{F}$ be consistent with a global linear function $\mathcal{H}\colon V \to \mathbb{F}_2$? At least when interpreted in the most straightforward sense, the answer to the above question is negative: there exists a table $\mathcal{F}$ that has constant consistency $\delta > 0$, such that for every linear function $\mathcal{H}\colon V \to \mathbb{F}_2$ the probability $\mathcal{F}[L], \mathcal{H}|_L$ agree on a randomly chosen ℓ-dimension space L is negligible: $O(2^{\ell-k})$.

Nevertheless, by considering the structure of known examples of the type mentioned above, one can come up with some weaker form of large-scale consistency that might be implied by the fact that a table passes the Grassmann test with non-negligible probability. Indeed, the following hypothesis was proposed in [DKK$^+$16]:

Hypothesis 1.1. *For every $\delta > 0$, there exists $\varepsilon > 0$, and $r > 0$ integer such that the following holds for sufficiently large $k \gg \ell$. Let V be a k dimensional subspaces over $\mathbb{F}_2$, and suppose $\mathcal{F}$ is a labeling of $G(V, \ell)$ by linear functions that has δ consistency with respect to the Grassmann test. Then there exist subspaces $Q, W \subseteq V$, $\dim(Q) = r, \dim(W) = k - r$ and a linear function $\mathcal{H}\colon W \to \mathbb{F}_2$, such that*

$$\Pr_{Q \subseteq L \subseteq W}[\mathcal{F}[L] \equiv \mathcal{H}|_L] \geqslant \varepsilon.$$

In words, if $\mathcal{F}$ has non-negligible consistency, then there exists Q of small dimension and W of small co-dimension, so that $\mathcal{F}$ has non-negligible agreement with a legitimate linear function on subspaces containing Q and contained in W. Considering spaces containing Q is referred to as "zoom-in", and restricting to spaces contained in W is referred to as "zoom-out".

Hypothesis 1.1 was left as an open question in [DKK$^+$16]. A variant of the above hypothesis was made earlier by [KMS16], where it was suggested that studying vertex expansion properties of $G(V, \ell)$ may be a good starting point for studying the Grassmann test.

1.2 The Grassmann test and Edge expansion

In order to prove Hypothesis 1.1 wrong, one needs to construct a labeling of $G(V, \ell)$ that has non-negligible consistency yet does not posses global structure. One possible approach for doing this is by covering the vertices of $G(V, \ell)$ by a large number of less-than-optimally expanding sets.

Definition 1.2. *Let $G = (U, E)$ be a d-regular graph, and $S \subseteq U$ a set of vertices. The edge expansion of S is the fraction of edges in S that go outside of it, namely*

$$\Phi_G(S) = \frac{|E(S, U \setminus S)|}{d\,|S|},$$

where $E(S, U \setminus S)$ denotes the set of edges between S and $U \setminus S$.

Equivalently, the expansion of a set in a regular graph is equal to the probability of picking a uniformly chosen vertex from it, taking a random edge from it and reaching a vertex outside the set.

Take $A_1, ..., A_r$ to be disjoint subsets of vertices of roughly equal sizes that cover a non-negligible fraction $G(V, \ell)$, and suppose that these sets have small expansion. Choosing $\mathcal{H}_i \colon V \to \mathbb{F}_2$ to be random global linear functions, one can construct a table $\mathcal{F}$ by assigning the elements of A_i according to $\mathcal{H}_i$ (vertices which are not in any A_i can be assigned randomly). We now have an assignment where no global linear function agrees with it on significantly more that $\frac{1}{r}$ of the vertices. However, since the set A_i are non-expanding (we leave exact parameters for later), the Grassmann test still has a good chance of picking an edge that lies within one the A_i's, and thus to accept.

If, moreover, one could avoid any zoom-in/zoom-out structure while constructing these sets $A_1, ..., A_r$, then one would effectively refute Hypothesis 1.1. It thus seems that determining if such small non-expanding structure-less sets A_i exist is a prerequisite for resolving Hypothesis 1.1. This paper tries to make a first step to answer this question.

While we do not know a formal way to obtain Hypothesis 1.1 from expansion properties of the Grassmann Graph, there seems to be a strong connection between the two. For example, using the fact that a small set of vertices S in the Grassmann Graph has expansion at least nearly half [1], one can prove the following theorem, proved in Section D.

Theorem 1.3. *For every $\delta > \frac{1}{2}$ there exists $\varepsilon > 0$ such that the following holds for sufficiently large k, ℓ. If $\mathcal{F}$ is a labeling of $G(V, \ell)$ by linear functions that has δ-consistency in the Grassmann test, then there exists a linear function $\mathcal{H} \colon V \to \mathbb{F}_2$ such that*

$$\Pr_L \left[\mathcal{F}[L] \equiv \mathcal{H}|_L\right] \geqslant \varepsilon.$$

Note that Theorem 1.3 implies that Hypothesis 1.1 holds for $\delta > \frac{1}{2}$.This may interpreted as a further suggestion that understanding the structure of non-expanding sets in the Grassmann graph may contribute to resolving Hypothesis 1.1.

1.3 Non-Expanding sets in the Grassmann Graph

Theorem 1.3 relies on the fact that a set whose expansion is very small must be relatively large. But this does not necessarily hold for sets whose expansion is just slightly smaller than optimal, which seems to be the relevant case for Hypothesis 1.1.

[1] This fact is proved later in the paper in Section D.

We would therefore like to understand what we can say about the structure of non-optimally expanding, small subsets of vertices in $G(V, \ell)$. Some natural examples are sets of vertices that induce a subgraph that is by itself (isomorphic to) a Grassmann graph of smaller dimension.

Zoom-out. One type of such sets results from taking the set of subspaces of a hyperplane (or more generally, a small co-dimension subspace), namely taking $G(W, \ell)$ for a hyperplane $W \subseteq V$: Given a vertex in it $L \subseteq W$ and a random edge (L, L'), the probability that this edge stays inside $G(W, \ell)$ is $1/2$. This is seen by observing that L' is obtained from $L \cap L'$ by "adding a random vector", which belongs to W with probability $1/2$ since it contains half of the points in V. Below we refer to these type of examples as *zoom-outs*.

Zoom-in. A different set of examples result from induced subgraphs isomorphic to $G(W, \ell - 1)$ for hyperplane W. For examples consider the graph induced by all vertices that correspond to subspaces that include a particular vector $x \neq 0$. One can observe that the degree of each vertex in this induced subgraph is roughly half of its degree on $G(V, \ell)$: for any L in this subgraph, the probability a random neighbour L' also contains x is roughly half, since L, L' share nearly half of their non-zero vectors.

More generally, the set of vertices that correspond to subspaces that contain some particular r vectors, result in poorly-expanding induced subgraph, which by itself is isomorphic to $G(W', k - r)$, for W' of co-dimension r. Below we refer to these type of examples as *zoom-in*.

1.3.1 Pseudo-random sets

What can we say about the structure of general sets that are non-optimally expanding? The above examples are obviously not general, as once can obtain such sets using a combination of zoom-ins and zoom-outs: Namely take the set of subspaces that are contained in some hyperplane as well as contain few specific vectors. Are there any inherently different non-optimally expanding sets of vertices? The main question of this paper is whether these are, in a sense, the only non-optimally expanding sets (see Hypothesis 1.7 below).

To formulate this question more precisely we need to define the density of sets relative to zoom-ins and zoom-outs.

Definition 1.4. *Let $G = (U, E)$ be a graph. The density of a set of vertices $S \subseteq U$, denoted $\mu(S)$, is the fractional size of S in U.*

Definition 1.5. *Let V be a k-dimensional space over $\mathbb{F}_2$, consider the Grassmann Graph $G(V, \ell)$ and let S be a set of vertices in it. For $Q, W \subseteq V$ of dimensions $\mathsf{dim}(Q) < \ell < \mathsf{dim}(W)$, the density of S among spaces containing Q contained in W is denoted by $\mu_{(Q,\mathsf{in}),(W,\mathsf{out})}(S)$ and equals*

$$\frac{|S \cap \{ L \in G(V, \ell) \mid Q \subseteq L \subseteq W \}|}{|\{ L \in G(V, \ell) \mid Q \subseteq L \subseteq W \}|}.$$

Note that setting $r = \mathsf{dim}(Q)$, the induced subgraph on the set of ℓ-dimensional spaces containing Q and contained in W, is isomorphic to $G(W, \ell - r)$. Thus $\mu_{(Q,\mathsf{in}),(W,\mathsf{out})}(S)$ is nothing but the density of S in that subgraph.

We have seen examples of sets that are non-optimally expanding that are small. However, one may theorise that such a set must be significantly denser inside a zoom-in/zoot-out combination: this would mean

that such a set is, in a sense, close to the examples discussed above. The following definition encapsulates this idea.

Definition 1.6 (Pseudo-randomness)**.** *Let V be a vector space of dimension k over $\mathbb{F}_2$, and consider the Grassmann Graph $G(V,\ell)$. We say a set of vertices S is (m,ε) pseudo-random if for every integers q,r such that $q+r=m$, Q a q-dimensional subspace and W of co-dimension r, $\mu_{(Q,\mathsf{in}),(W,\mathsf{out})}(S) \leqslant \mu(S)+\varepsilon$.*

We are now ready to formulate a precise version of the question leading this paper. It asks whether a small set of vertices that have that have less than optimal expansion, must be correlated with a combination of the above examples.

Hypothesis 1.7. *For every $\eta > 0$ there exist $\delta, r, \varepsilon > 0$, such that for large enough $k \gg \ell \gg 1$ the following holds. Let S is a set of vertices in $G(V,\ell)$ of density at most δ. If S is (r,ε)-pseudo random, then $\Phi(S) \geqslant 1-\eta$.*

Equivalently, the above questions asks whether a small set with expansion bounded away from 1 is necessarily not pseudo-random.

1.4 Main Results

Our main results can be seen as partial answer to Hypothesis 1.7. Below we state slightly informal (and quantitatively weaker) versions of them. Our first result states that a $(1,\varepsilon)$ pseudo-random set has expansion at least close to $3/4$:

Theorem 1.8 (Informal version of Theorems 3.3,3.5)**.** *For sufficiently small $\delta,\eta > 0$, there exists $\varepsilon > 0$ such that the following holds for large enough k,ℓ. If S is a $(1,\varepsilon)$ pseudo random set of vertices in $G(V,\ell)$ of density δ, then $\Phi(S) \geqslant \frac{3}{4}-\eta$.*

Our second result states that a $(2,\varepsilon)$ pseudo-random set has expansion at least close to $7/8$:

Theorem 1.9 (Informal version of Theorem 3.7)**.** *For sufficiently small $\delta,\eta > 0$, there exists $\varepsilon > 0$ such that the following holds for large enough k,ℓ. If S is a $(2,\varepsilon)$ pseudo random set of vertices in $G(V,\ell)$ of density δ, then $\Phi(S) \geqslant \frac{7}{8}-\eta$.*

In light of the above theorems, one would expect that a (r,ε) pseudo-random set must have expansion close to $1-2^{-r}$. While this seems plausible, at this point we are unable to apply our techniques to this general case.

Pushing to the limit. Note that a set S of density δ can potentially have expansion $1-\delta$. This follows by observing that it must contain at least $\delta^2 - o(1)$ fraction of the edges, since this fraction equals

$$\begin{aligned}\mathop{\mathbb{E}}_{A\subseteq V,\mathsf{dim}(A)=\ell-1}\left[\Pr_{L,L'\supseteq A}\left[L,L'\in S\right]\right]-o(1) &= \mathop{\mathbb{E}}_{A\subseteq V,\mathsf{dim}(A)=\ell-1}\left[\Pr_{L\supseteq A}\left[L\in S\right]^2\right]-o(1)\\ &\geqslant \mathop{\mathbb{E}}_{A\subseteq V,\mathsf{dim}(A)=\ell-1}\left[\Pr_{L\supseteq A}\left[L\in S\right]\right]^2-o(1)\\ &= \delta^2-o(1).\end{aligned}$$

In this light, the most ambitious form of Hypothesis 1.7 asserts that if $\Phi(S) \leqslant 1-\delta-\varepsilon$, then S is not pseudo-random.

1.5 Techniques

1.5.1 A spectral approach for expansion

We study the structure of non-expanding sets in $G(V,\ell)$ via spectral analysis: Identifying sets of vertices with their indicator functions, we study representations of functions over $G(V,\ell)$ as sums of eigenvectors.

Since $G(V,\ell)$ is undirected regular graph, we know that the space of real-valued functions over it can be written as a direct sum of orthogonal eigenspaces. Finding such a decomposition is often straightforward for Cayley graphs of nicely structured groups, however this is not the case with $G(V,\ell)$. In Section 2 we show that when $\ell \ll k$, the normalized adjacency operator of $G(V,\ell)$, $A_{G(V,\ell)}$, has eigenspaces $J_{=0}, ..., J_{=\ell}$ with eigenvalues $\lambda_0, ..., \lambda_\ell$, where $\lambda_i \approx 2^{-i}$. The first eigenspace consists of the constant functions, and the corresponding eigenvalue is $\lambda_0 = 1$.

Let S be a set of density δ, and assume $\Phi(S) \leqslant 1 - \delta - \varepsilon$ for $\varepsilon > 0$. Denote by F the indicator function of S. It is easy to see that the quantity $\langle F, A_{G(V,\ell)}F\rangle$ counts the fraction of edges of the graph with both endpoints in S. Therefore, the expansion of the set S equals $1 - \frac{1}{\delta}\langle F, A_{G(V,\ell)}F\rangle$, and the assumption on the expansion of S can be rewritten as

$$\langle F, A_{G(V,\ell)}F\rangle \geqslant (\delta + \varepsilon)\delta.$$

Writing $F = F_{=0} + ... + F_{=\ell}$ where $F_{=i} \in J_{=i}$, the above inequality can be rewritten as:

$$\sum_{i=0}^{\ell} \lambda_i \langle F_{=i}, F_{=i}\rangle \geqslant (\delta + \varepsilon)\delta.$$

Noting that the summand corresponding to $i = 0$ equals δ^2, we thus have

$$\sum_{i=1}^{\ell} \lambda_i \langle F_{=i}, F_{=i}\rangle \geqslant \delta\varepsilon. \tag{1}$$

Recalling that $\lambda_i \approx 2^{-i}$, we see that for (1) to hold the weight of F on low-index eigenspaces must be significant. In conclusion, F must have non-negligible correlation with an eigenspace $J_{=i}$ for relatively small i. If we could show that having significant weight on low-index eigenspaces implies not being pseudo-random, we would be done.

We do not prove such a general statement for any low-index eigenspace – the bulk of this paper is focused on partial results, namely proving such a statement specifically for the cases $i = 1$ and $i = 2$. Theorem 1.8 and Theorem 1.9 are relatively direct corollaries of these cases.

1.5.2 Correlation with small order eigenspace implies non pseudo-randomness

We would like to describe the main ideas utilized in the proof that if F has significant weight on one of the levels 1 or 2, it must not be pseudo-random. We need some observations about these eigenspaces first.

Structure of eigenspaces We have already seen that $J_{=0}$ consists of constant functions. We can also give a simple characterization of functions in $J_{=1}$ — these are the functions of the form

$$G[L] = \sum_{x \in L \setminus \{0\}} g_{=1}(x),$$

where $g_{=1} \colon V \setminus \{0\} \to \mathbb{R}$ has average 0. Similar characterizations in fact hold for all eigenspaces, as discussed in Section 2.

Idea of the proof Let us focus on the case $i = 1$, and for simplicity let us further assume that F is entirely supported on $J_{=0}$ and $J_{=1}$ (as opposed to only having non-negligible correlation with $J_{=1}$) – that is, $F = F_{=0} + F_{=1}$. Note that $F_{=0} \equiv \delta$ is constant, and since F is Boolean valued, we have that $F_{=1}$ receives only two distinct, far apart values: $-\delta$ and $1 - \delta$. As previously noted, we can write $F_{=1} \in J_{=1}$ as

$$F_{=1}[L] = \sum_{x \in L \setminus \{0\}} f_{=1}(x) = \sum_{x \in V \setminus \{0\}} \mathbb{1}_{x \in L} f_{=1}(x),$$

for some $f_{=1} \colon V \setminus \{0\} \to \mathbb{R}$ with average 0.

Consider this sum from a probabilistic perspective: think of L as being uniformly chosen among the vertices of $G(V, \ell)$, and of the summands $\mathbb{1}_{x \in L} f_{=1}(x)$ as random variables. Note that since ℓ-dimensional subspaces are almost pairwise independent (namely the probability that a random L contains two distinct points is roughly the square of the probability of containing one point), we have that $F_{=1}[L]$ is a sum of almost pairwise independent random variables. In fact, since $\ell > 4$ the summands which correspond to linearly independent points are even nearly 4-wise independent.

We would thus expect $F_{=1}[L]$ to have a fourth moment which roughly equals the square of its second moment, unless one of the following cases occur: either there is a large contribution to $F_{=1}[L]$ from summands $\mathbb{1}_{x \in L} f_{=1}(x)$ that correspond to x's that are *not* linearly independent; or else there are some x's such that the contribution of $\mathbb{1}_{x \in L} f_{=1}(x)$ is significantly higher than that of typical summands.

Roughly speaking, we show that in the first case the values of $f_{=1}$ must be non-negligibly correlated with a hyperplane $W \subseteq V$. In turn, we show that this implies that zooming out on W increases the density of S significantly. This last step is done by viewing $f_{=1}$ as a function over a hypercube and analyzing its fourier coefficients.

The second case turns out to be rather straightforward: in that case $f_{=1}$ must have high magnitude on at least one $x \in V$. We show that in this case zooming-in to x increases the density of S significantly.

2 Preliminaries

In this section we present the necessary background on the Grassmann Graph.

Definition 2.1. *Let V be a vector space of dimension k over $\mathbb{F}_2$, and let $0 \leqslant \ell \leqslant k$ be an integer. The Grassmann Graph $G(V, \ell)$ is defined as follows: its vertex set is the set of all ℓ dimensional subspaces of V, and two vertices L, L' are connected by an edge if* $\dim(L \cap L') = \ell - 1$.

This paper largely deals with real-valued functions on the vertices of the Grassmann Graph, $F \colon G(V, \ell) \to \mathbb{R}$. We often consider normalized adjacency operator of $G(V, \ell)$, $A_{G(V,\ell)}$ [2] and consider its action on real-valued functions F on $G(V, \ell)$ [3]:

$$(A_{G(V,\ell)} F)[L] = \mathop{\mathbb{E}}_{L' : \dim(L \cap L') = \ell - 1} \big[F[L']\big].$$

[2] That is, an n by n matrix (n is the number of vertices in $G(V, \ell)$) where $A_{G(V,\ell)}(u, v) = \frac{1}{d}$ when u, v are connected by an edge and otherwise 0, when d is the degree of $G(V, \ell)$.

[3] Which can be viewed as vectors in $\mathbb{R}^{G(V,\ell)}$.

2.1 Spectral analysis on the Grassmann Graph

The Grassmann Graph has been considered in the context of spectral theory [Sri12, Sri14, FW86], however its analysis is more complex than the spectral analysis done on the hypercube [4]. One key difference is that there is no "canonical orthonormal spectral basis" for the transition operator $A_{G(V,\ell)}$; thus one has to settle for a less explicit "block decomposition". It is folklore fact that such decomposition exists and can be obtained; for the sake of completeness, we develop such a block decomposition in this section.

We endow the space of real-valued functions over $G(V,\ell)$ with the inner product

$$\langle F, G\rangle = \mathop{\mathbb{E}}_{L\in G(V,\ell)}[F[L]G[L]].$$

Definition 2.2. *Let $i \geqslant 0$. We say a function $F\colon G(V,\ell)\to\mathbb{R}$ is spanned by the first i levels, if there is $f\colon \binom{V}{i}\to\mathbb{R}$ such that*

$$F[L] = \sum_{\substack{L_i\subseteq L\\ \mathsf{dim}(L_i)=i}} f(L_i).$$

We denote the set of functions spanned by the first i levels by $J_{\leqslant i}$.

Note that $J_{\leqslant \ell}$ is the set of all real-valued functions on $G(V,\ell)$.

Let us look at the first few levels. The 0-level space, namely $J_{\leqslant 0}$, contains only constant functions. How do functions from $J_{\leqslant 1}$ look like? First, observe there exists a natural identification of $V\setminus\{0\}$ with $\binom{V}{1}$ by the mapping $x\to\mathsf{Span}(\{x\})$, thus functions spanned by level $0,1$ are determined by functions assigning values to $V\setminus\{0\}$.

One natural example for a function in $J_{\leqslant 1}$, is the indicator function of the set $\{L\in G(V,\ell)\mid x\in L\}$ for any non-zero $x\in V$. Another important example is the following: Consider $W\subseteq V$ a hyperplane (i.e. subspace of dimension $(\mathsf{dim}(V)-1)$), and consider F the indicator function of the set $\{L\in G(V,\ell)\mid L\subseteq W\}$. Intuitively, this function is of level one, since $L\subseteq W$ depends solely on L being perpendicular to $W^\perp$, which is a subspace of dimension 1. More precisely, define $f_1\colon V\setminus\{0\}\to\mathbb{R}$ by

$$f_1(x) = \begin{cases} 2^{-\ell} & x\in W,\\ -2^{-\ell} & x\notin W.\end{cases}$$

Then if $L\subseteq W$,

$$\sum_{x\in L\setminus\{0\}} f_1(x) = (2^\ell-1)2^{-\ell} = 1-2^{-\ell}.$$

If $L\not\subseteq W$, then it contains $2^{\ell-1}-1$ non-zero points from W and $2^{\ell-1}$ non-zero points outside W, thus

$$\sum_{x\in L\setminus\{0\}} f_1(x) = (2^{\ell-1}-1)2^{-\ell} - 2^{\ell-1}2^{-\ell} = -2^{-\ell}.$$

In conclusion,

$$2^{-\ell} + \sum_{x\in L\setminus\{0\}} f_1(x) = \begin{cases} 1 & L\subseteq W,\\ 0 & L\not\subseteq W.\end{cases}$$

The left hand side can be rewritten as $\sum_{x\in L\setminus\{0\}}\frac{1}{(2^\ell-1)2^\ell} + f_1(x)$, and $F\in J_{\leqslant 1}$.

[4] And more in the flavor of the spectral analysis done on the slice of the hypercube [Fil16b, Fil16a, FKMW16, FM16] with key distinctions.

Fact 2.3. *Let $k > \ell > 0$ be integers. Then for every $0 \leqslant i \leqslant \ell - 1$, $J_{\leqslant i} \subseteq J_{\leqslant i+1}$.*

Proof. Let $F \in J_{\leqslant i}$ be given by $F[L] = \sum_{L_i \subseteq L} f(L_i)$ for $f\colon \binom{V}{i} \to \mathbb{R}$. Note that each i-dimensional space L_i and ℓ-dimensional space L, there are precisely $2^{\ell-i+1} - 1$ $(i+1)$ dimensional subspaces L_{i+1} such that $L_i \subseteq L_{i+1} \subseteq L$. Define $g\colon G(V, i+1) \to \mathbb{R}$ by

$$g(L_{i+1}) = \frac{1}{2^{\ell-i+1} - 1} \sum_{L_i \subseteq L_{i+1}} f(L_i).$$

Then for any $L \in G(V, \ell)$,

$$\sum_{L_{i+1} \subseteq L} g(L_{i+1}) = \frac{1}{2^{\ell-i+1} - 1} \sum_{L_{i+1} \subseteq L} \sum_{L_i \subseteq L_{i+1}} f(L_i) = \sum_{L_i \subseteq L} f(L_i) = F[L],$$

the second equality holds since L_i is counted $2^{\ell-i+1} - 1$ times. Therefore by definition $F \in J_{\leqslant i+1}$. □

We will be interested in functions strictly on the i-th level, namely functions in $J_{\leqslant i}$ that are perpendicular to $J_{\leqslant i-1}$.

Definition 2.4. *The set of level i functions is*

$$J_{=i} = J_{\leqslant i} \cap \left(J_{\leqslant i-1}\right)^{\perp}.$$

It is a well known fact that the sets $J_{=i}$'s are eigenspaces of $A_{G(V,\ell)}$. For the sake of completeness we provide a proof later in this section.

Unraveling the definitions,

$$J_{\leqslant \ell} = J_{=\ell} \oplus J_{\leqslant \ell-1} = \cdots = J_{=\ell} \oplus J_{=\ell-1} \oplus \cdots \oplus J_{=0},$$

and thus we can decompose each function over $G(V, \ell)$ according to its projection to those subspaces.

Definition 2.5. *Let V be a vector space over $\mathbb{F}_2$ of dimension k, and $0 < \ell < k$ an integer. For a function $F\colon G(V, \ell) \to \mathbb{R}$ and $i \in \{0, 1, \ldots, \ell\}$ we define $F_{=i}$ to be the projection of F onto $J_{=i}$. For $i \leqslant \ell - 1$, we define $f_{=i}\colon \binom{V}{i} \to \mathbb{R}$ to be the function such that*

$$F_{=i}[L] = \sum_{L_i \subseteq L} f_{=i}(L_i).$$

Using the notation in the above definition, the block decomposition we are interested in is $F = F_{=0} + \ldots + F_{=\ell}$. For our purposes we require more information about this decomposition. Coming up with exact formulas for the projections $F_{=i}$ turns out to be tricky; instead we work out simple approximations for each $F_{=i}$.

In the remainder of this section we show that each $J_{=i}$ is an eigenspace of $A_{G(V,\ell)}$, present exact formulas for the projection onto $J_{=0}, J_{=1}$, and develop the approximations for all levels. Towards this end, we first define the notions of zoom-ins and zoom-outs that play an important role in the projections.

2.2 Zoom ins and zoom outs

We will often be interested in the effect of events of the form $L \ni x$ or $L \subseteq W$ on the average of F. For that, we introduce the notion of zoom-ins and zoom-outs.

Definition 2.6 (Zoom-in)**.** *Let $F\colon G(V,\ell) \to \mathbb{R}$ be a function, and $Q \subseteq V$ be a q-dimensional subspace, where $q \leqslant \ell$. Define*

$$\mu_{Q,\mathsf{in}}(F) = \mathop{\mathbb{E}}_{\substack{L \in G(V,\ell) \\ Q \subseteq L}} [F[L]].$$

In words, it is the average of the function F on all subspaces that contain Q.

The following claim is easy to verify and will be used throughout this section. We omit the proof.

Claim 2.7. *Let Q be a q-dimensional subspace, and $q' > q$ be an integer. Then*

$$\mathop{\mathbb{E}}_{\substack{Q' \in G(V,q') \\ Q \subseteq Q'}} \left[\mu_{Q',\mathsf{in}}(F)\right] = \mu_{Q,\mathsf{in}}(F).$$

Definition 2.8 (Zoom-Out)**.** *Let $F\colon G(V,\ell) \to \mathbb{R}$ be a function, and a $W \subseteq V$ a subspace of dimension at least ℓ. Define*

$$\mu_{W,\mathsf{out}}(F) = \mathop{\mathbb{E}}_{L \subseteq W} [F[L]].$$

In words, it is the average of the function F on all subspaces that are contained in W - namely on the subgraph $G(W,\ell)$.

Definition 2.9. *Let $F\colon G(V,\ell) \to \mathbb{R}$. For subspaces $Q, W \subseteq V$ where $\mathsf{dim}(Q) \leqslant \ell \leqslant \mathsf{dim}(W)$, define*

$$\mu_{(Q,\mathsf{in}),(W,\mathsf{out})}(F) = \mathop{\mathbb{E}}_{Q \subseteq L \subseteq W} [F[L]].$$

In words, it is the average of F on subspaces contained in W and containing Q.

Gaussian Coefficients

The Gaussian Binomial Coefficients $\begin{bmatrix} k \\ i \end{bmatrix}_p$ count the number of i-dimensional subspaces of $\mathbb{F}_p^k$. As we are only interested in $p = 2$ throughout this paper, we omit the p subscript.

Definition 2.10. *Let $k \geqslant i \geqslant 0$ be integers. The Gaussian Binomial Coefficient is*

$$\begin{bmatrix} k \\ i \end{bmatrix} = \frac{(2^k - 2^0)\cdots(2^k - 2^{i-1})}{(2^i - 2^0)\cdots(2^i - 2^{i-1})}.$$

We will often abuse notation, and denote the set of i-dimensional linear subspaces of V by $\begin{bmatrix} V \\ i \end{bmatrix}$.

Observe that the Gaussian coefficients have the same symmetry as the binomial coefficients, namely $\begin{bmatrix} k \\ i \end{bmatrix} = \begin{bmatrix} k \\ k-i \end{bmatrix}$ for every i, k. This follows by the natural bijection $A \mapsto A^\perp$ of i dimensional subspaces with $k - i$ dimensional subspaces. They posses additional properties of the binomial coefficients, such as unimodality with mode at $k/2$, and the following Pascal-type identity.

Fact 2.11.

$$\begin{bmatrix} k \\ i \end{bmatrix} = \begin{bmatrix} k-1 \\ i \end{bmatrix} + 2^{k-i}\begin{bmatrix} k-1 \\ i-1 \end{bmatrix} = 2^i\begin{bmatrix} k-1 \\ i \end{bmatrix} + \begin{bmatrix} k-1 \\ i-1 \end{bmatrix}.$$

Proof. Both equalities are easy to prove algebraically, however for clarity we provide combinatorial proofs. Let V be a k-dimensional space over $\mathbb{F}_2$ and let $W \subseteq V$ be of co-dimension 1. We count the number of i dimensional subspaces of V in two ways. Clearly, by definition it is $\begin{bmatrix} k \\ i \end{bmatrix}$. On the other hand those spaces can be partitioned to two: there are $\begin{bmatrix} k-1 \\ i \end{bmatrix}$ subspaces contained in W, and $\begin{bmatrix} k-1 \\ i-1 \end{bmatrix}(\begin{bmatrix} k-i+1 \\ 1 \end{bmatrix} - \begin{bmatrix} k-i \\ 1 \end{bmatrix})$ not contained in W (the first factor chooses the intersection with W - which is $i-1$ dimensional, and the second chooses the last basis vector without over-counting). This shows

$$\begin{bmatrix} k \\ i \end{bmatrix} = \begin{bmatrix} k-1 \\ i \end{bmatrix} + 2^{k-i} \begin{bmatrix} k-1 \\ i-1 \end{bmatrix}.$$

The second equality follows by plugging in $i = k - j$, yielding

$$\begin{bmatrix} k \\ j \end{bmatrix} = \begin{bmatrix} k \\ i \end{bmatrix} = \begin{bmatrix} k-1 \\ i \end{bmatrix} + 2^{k-i} \begin{bmatrix} k-1 \\ i-1 \end{bmatrix} = \begin{bmatrix} k-1 \\ j-1 \end{bmatrix} + 2^{j} \begin{bmatrix} k-1 \\ j \end{bmatrix}.$$

□

The following fact can be verified by an easy calculation.

Fact 2.12.

$$\frac{\begin{bmatrix} \ell-1 \\ i \end{bmatrix}}{\begin{bmatrix} \ell \\ i \end{bmatrix}} = 2^{-i} \frac{2^\ell - 2^i}{2^\ell - 1} = 2^{-i} + O(2^{-\ell}).$$

Instead of calculating, it is often useful to consider such terms probabilistically. Fix L to be an ℓ-dimensional space, $L' \subseteq L$ to be an $\ell - 1$ dimensional subspace, and let us pick an i-dimensional subspace of L uniformly at random. What is the probability it is contained in L'? Clearly, it is the expression on the left hand side of the above Fact. On the other hand, picking an i dimensional subspace of L amounts to picking $v_1, ..., v_i$ linearly independent. If we ignore the dependency in choice of the v's, the probability they all fall in L' is $\approx 2^{-i}$, since it has roughly half the number of non-zero points of L.

2.3 Eigenspaces of the Grassmann Graph

In this subsection we characterize the eigenvalues and eigenspaces of $A_{G(V,\ell)}$. We show below that the eigenvalues of this operator can be expressed as follows.

Definition 2.13. *For integers $0 \leqslant i \leqslant \ell - 1, \ell \leqslant k$, denote*

$$\lambda_i(k,\ell) \stackrel{def}{=} \frac{\begin{bmatrix} \ell-1 \\ i \end{bmatrix}}{\begin{bmatrix} \ell \\ i \end{bmatrix}} - \frac{(\begin{bmatrix} \ell \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix})(\begin{bmatrix} \ell \\ i \end{bmatrix} - \begin{bmatrix} \ell-1 \\ i \end{bmatrix})}{\begin{bmatrix} \ell \\ i \end{bmatrix}(\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} \ell \\ 1 \end{bmatrix})},$$

For $i = \ell$, denote

$$\lambda_\ell(k,\ell) \stackrel{def}{=} -\frac{1}{\begin{bmatrix} k-\ell+1 \\ 1 \end{bmatrix} - 1}.$$

We often write just λ_i when k, ℓ are clear from context. Morally speaking, we encourage the reader should think of λ_i as 2^{-i} for $i \leqslant \ell - 1$ and as 0 for $i = \ell$. More precisely, a basic calculation can show that the following holds:

Fact 2.14. *For $i \leqslant \ell - 1$, $k \geqslant 2\ell$, $\ell \geqslant 2$,*

$$\lambda_i = \frac{\begin{bmatrix}\ell-1\\ i\end{bmatrix}}{\begin{bmatrix}\ell\\ i\end{bmatrix}} - \frac{(2^{\ell-i+1}-1)(2^i-1)}{(2^\ell-1)(2^{k-\ell+1}-2)},$$

and in particular

$$2^{-i} - 2^{2-\ell} \leqslant \lambda_i \leqslant 2^{-i}.$$

Fact 2.15. *Suppose $k \geqslant 2\ell$. Then $\lambda_0 > \lambda_1 > .... > \lambda_\ell$.*

Proof. We first prove that $\lambda_i > \lambda_{i+1}$ for all $0 \leqslant i \leqslant \ell - 2$. By the expression for λ_i in Fact 2.14, we get that

$$\lambda_i = \frac{2^{k-i+1} + 2^i - 2^{k-\ell+1} - 2^{\ell+1} + 1}{(2^\ell - 1)(2^{k-\ell+1} - 2)}.$$

A close inspection reveals that this formula is correct for $i = \ell$ as well. Since only the numerator depends on i and is decreasing as long as $k - i + 1 > i$, i.e. as long as $i \leqslant k/2$ (which holds since $i \leqslant \ell \leqslant k/2$), we have that λ_i is decreasing in $0, ..., \ell$. □

The main objective of this section is to prove the following theorem:

Theorem 2.16. *Let $k, \ell > 0$ and $0 \leqslant i \leqslant \ell$ be integers. If $k \geqslant 7\ell^2 + 1$ [5], then $J_{=i}$ is an eigenspace of $A_{G(V,\ell)}$ with eigenvalue $\lambda_i(k,\ell)$ and dimension $\begin{bmatrix}k\\ i\end{bmatrix} - \begin{bmatrix}k\\ i-1\end{bmatrix}$.*

The following claim identifies a (rather straightforward) spanning set of $J_{\leqslant i}$.

Claim 2.17. *Let $k > 2\ell > 0$ be integers. Then $\{G_{L_i}\}_{L_i \in \begin{bmatrix}V\\ i\end{bmatrix}}$ spans $J_{\leqslant i}$, where*

$$G_{L_i}[L] = \begin{cases} 1 & L_i \subseteq L, \\ 0 & \textit{else}. \end{cases}$$

Proof. Let $F \in J_{\leqslant i}$. Then there is $f\colon \begin{bmatrix}V\\ i\end{bmatrix} \to \mathbb{R}$ such that

$$F[L] = \sum_{L_i \in L} f(L_i) = \sum_{L_i} f(L_i) G_{L_i}(L) \in \mathsf{Span}(\{G_{L_i}\}_{L_i \in \begin{bmatrix}V\\ i\end{bmatrix}}).$$

Let $L_i \in \begin{bmatrix}V\\ i\end{bmatrix}$ and define $f\colon \begin{bmatrix}V\\ i\end{bmatrix} \to \mathbb{R}$ by $f(R_i) = 1$ i $R_i = L_i$, and otherwise 0. Then

$$G_{L_i}[L] = \sum_{R_i \in L} f(R_i) \in J_{\leqslant i}.$$

□

The following technical claim asserts that the average of a function $F \in J_{\leqslant i}$ over spaces containing L_{i-1} can be expressed in terms of averages of f over spaces containing subspaces of L_{i-1}. The important point (and the way it is applied later on) is that any information about near-orthogonality of F to $J_{\leqslant i-1}$ can be expressed in terms of properties of f, and vice versa.

[5] This condition is not needed and is artifact of the proof presented herein. We present this proof since some of the elements in it are needed in later proofs.

Claim 2.18. *Suppose $0 \leqslant j < i \leqslant \ell$, $k \geqslant 7\ell^2+1$. There exists $\beta_0, ..., \beta_j \in \mathbb{R}$ such that the following holds. For every $F \in J_{\leqslant i}$ given by $F[L] = \sum_{R_i \subseteq L} f(R_i)$ and $L_j \in \binom{V}{j}$,*

$$\mathop{\mathbb{E}}_{L \supseteq L_j}[F[L]] = \left(\begin{bmatrix} \ell - j \\ i - j \end{bmatrix} + \beta_j \right) \cdot \mu_{L_j,\mathsf{in}}(f) + \sum_{r=0}^{j-1} \beta_r \sum_{R_r \subseteq L_j} \mu_{R_r,\mathsf{in}}(f).$$

Additionally, the β's have the following properties:

- *for $r = 0, ..., j-1$, $|\beta_r| \leqslant 2^{6\ell^2}$.*
- *$|\beta_j| \leqslant 2^{7\ell^2 - k}$.*

In particular, the coefficient of $\mu_{L_j,\mathsf{in}}(f)$ is not 0.

We defer the proof of this claim to Section B.

Lemma 2.19. *Suppose $i \leqslant \ell$, $k \geqslant 7\ell^2+1$. Let $F \in J_{\leqslant i}$ be given by $F[L] = \sum_{R_i \subseteq L} f(R_i)$.*

Then $F \in J_{=i}$ if and only if for every $j \leqslant i-1$ and $L_j \in \binom{V}{j}$, $\mu_{L_j,\mathsf{in}}(f) = 0$.

Proof. Suppose that for every $j \leqslant i-1$ and $L_j \in \binom{V}{j}$ we have $\mu_{L_j,\mathsf{in}}(f) = 0$. Then by Claim 2.18 we have that for every $L_{i-1} \in \binom{V}{i-1}$,

$$\mathop{\mathbb{E}}_{L \supseteq L_{i-1}}[F[L]] = \beta_{i-1} \cdot \mu_{L_{i-1},\mathsf{in}}(f) + \sum_{r=0}^{i-2} \beta_r \sum_{R_r \subseteq L_{i-1}} \mu_{R_r,\mathsf{in}}(f) = 0.$$

Hence F is perpendicular to $G_{L_{i-1}}$ from Claim 2.17, and so $F \in (J_{\leqslant i-1})^{\perp}$. It follows that $F \in J_{=i}$.

Suppose that $F \in J_{=i}$. We prove by induction on j that $\mu_{L_j,\mathsf{in}}(f) = 0$ for all $j \leqslant i-1$ and $L_j \in \binom{V}{j}$. We may assume that $i \geqslant 1$, otherwise the claim holds trivially. For $j = 0$, we get that from Claim 2.18 (applied to $j = 0$ and $L_j = \{0\}$) that

$$\mathop{\mathbb{E}}_{L}[F[L]] = \beta_0 \cdot \mu(f).$$

Since $F \in J_{=i}$, it is perpendicular to constant functions, that is $\mathbb{E}_L[F[L]] = 0$. Since $\beta_0 \neq 0$ we conclude that $\mu(f) = 0$. Assume the claim is true for all $j \leqslant n$ where $n < i-1$, and prove for $j \leqslant n+1$. Fix $L_{n+1} \in \binom{V}{n+1}$, then by Claim 2.18,

$$\mathop{\mathbb{E}}_{L \supseteq L_{n+1}}[F[L]] = \beta_{n+1} \cdot \mu_{L_{n+1},\mathsf{in}}(f) + \sum_{r=0}^{n} \beta_r \sum_{R_r \subseteq L_{i-1}} \mu_{R_r,\mathsf{in}}(f),$$

where $\beta_{n+1} \neq 0$. The left hand side is, up to normalization, the inner product of F with $G_{L_{n+1}}$, which is 0 since $F \in J_{=i}$. By the induction hypothesis, $\mu_{R_r,\mathsf{in}}(f) = 0$ for every R_r such that $\mathsf{dim}(R_r) \leqslant n$. We conclude that the last equation implies

$$\beta_{n+1} \cdot \mu_{L_{n+1},\mathsf{in}}(f) = 0,$$

since $\beta_{n+1} \neq 0$, we conclude that $\mu_{L_{n+1},\mathsf{in}}(f) = 0$. □

We next show that $J_{=i}$ is an eigenspace for all $i = 0, ..., \ell$. The proof considers the cases $i \leqslant \ell - 1$ and $i = \ell$ separately.

Lemma 2.20. *Suppose $k \geqslant 7\ell^2 + 1$, and let $F_i \in J_{=i}$. Then for every $0 \leqslant i \leqslant \ell - 1$, $L \in \begin{bmatrix} V \\ \ell \end{bmatrix}$,*

$$(A_{G(V,\ell)}F_i)[L] = \lambda_i(k,\ell)F_i[L].$$

Proof. Denote $F_i[L] = \sum_{L_i \subseteq L} f_i(L_i)$ for some $f_i \colon \begin{bmatrix} V \\ i \end{bmatrix} \to \mathbb{R}$.

$$(A_{G(V,\ell)}F_i)[L] = \mathop{\mathbb{E}}_{L',\, \dim(L\cap L')=\ell-1} \left[F_i[L']\right] = \mathop{\mathbb{E}}_{L',\, \dim(L\cap L')=\ell-1} \left[\sum_{L_i \subseteq L'} f_i(L_i)\right].$$

In the last sum, consider L_i contained in $L \cap L'$ and those that are not separately; let us denote their contributions by A, B respectively.

$$A = \mathop{\mathbb{E}}_{L',\, \dim(L\cap L')=\ell-1} \left[\sum_{L_i \subseteq L' \cap L} f_i(L_i)\right] = \begin{bmatrix} \ell - 1 \\ i \end{bmatrix} \mathop{\mathbb{E}}_{L',\, \dim(L\cap L')=\ell-1} \left[\mathop{\mathbb{E}}_{L_i \subseteq L' \cap L} [f_i(L_i)]\right]$$

Note that L_i is distributed uniformly over all i-dimensional subspaces of L. Therefore,

$$A = \begin{bmatrix} \ell - 1 \\ i \end{bmatrix} \mathop{\mathbb{E}}_{L_i \subseteq L} [f_i(L_i)] = \frac{\begin{bmatrix} \ell-1 \\ i \end{bmatrix}}{\begin{bmatrix} \ell \\ i \end{bmatrix}} F_i[L].$$

As to the contribution of B, write

$$B = \mathop{\mathbb{E}}_{L',\, \dim(L\cap L')=\ell-1} \left[\sum_{L_i \subseteq L', L_i \not\subseteq L} f_i(L_i)\right] = \left(\begin{bmatrix} \ell \\ i \end{bmatrix} - \begin{bmatrix} \ell - 1 \\ i \end{bmatrix}\right) \mathop{\mathbb{E}}_{L',\, \dim(L\cap L')=\ell-1} \left[\mathop{\mathbb{E}}_{L_i \subseteq L', L_i \not\subseteq L} [f_i(L_i)]\right].$$

Each contributing L_i intersects $L \cap L'$ in subspace of dimension $i - 1$; we partition B according to this intersection $R_{i-1} = L \cap L' \cap L_i$. We have that

$$B = \left(\begin{bmatrix} \ell \\ i \end{bmatrix} - \begin{bmatrix} \ell - 1 \\ i \end{bmatrix}\right) \mathop{\mathbb{E}}_{R_{i-1} \subseteq L} \left[\mathop{\mathbb{E}}_{x \notin L} [f_i(R_{i-1} \oplus \mathsf{Span}(x))]\right] \tag{2}$$

Denote the last expectation by E. By Lemma 2.19, since $F_i \in J_{=i}$ for every $R_{i-1} \in \begin{bmatrix} V \\ i-1 \end{bmatrix}$ we have $\mathbb{E}_{L_i \supseteq R_{i-1}} [f_i(L_i)] = 0$. Therefore using conditional expectation:

$$\begin{aligned}
0 &= \mathop{\mathbb{E}}_{R_{i-1} \subseteq L} \left[\mathop{\mathbb{E}}_{L_i \supseteq R_{i-1}} [f_i(L_i)]\right] \\
&= \mathop{\mathbb{E}}_{R_{i-1} \subseteq L} \left[\mathop{\mathbb{E}}_{x \notin R_{i-1}} [f_i(R_{i-1} \oplus \mathsf{Span}(x))]\right] \\
&= \Pr_{\substack{R_{i-1} \subseteq L \\ x \notin R_{i-1}}} [x \in L] \mathop{\mathbb{E}}_{\substack{R_{i-1} \subseteq L \\ x \in L \setminus R_{i-1}}} [f_i(R_{i-1} \oplus \mathsf{Span}(x))] + \Pr_{\substack{R_{i-1} \subseteq L \\ x \notin R_{i-1}}} [x \notin L] \cdot \mathop{\mathbb{E}}_{\substack{R_{i-1} \subseteq L \\ x \notin L}} [f_i(R_{i-1} \oplus \mathsf{Span}(x))] \\
&= \frac{\begin{bmatrix} \ell \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix}}{\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix}} \cdot \frac{1}{\begin{bmatrix} \ell \\ i \end{bmatrix}} \sum_{L_i \subseteq L} f_i(L_i) + \frac{\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} \ell \\ 1 \end{bmatrix}}{\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix}} E \\
&= \frac{\begin{bmatrix} \ell \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix}}{\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix}} \cdot \frac{1}{\begin{bmatrix} \ell \\ i \end{bmatrix}} F_i[L] + \frac{\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} \ell \\ 1 \end{bmatrix}}{\begin{bmatrix} k \\ 1 \end{bmatrix} - \begin{bmatrix} i-1 \\ 1 \end{bmatrix}} E
\end{aligned}$$

Rearranging yields

$$E = -\frac{\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}i-1\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}i-1\\1\end{bmatrix}} \frac{1}{\begin{bmatrix}\ell\\i\end{bmatrix}} F_i[L].$$

Plugging this into equation (2) yields

$$B = -\frac{(\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}i-1\\1\end{bmatrix})(\begin{bmatrix}\ell\\i\end{bmatrix} - \begin{bmatrix}\ell-1\\i\end{bmatrix})}{(\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}i-1\\1\end{bmatrix})\begin{bmatrix}\ell\\i\end{bmatrix}} F_i[L]$$

and overall

$$A + B = \left(\frac{\begin{bmatrix}\ell-1\\i\end{bmatrix}}{\begin{bmatrix}\ell\\i\end{bmatrix}} - \frac{(\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}i-1\\1\end{bmatrix})(\begin{bmatrix}\ell\\i\end{bmatrix} - \begin{bmatrix}\ell-1\\i\end{bmatrix})}{(\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}i-1\\1\end{bmatrix})\begin{bmatrix}\ell\\i\end{bmatrix}} \right) F_i[L] = \lambda_i(k,\ell) F_i[L].$$

□

Lemma 2.21. *Let $F_\ell \in J_{=\ell}$. For every $L \in \begin{bmatrix}V\\\ell\end{bmatrix}$,*

$$(A_{G(V,\ell)} F_\ell)[L] = \lambda_\ell(k,\ell) F_\ell[L].$$

Proof. Fix L. By definition,

$$(A_{G(V,\ell)} F_\ell)[L] = \mathop{\mathbb{E}}_{R_{\ell-1} \subseteq L} \left[\mathop{\mathbb{E}}_{x \notin L} [F[R_{\ell-1} \oplus \mathsf{Span}(x)]] \right],$$

denote the expression on the right hand side by E.

The rest of the proof examines the following quantity, and computes it in two ways. On the one hand, using conditional expectation,

$$\mathop{\mathbb{E}}_{R_{\ell-1} \subseteq L} \left[\mu_{R_{\ell-1},\mathsf{in}}(F_\ell) \right]$$
$$= \mathop{\mathbb{E}}_{R_{\ell-1} \subseteq L} \left[\Pr_{x \notin R_{\ell-1}} [x \notin L] \cdot \mathop{\mathbb{E}}_{x \notin L} [F_\ell[R_{\ell-1} \oplus \mathsf{Span}(x)]] + \Pr_{x \notin R_{\ell-1}} [x \in L] \mathop{\mathbb{E}}_{x \in L \setminus R_{\ell-1}} [F_\ell[R_{\ell-1} \oplus \mathsf{Span}(x)]] \right].$$

Computing the above probabilities and using linearity of expectation one sees that the above equals

$$\begin{aligned}\mathop{\mathbb{E}}_{R_{\ell-1} \subseteq L} \left[\mu_{R_{\ell-1},\mathsf{in}}[F_\ell] \right] &= \frac{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}} E + \frac{\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}} \mathop{\mathbb{E}}_{\substack{R_{\ell-1} \subseteq L \\ x \in L \setminus R_{\ell-1}}} [F_\ell[R_{\ell-1} \oplus \mathsf{Span}(x)]] \\ &= \frac{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}} E + \frac{\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}} F_\ell[L].\end{aligned}$$

On the other hand, since $F_\ell \perp J_{\leqslant \ell-1}$,

$$\mathop{\mathbb{E}}_{R_{\ell-1} \subseteq L} \left[\mu_{R_{\ell-1},\mathsf{in}}[F] \right] = 0.$$

Combining the last two equations, we get that

$$\frac{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}} E + \frac{\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}} F[L] = 0,$$

which by rearranging implies that

$$E = -\frac{\begin{bmatrix}\ell\\1\end{bmatrix} - \begin{bmatrix}\ell-1\\1\end{bmatrix}}{\begin{bmatrix}k\\1\end{bmatrix} - \begin{bmatrix}\ell\\1\end{bmatrix}} F[L].$$

Note that the coefficient of $F[L]$ in the previous equation is precisely

$$-\frac{2^{\ell-1}}{2^k - 2^\ell} = -\frac{1}{2^{k-\ell+1} - 2} = \lambda_\ell(k,\ell),$$

and we are done. □

We are now ready to prove Theorem 2.16:

Theorem 2.16 (Restated) . *Let k, ℓ and $0 \leqslant i \leqslant \ell$ be integers. If $k \geqslant 7\ell^2 + 1$ [6], then $J_{=i}$ is an eigenspace of $A_{G(V,\ell)}$ with eigenvalue $\lambda_i(k,\ell)$ and dimension $\begin{bmatrix}k\\i\end{bmatrix} - \begin{bmatrix}k\\i-1\end{bmatrix}$.*

Proof of Theorem 2.16. By Lemmas 2.20, 2.21 we have that $J_{=i}$ are eigenspaces, and their direct sum is the entire space of functions over $G(V,\ell)$. Let A_i be the eigenspaces of λ_i. By [BCN12, Theorem 9.3.3] we have that $\mathsf{dim}(A_i) = \begin{bmatrix}k\\i\end{bmatrix} - \begin{bmatrix}k\\i-1\end{bmatrix}$. Since by Fact 2.15 all λ_i are different, we conclude that $J_{=i} \subseteq A_i$. If we have equalities for all i, we are done. Otherwise, we have strict containment for at least one i and so

$$\sum_{i=0}^{\ell} \mathsf{dim}(J_{=i}) < \sum_{i=0}^{\ell} \mathsf{dim}(A_i) = \begin{bmatrix}k\\\ell\end{bmatrix}.$$

On the other hand, since all spaces $J_{=i}$ are mutually orthogonal, the left hand side equals $\mathsf{dim}(J_{=0} \oplus ... \oplus J_{=\ell}) = \mathsf{dim}(J_{\leqslant\ell}) = \begin{bmatrix}k\\\ell\end{bmatrix}$, the last equality holds since the space of all functions on $G(V,\ell)$, $J_{\leqslant\ell}$, is of dimension $\begin{bmatrix}k\\\ell\end{bmatrix}$. Contradiction. □

2.4 Explicit projections

In this section we derive some expressions related to the projections of a function $F\colon \begin{bmatrix}V\\\ell\end{bmatrix} \to \mathbb{R}$ to $J_{=0} \oplus J_{=1} \oplus ... \oplus J_{=\ell}$, as in Definition 2.5. We give exact expressions for $F_{=0}, F_{=1}$, and approximated versions $F_{\approx i}$ for all i. While this may already be known, we include the proofs as we are unaware of a published source.

2.4.1 Exact projections

Let $F\colon \begin{bmatrix}V\\\ell\end{bmatrix} \to \mathbb{R}$, and denote $\mu(F) \stackrel{def}{=} \mathbb{E}_{L\in\begin{bmatrix}V\\\ell\end{bmatrix}}[F[L]]$. Let us find $F_{=0}$. Clearly, F_0 is a constant function such that $F - F_{=0}$ is perpendicular to $J_{=0}$. Therefore, we need $\mathbb{E}_{L\in\begin{bmatrix}V\\\ell\end{bmatrix}}[F[L] - F_{=0}[L]] = 0$ and thus we get that $F_{=0} \equiv \mu(F)$.

More generally, assuming we have computed $F_{=0}, ..., F_{=i-1}$, $F_{=i}$ is the only function from $J_{\leqslant i}$ such that $F - F_{=i} - F_{=i-1} - .... - F_{=0}$ is perpendicular to $J_{\leqslant i}$. I.e., for every $L_i \in \begin{bmatrix}V\\i\end{bmatrix}$,

$$\mathop{\mathbb{E}}_{L\supseteq L_i}[(F - F_{=i} - F_{=i-1} - .... - F_{=0})[L]] = 0.$$

Thus, in theory, deriving a formula for $F_{=i}$ is possible, albeit involves unpleasant computations. We shall demonstrate it in the simplest case $i = 1$.

[6] This condition is an artifact of the proof presented herein, and can probably be relaxed.

Level 1 projection

Define $f_1(x) = \frac{2^k-2}{2^k-2^\ell}\left(\mu_{\mathsf{Span}(x),\mathsf{in}}(F) - \mu(F)\right)$.

Lemma 2.22. *Let $k > \ell \geqslant 1$ be integers, and $F\colon \binom{V}{\ell} \to \mathbb{R}$. Then*

$$F_{=1}[L] = \sum_{x \in L \setminus \{0\}} f_1(x),$$

in other words $f_{=1} = f_1$.

Proof. It suffices to show that $G[L] \stackrel{def}{=} F[L] - \sum_{x \in L\setminus\{0\}} f_1(x) - F_0[L]$ is perpendicular to $J_{\leqslant 1}$. By Claim 2.17, it suffices to show that for every $x \neq 0$,

$$\mathop{\mathbb{E}}_{L \ni x}[G[L]] = 0.$$

Using linearity of expectation,

$$\begin{aligned}
\mathop{\mathbb{E}}_{L\ni x}[G[L]] &= \mathop{\mathbb{E}}_{L\ni x}\left[F[L] - \sum_{y\in L\setminus\{0\}} f_1(y) - F_0[L]\right] \\
&= \mathop{\mathbb{E}}_{L\ni x}[F[L]] - \mathop{\mathbb{E}}_{L\ni x}\left[\sum_{y\in L\setminus\{0\}} f_1(y)\right] - \mu(F) \\
&= \mu_{\mathsf{Span}(x),\mathsf{in}}(F) - \mu(F) - f_1(x) - \mathop{\mathbb{E}}_{L\ni x}\left[\sum_{y\in L\setminus\{0,x\}} f_1(y)\right] \\
&= \mu_{\mathsf{Span}(x),\mathsf{in}}(F) - \mu(F) - f_1(x) - (2^\ell - 2)\mathop{\mathbb{E}}_{L\ni x}\left[\mathop{\mathbb{E}}_{y\in L\setminus\{0,x\}}[f_1(y)]\right].
\end{aligned}$$

Let us evaluate the last expectation. Note that y is distributed uniformly from $V \setminus \{0, x\}$. It holds that

$$0 = \mathop{\mathbb{E}}_{y\in V\setminus\{0\}}[f_1(y)] = \frac{2^k-2}{2^k-1}\mathop{\mathbb{E}}_{y\in V\setminus\{0,x\}}[f_1(y)] + \frac{1}{2^k-1}f_1(x),$$

and therefore $\mathbb{E}_{y\in V\setminus\{0,x\}}[f_1(y)] = -\frac{1}{2^k-2}f_1(x)$. Plugging it into the previous yields

$$\mathop{\mathbb{E}}_{L\ni x}[G[L]] = \mu_{\mathsf{Span}(x),\mathsf{in}}(F) - \mu(F) - f_1(x) + \frac{2^\ell-2}{2^k-2}f_1(x).$$

Plugging in the definition of f_1 and simplifying shows that the above expectation equals 0, as desired. □

Formulas for higher levels

The above described method can be used to obtain explicit formula for higher levels as well, however as explained, it requires tiresome calculations. Instead, we choose to work with approximated forms of the higher level components, presented in the next section.

2.4.2 Approximate Projections

Let $F\colon \binom{V}{\ell} \to \mathbb{R}$ be a function. After computing the projection of F to the eigenspaces $J_{=0}$ and $J_{=1}$, we would like to continue with the projections to $J_{=i}$ for larger i's. But the exact projections for higher levels turn out to be difficult to work with, and thus we turn to computing approximate projections instead.

So as to get an initial intuition, assume we already subtracted from F its projections on $J_{=j}$ for $j < i$, namely we are left with $F_{\geqslant i} = F_{=i} + F_{=i+1} + \ldots$. Suppose we would now like to find, or at least approximate, $F_{=i}$. A natural approach would be to start by averaging $F_{\geqslant i}$ down on i-dimensional subspaces, namely taking $f(L_i) = \mathbb{E}_{L \supseteq L_i}[F_{\geqslant i}[L]]$. By Lemma 2.19, the components of $F_{\geqslant i}$ of higher levels than i are zeroed out by this averaging, hence we are left only with $\mathbb{E}_{L \supseteq L_i}[F_{=i}[L]]$. One might have hoped that this function equals $f_{=i}(L_i)$; this is false, nonetheless we prove that they are close.

There are sources of errors other than the above: we defined $F_{\geqslant i}$ above using the exact versions of $F_{=j}$, which we do not have access to. Instead we have access to approximations $f_{\approx j}$ of $f_{=j}$ and instead of working with $F_{\geqslant i}$ we work with $F - F_{\approx 0} - \ldots - F_{\approx i-1}$ where $F_{\approx j}[L] = \sum_{L_j \subseteq L} f_{\approx j}(L_j)$ are the approximated projections previously calculated. In turn, approximating $\mathbb{E}_{L \supseteq L_i}[(F - F_{\approx i-1} - \ldots - F_{\approx 0})[L]]$ boils down to the fact that

$$\mathop{\mathbb{E}}_{L_j \subseteq L_i}[F_{\approx j}[L_j]] \approx \sum_{L_j \subseteq L_i} f_{\approx j}(L_j).$$

Definition 2.23. *Let $F\colon \binom{V}{\ell} \to \mathbb{R}$ be a function. Define $f_{\approx 0}\colon \binom{V}{0} \to \mathbb{R}$ by $f_{\approx 0}(\emptyset) \stackrel{def}{=} \mu$ and $f_{\approx 1}\colon \binom{V}{1} \to \mathbb{R}$ by*

$$f_{\approx 1}(L_1) \stackrel{def}{=} \mu_{L_1,\mathsf{in}}(F) - \mu(F).$$

Inductively, once $f_{\approx i}\colon \binom{V}{i} \to \mathbb{R}$ has been defined, define $f_{\approx i+1}\colon \binom{V}{i+1} \to \mathbb{R}$ by

$$f_{\approx i+1}(L_{i+1}) \stackrel{def}{=} \mu_{L_{i+1},\mathsf{in}}(F) - \sum_{j=0}^{i} \sum_{\substack{L_j \subseteq L_{i+1} \\ \dim(L_j)=j}} f_{\approx j}(L_j).$$

Comparing $f_{\approx 1}$ and $f_{=1}$, one can already observe that while the two functions are different, they are very close to each other. Hence $F_{\approx 1}$ given below is a good approximation for F_1.

Definition 2.24. *Let $F\colon \binom{V}{\ell} \to \mathbb{R}$ be a function. Define*

$$F_{\approx 0}[L] = \mu(F).$$

For $0 < i \leqslant \ell - 1$, define

$$F_{\approx i}[L] \stackrel{def}{=} \sum_{\substack{L_i \subseteq L \\ \dim(L_i)=i}} f_{\approx i}(L_i),$$

and for $i = \ell$, define

$$F_{\approx \ell}[L] = F[L] - \sum_{i=0}^{\ell-1} F_{\approx i}[L].$$

The following theorem, which is the main purpose of this section, asserts that the $F_{\approx i}$'s are close to the $F_{=i}$'s in ℓ_2-norm, given that the function F is bounded.

Theorem 2.25. *Assume $k \geqslant 13\ell^2 + 1$, let V be a k-dimensional vector space over $\mathbb{F}_2$. Then for every $F\colon \binom{V}{\ell} \to \mathbb{R}$, $0 \leqslant i \leqslant \ell$*

$$\|F_{=i} - F_{\approx i}\|_2^2 \leqslant 2^{26\ell^4 - k}\|F\|_\infty^2.$$

Roughly speaking, the way this theorem is proved is by showing that $F - \tilde{F}_i - .... - \tilde{F}_0$ is nearly perpendicular to $J_{\leqslant i}$ for all i. Morally speaking, this implies the projection of F onto $J_{\leqslant i}$ is nearly $F_{\approx 0} + ... + F_{\approx i}$. Given this holds for every i, one can prove by induction that $F_{\approx i}$ is close to $F_{=i}$.

The actual proof is more involved, and is deferred to Section A. We remark that we did not attempt to optimize the upper bound in terms of the dependence on ℓ; in fact it is clear they are not tight:

Claim 2.26. *For all $0 < \ell < k$, $F\colon \binom{V}{\ell} \to \mathbb{R}$ it holds that*

$$\|f_{=1} - f_{\approx 1}\|_\infty \leqslant 2^{\ell + 2 - k}\|F\|_\infty,$$

2.23. In particular,

$$\|F_{=1} - F_{\approx 1}\|_\infty \leqslant 2^{2\ell + 2 - k}\|F\|_\infty.$$

Proof. The first is straightforward by definitions, and the second follows from the first by the triangle inequality. □

3 Expansion on the Grassmann Graph

In this section we describe our main results.

We begin with a function-version definition of pseudo-randomness.

Definition 3.1. *Let V be a vector space of dimension k over $\mathbb{F}_2$, $\ell \ll k$. Let m be an integer and $\varepsilon > 0$. We say a function $F\colon \binom{V}{\ell} \to \mathbb{R}$ is (m, ε) pseudo-random if:*

- *For every integers q, r such that $q + r = m$ and q-dimensional subspace Q and W of co-dimension r, $\left|\mu_{(Q,\mathsf{in}),(W,\mathsf{out})}(F) - \mu(F)\right| \leqslant \varepsilon$.*

Identifying sets and their indicator functions, we see that the above definition is a generalization of Definition 1.5, with the exception that we require a pseudo-random function to not deviate from its mean to any direction. In Definition 1.5 we only required that a set is not significantly denser on zooms for it to be pseudo-random. Note that whenever $\varepsilon \geqslant \mu(F)$ the definitions are equivalent for non-negative functions, and in particular for indicators of sets.

We have already seen several examples of non pseudo-random functions, namely the indicator function of $\{L \mid x \in L\}$ for a fixed $x \in V \setminus \{0\}$ and $\{L \mid L \subseteq W\}$ for a hyperplane $W \subseteq V$. The main goal of this paper is to show that the indicator function of a non-optimally expanding set cannot be pseudo-random. As discussed in the introduction, we approach this question via the spectral decomposition of the indicator function.

3.1 Spectral approach for expansion

Recall that by inequality (1), if F is the indicator of a set with expansion bounded by $1 - \delta - \varepsilon$,

$$\sum_{i=1}^{\ell} \lambda_i \langle F_{=i}, F_{=i} \rangle \geqslant \varepsilon\delta.$$

We will now use the fact that the eigenvalues of $G(V,\ell)$ are exponentially decaying to conclude there is a low-level i on which F has large mass. First, we define the weight of a function on a single level.

Definition 3.2. *The weight of F on level i is denoted by $W^{=i}[F]$, and defined by*

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

By orthogonality, the weight of F on level i equals $\langle F, F_{=i} \rangle$.

To make the above discussion quantitative, let us estimate the tail of the sum on the left hand side. By Fact 2.12 $\lambda_i \leqslant 2^{-i}$, and so

$$\sum_{i=\lceil \log \frac{2}{\varepsilon} \rceil}^{\ell} \lambda_i \langle F_{=i}, F_{=i} \rangle \leqslant 2^{-\log \frac{2}{\varepsilon}} \sum_{i=\lceil \log \frac{2}{\varepsilon} \rceil}^{\ell} \langle F_{=i}, F_{=i} \rangle \leqslant \frac{1}{2}\varepsilon \sum_{i=0}^{\ell} \langle F_{=i}, F_{=i} \rangle = \frac{1}{2}\varepsilon \cdot \|F\|_2^2 \leqslant \frac{1}{2}\varepsilon\delta.$$

We used Parseval in the last equality. Therefore we conclude that

$$\sum_{i=1}^{\lceil \log \frac{2}{\varepsilon} \rceil - 1} \lambda_i \langle F_{=i}, F_{=i} \rangle \geqslant \varepsilon\delta - \sum_{i=1}^{\lceil \log \frac{2}{\varepsilon} \rceil} \lambda_i \langle F_{=i}, F_{=i} \rangle \geqslant \frac{1}{2}\varepsilon\delta,$$

and in particular there is $i^* \in \left\{1, 2, ..., \lceil \log \frac{2}{\varepsilon} \rceil - 1\right\}$ such that the weight of F on the i^* level is large:

$$\langle F_{=i^*}, F_{=i^*} \rangle \geqslant \frac{1}{\lambda_{i^*}} \frac{\varepsilon}{2\lceil \log \frac{2}{\varepsilon} \rceil}\delta.$$

It follows that in order to characterize sets with non-perfect expansion, it suffices to prove a good enough structural result for Boolean functions with large weight on their low-level spectrum. In this paper we make some progress in this direction, giving some structural results for functions with large weight on their first or second levels. We hope our techniques might be relevant to the general case.

We begin with our results for the first level. The first one asserts that a function that has at least $\Omega(\delta^{4/3})$ weight on its first level is not pseudo-random:

Theorem 3.3. *There exists $c_0 > 0$, such that the following holds. Let V be a k-dimensional vector space, $10 < \ell < k$ be an integer, and $\delta, \eta > 0$ such that $k \geqslant 6\ell + 2\log \frac{1}{\eta} + c_0$. Let $F \colon \begin{bmatrix} V \\ \ell \end{bmatrix} \to \{0,1\}$, and assume $W^{=1}[F] = \eta$, $\mu(F) = \delta$.*

If $\eta \geqslant 30\delta^{4/3}$, then F is not $(1, \frac{\eta^2}{200\delta^2})$ pseudo-random.

The above theorem asserts that any Boolean-valued function that has significant weight on its first level must have noticeable correlation with the indicator function of $\{ L \mid x \in L \}$ or $\{ L \mid L \subseteq W \}$ for some $x \in V$ or $W \subseteq V$ hyperplane. The above theorem is proved in Section 5.

It will be useful for us to phrase the above theorem in a counter-positive form, which reads:

Corollary 3.4. *There exists $c_0 > 0$, such that the following holds. Let V be a k-dimensional vector space, $10 < \ell < k$ an integer and $\varepsilon, \delta > 0$ such that $k \geqslant 6\ell + 3\log \frac{1}{\delta} + c_0$. Let $F \colon \begin{bmatrix} V \\ \ell \end{bmatrix} \to \{0,1\}$, and assume $\mu(F) = \delta$, $\varepsilon \geqslant \frac{9}{2}\delta^{2/3}$.*

If F is $(1, \varepsilon)$ pseudo-random, then $W^{=1}[F] \leqslant \sqrt{200\varepsilon}\delta$.

Our next main result is concerned with a quantitative improvement of Theorem 3.3 with regards to the weight needed on level 1 so as to deduce non-pseudorandomness.

Theorem 3.5. *There exists $\ell_0 > 0$, such that for all $\ell > \ell_0$ the following holds. Let V be a k-dimensional vector space, and $\varepsilon, \delta, \eta > 0$ such that $k \geqslant 30\ell^4 + \frac{100}{\varepsilon^3}$, $\sqrt{\frac{40}{\log(1/\delta)}} < \varepsilon < 1$. Let $F\colon \binom{V}{\ell} \to \{0,1\}$, and assume $W^{=1}[F] = \eta$, $\mu(F) = \delta$.*
If $\eta \geqslant \delta^{2-\varepsilon}$, then F is not $(1, 2^{-\frac{20}{\varepsilon^2}} \frac{\eta^{1+\frac{\varepsilon}{4}}}{\delta})$ pseudo-random.

Counter positively, a small pseudo-random set cannot have significantly higher weight than δ^2 on its first level, where δ is the density of the set. The above theorem is proved in Section 6

We complement this last result with a matching example, showing that δ^2 weight on the first level is necessary to conclude non pseudo-randomness. More precisely

Theorem 3.6. *There exists constants $c_1, c_2, c_3 > 0$ such that the following holds. Let $\delta > 0$ and $k \gg \ell$ sufficiently large, and let V be a k-dimensional subspace over $\mathbb{F}_2$. Then there exists a set of vertices S in $G(V, \ell)$ such that*

- *S has density $\Theta(\delta)$: $\mu(S) = c_1\delta$.*
- *S has weight $\Omega(\delta^2)$ on the first level: $W^{=1}[S] \geqslant c_2\delta^2$.*
- *S is $(1, c_3 2^{-\ell/2})$ pseudo-random.*

The above theorem is proved in Section C.

Our final result is analogous to theorem 3.3 on the second level. It asserts that a pseudo-random function cannot have large weight on the second level. More precisely:

Theorem 3.7. *There exists $\ell_0 > 0$, such that for all integer $\ell \geqslant \ell_0$ the following holds. Let k be an integer, V be a k-dimensional vector space, an integer and $\varepsilon, \delta > 0$ such that $k \geqslant 27\ell^4 + 10\log\frac{1}{\delta}$.*
If $\eta > 2^{17} \cdot \delta^{4/3}$, then F is not $(2, 2^{-52}\left(\frac{\eta}{\delta}\right)^3)$ pseudo-random.

The above theorem is proved in Section 7 (note that we have made no significant effort optimizing the various constants and dependencies between k, ℓ, δ in our main results).

4 Analysis on The Boolean Hypercube

In this section we show each $f_{\approx i}$ can be viewed as a function over the Boolean hypercube, and and develop a connection between its hypercube fourier transform and zoom outs. This connection exhibits strong duality between zoom ins, that in a sense define the values of $f_{\approx i}$, and zoom outs, that in a sense define its fourier coefficients $\widehat{f_{\approx i}}$.

Let V be a vector space over $\mathbb{F}_2$ of dimension k. We can naturally identify V with $\mathbb{F}_2^k$: let $e_1, ..., e_k$ be a basis of V, and consider the transformation

$$x = \sum_{i=1}^{k} x_i e_i \longrightarrow (x_1, x_2, ..., x_k).$$

Thus we may think of x as having coordinates $(x_1, ..., x_k)$. This view gives rise to the decomposition of each $f\colon V \to \mathbb{R}$ according multiplicative functions:

$$f(x) = \sum_{S \subseteq [k]} \widehat{f}(S)\chi_S(x),$$

where $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$, and

$$\widehat{f}(S) = \mathop{\mathbb{E}}_{x \in V}[f(x)\chi_S(x)].$$

It turns out that when f is a level function of F, there is a tight relation between its fourier coefficients and the notion of zoom-outs from Definition 2.8.

Definition 4.1. *The associated hyperplane W_S of $S \neq \emptyset$ is defined by*

$$W_S \stackrel{def}{=} \{\, x \mid \chi_S(x) = 1 \}\,.$$

Lemma 4.2. *Let $F\colon \binom{V}{\ell} \to \mathbb{R}$, and consider $f_{\approx 1}\colon V \to \mathbb{R}$ (We extend $\tilde{f}_1(0) = 0$). Then*

$$\widehat{f}_{\approx 1}(S) = c_1(k,\ell)(\mu_{W_S,\mathsf{out}}(F) - \mu(F)),$$

where $c_1(k,\ell) = \frac{\begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix} - \begin{bmatrix} k-1 \\ \ell \end{bmatrix}} \leqslant \frac{1}{\begin{bmatrix} \ell \\ 1 \end{bmatrix}} = 2^{-\ell} + O(2^{-2\ell})$.

Proof. By definition,

$$\widehat{f}_{\approx 1}(S) = \mathop{\mathbb{E}}_{x \in V}[f_{\approx 1}(x)\chi_S(x)] = \frac{1}{2}\mathop{\mathbb{E}}_{x \in W_S}[f_{\approx 1}(x)] - \frac{1}{2}\mathop{\mathbb{E}}_{x \notin W_S}[f_{\approx 1}(x)].$$

Note that the above two expectations sum up to $\mathbb{E}_{x \in V}[f_{\approx 1}(x)] = 0$ and hence are equal in absolute values but negated in signs, and it holds that

$$\widehat{f}_{\approx 1}(S) = -\mathop{\mathbb{E}}_{x \notin W_S}[f_{\approx 1}(x)].$$

Let us expand out the previous.

$$\widehat{f}_{\approx 1}(S) = -\mathop{\mathbb{E}}_{x \notin W_S}[f_{\approx 1}(x)] = -\mathop{\mathbb{E}}_{x \notin W_S}\left[\mu_{\mathsf{Span}(x),\mathsf{in}} - \delta\right] = -\mathop{\mathbb{E}}_{x \notin W_S}\left[\mathop{\mathbb{E}}_{L \ni x}[F[L] - \delta]\right] = -\mathop{\mathbb{E}}_{L \not\subseteq W_S}[F[L] - \delta]. \tag{3}$$

The first equality is by definition of $f_{\approx 1}$, the second by Definition 2.6, and the third follows since each L that is not contained in W_S, contains precisely $2^{\ell-1}$ points outside it.

Using conditional expectations,

$$\begin{aligned} 0 = \mathop{\mathbb{E}}_{L}[F[L] - \delta] &= \Pr_{L}[L \subseteq W_S]\mathop{\mathbb{E}}_{L \subseteq W_S}[F[L] - \delta] + \Pr_{L}[L \not\subseteq W_S]\mathop{\mathbb{E}}_{L \not\subseteq W_S}[F[L] - \delta] \\ &= \frac{\begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix}}(\mu_{W_S,\mathsf{out}} - \delta) + \frac{\begin{bmatrix} k \\ \ell \end{bmatrix} - \begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix}}\mathop{\mathbb{E}}_{L \not\subseteq W_S}[F[L] - \delta]. \end{aligned}$$

Rearranging yields

$$-\mathop{\mathbb{E}}_{L \not\subseteq W_S}[F[L] - \delta] = \frac{\begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix} - \begin{bmatrix} k-1 \\ \ell \end{bmatrix}}(\mu_{W_S,\mathsf{out}} - \delta)$$

Plugging the above equation into Equation 3 finishes the proof. □

Asymptotically, the coefficient in the last Lemma can be shown to be $2^{-\ell} + O(2^{-2\ell})$, which is the main point here.

Fourier analysis for higher levels. One can derive similar expressions for the fourier coefficients of $\tilde{f}_r$, for $r > 1$. Indeed we develop the case $r = 2$ in Section 7.1 for the analysis of the second level.

5 Results for The First Level

In this section we prove Theorem 3.3, slightly restated as follows:

Theorem 3.3 (Restated) . *There exists $c_0 > 0$, such that the following holds. Let V be a k-dimensional vector space, $10 < \ell < k$ be an integer, and $\delta, \eta > 0$ such that $k \geqslant 6\ell + 2\log\frac{1}{\eta} + c_0$. Let $F\colon \binom{V}{\ell} \to \{0,1\}$, and assume $W^{=1}[F] = \eta$, $\mu(F) = \delta$.*

If $\eta \geqslant 30\delta^{4/3}$, then at least one of the two must happen:

1. $$\Pr_x\left[\mu_{\mathsf{Span}(x),\mathsf{in}}(F) \geqslant \delta + \frac{\eta^2}{100\delta^2}\right] \geqslant c(\eta,\delta)2^{-\ell},$$
   where $c(\eta,\delta) > 0$ depends only on η, δ.
2. *There exists W of co-dimension 1, such that $\mu_{W,\mathsf{out}}(F) \geqslant \delta + \frac{\eta^2}{100\delta^2}$.*

Throughout this section, whenever F is clear from the context, we also denote $\delta = \mu(F)$, $\eta = W^{=1}[F]$.

First, we establish a relation between the second moment of $F_{=1}$ and the second moment of $f_{=1}$. As we need this relation later for higher levels than 1, and since this relation is implied for higher levels with the same proof, we give it in full generality.

Lemma 5.1. *Let $k > \ell$ and $1 \leqslant r \leqslant \ell$ be integers, and $\varepsilon \geqslant 0$, and let $G_r[L] = \sum_{L_r \subseteq L} g_r(L_r)$ where $g_r\colon \binom{V}{r} \to \mathbb{R}$.*

Suppose that for all $L_{r-1} \in \binom{V}{r-1}$, $\left|\mathbb{E}_{L_r \supseteq L_{r-1}}\left[g_r(L_r)\right]\right| \leqslant \varepsilon$. Then

$$\|G_r\|_2^2 = \begin{bmatrix}\ell\\ r\end{bmatrix}\mathop{\mathbb{E}}_{L_r}\left[g_r^2(L_r)\right] + O(2^{4r\ell}(\varepsilon + 2^{-k}\|g_r\|_\infty^2)).$$

Proof.

$$\mathop{\mathbb{E}}_{L}\left[G_r^2[L]\right] = \mathop{\mathbb{E}}_{L}\left[\left(\sum_{L_r\subseteq L} g_r(L_r)\right)^2\right] = \mathop{\mathbb{E}}_{L}\left[\sum_{L_r, L_r' \subseteq L} g_r(L_r)g_r(L_r')\right].$$

We divide the last sum according to the dimension of $L_r \cap L_r'$, which is between 0 and r:

$$= \mathop{\mathbb{E}}_{L}\left[\sum_{i=0}^{r}\begin{bmatrix}\ell\\ i\end{bmatrix}\begin{bmatrix}\ell-i\\ r-i\end{bmatrix}\begin{bmatrix}\ell-r\\ r-i\end{bmatrix}\mathop{\mathbb{E}}_{\substack{L_r, L_r' \subseteq L\\ \dim(L_r\cap L_r')=i}}\left[g_r(L_r)g_r(L_r')\right]\right],$$

the first term is the choice of $L_r \cap L_r'$, and the other two correspond to the choice of the rest of the vectors in L_r such that the intersection remains of dimension i. Interchange the sum and the outer expectation, and note that the distribution over L_r, L_r' is uniform over those that intersect on dimension i to get

$$= \sum_{i=0}^{r} \begin{bmatrix} \ell \\ i \end{bmatrix} \begin{bmatrix} \ell - i \\ r - i \end{bmatrix} \begin{bmatrix} \ell - r \\ r - i \end{bmatrix} \mathop{\mathbb{E}}_{\substack{L_r, L'_r \in \begin{bmatrix} V \\ r \end{bmatrix} \\ \dim(L_r \cap L'_r) = i}} \left[g_r(L_r) g_r(L'_r) \right].$$

Divide the last sum into $i = r$ - for which we get $\begin{bmatrix} \ell \\ r \end{bmatrix} \mathbb{E}_{L_r} \left[g_r^2(L_r) \right]$, and the rest. Notice that for a fixed i, the distribution over L_r, L'_r is $O(2^{(r-i)r-k})$ close to picking $A \in \begin{bmatrix} V \\ i \end{bmatrix}$ and then picking $L_r, L'_r \supseteq A$ independently. Therefore,

$$\begin{aligned} \mathop{\mathbb{E}}_{L} \left[G_r^2[L] \right] &= \begin{bmatrix} \ell \\ r \end{bmatrix} \mathop{\mathbb{E}}_{L_r} \left[g_r^2(L_r) \right] + \sum_{i=0}^{r-1} \begin{bmatrix} \ell \\ i \end{bmatrix} \begin{bmatrix} \ell - i \\ r - i \end{bmatrix} \begin{bmatrix} \ell - r \\ r - i \end{bmatrix} \left(\mathop{\mathbb{E}}_{\substack{A \in \begin{bmatrix} V \\ i \end{bmatrix} \\ L_r, L'_r \supseteq A}} \left[g_r(L_r) g_r(L'_r) \right] + O(2^{(r-i)r-k} \|g_r\|_\infty^2) \right) \\ &= \begin{bmatrix} \ell \\ r \end{bmatrix} \mathop{\mathbb{E}}_{L_r} \left[g_r^2(L_r) \right] + \sum_{i=0}^{r-1} \begin{bmatrix} \ell \\ i \end{bmatrix} \begin{bmatrix} \ell - i \\ r - i \end{bmatrix} \begin{bmatrix} \ell - r \\ r - i \end{bmatrix} \mathop{\mathbb{E}}_{A \in \begin{bmatrix} V \\ i \end{bmatrix}} \left[\left(\mathop{\mathbb{E}}_{L_r \supseteq A} \left[g_r(L_r) \right] \right)^2 \right] + O(2^{4r\ell - k} \|g_r\|_\infty^2). \end{aligned}$$

We used the fact that each one of the Gaussian coefficients above is at most $2^{r\ell}$. Note that for every A, since $\dim(A) \leqslant r - 1$, we have that $\left| \mathbb{E}_{L_r \supseteq A} \left[g_r(L_r) \right] \right| \leqslant \varepsilon$. Therefore we get that

$$\begin{aligned} \mathop{\mathbb{E}}_{L} \left[G_r^2[L] \right] &= \begin{bmatrix} \ell \\ r \end{bmatrix} \mathop{\mathbb{E}}_{L_r} \left[g_r^2(L_r) \right] + \sum_{i=0}^{r-1} \begin{bmatrix} \ell \\ i \end{bmatrix} \begin{bmatrix} \ell - i \\ r - i \end{bmatrix} \begin{bmatrix} \ell - r \\ r - i \end{bmatrix} \mathop{\mathbb{E}}_{A \in \begin{bmatrix} V \\ i \end{bmatrix}} \left[O(\varepsilon^2) \right] + O(2^{4r\ell - k} \|g_r\|_\infty^2) \\ &= \begin{bmatrix} \ell \\ r \end{bmatrix} \mathop{\mathbb{E}}_{L_r} \left[g_r^2(L_r) \right] + O(r \cdot 2^{3r\ell} \varepsilon^2) + O(2^{4r\ell - k} \|g_r\|_\infty^2). \end{aligned}$$

□

We remark that in this section, the above lemma will always be used with $\varepsilon = 0$.

5.1 Deviation properties of $F_{=1}[L]$

We capture the idea that $F_{=1}$ has to oscillate between two far values using its various moments. The observations below are not specific for $F_{=1}$.

Claim 5.2.

$$\mathop{\mathbb{E}}_{L} \left[F_{=1}[L]^2 \right] = W^{=1}[F] = \eta.$$

Proof. By definition. □

Lemma 5.3.

$$\Pr_{L} \left[F_{=1}[L] \geqslant \frac{\eta}{4\delta} \right] \geqslant \frac{\eta}{2}$$

Proof. Observe that $\mathbb{E}_L \left[(F[L] - F_{=1}[L])^2 \right]$ is the weight of F outside the first level. Since by Parseval the weight of F sum up to $\|F\|_2^2 = \delta$ and the weights of F on the first level is η, we conclude that

$$\mathop{\mathbb{E}}_{L} \left[(F[L] - F_{=1}[L])^2 \right] = \delta - \eta.$$

Hence, by Markov's inequality,

$$\Pr_L\left[(F[L]-F_{=1}[L])^2 > 1-\frac{\eta}{2\delta}\right] \leqslant \frac{\delta-\eta}{1-\frac{\eta}{2\delta}} \leqslant \delta\left(1-\frac{\eta}{2\delta}\right) = \delta-\frac{1}{2}\eta.$$

Note that $F[L]=1$ with probability δ, and therefore $(1-F_{=1}[L])^2 \leqslant 1-\frac{\eta}{2\delta}$ with probability at least $\frac{1}{2}\eta$. Thus with probability at least $\frac{1}{2}\eta$,

$$|1-F_{=1}[L]| \leqslant \sqrt{1-\frac{\eta}{2\delta}} \leqslant 1-\frac{\eta}{4\delta},$$

(in the last inequality we used $\sqrt{1-\varepsilon} \leqslant 1-\frac{1}{2}\varepsilon$), implying that $\frac{\eta}{4\delta} \leqslant F_{=1}[L] \leqslant 2-\frac{\eta}{4\delta}$ with probability at least $\frac{1}{2}\eta$. □

Claim 5.4.

$$\mathop{\mathbb{E}}_L\left[F_{=1}[L]^4\right] \geqslant \frac{\eta^5}{512\delta^4}.$$

Proof. Using Lemma 5.3

$$\mathop{\mathbb{E}}_L\left[F_{=1}[L]^4\right] \geqslant \Pr_L\left[F_{=1}[L] \geqslant \frac{\eta}{4\delta}\right]\left(\frac{\eta}{4\delta}\right)^4 \geqslant \frac{\eta^5}{512\delta^4}.$$

□

Claim 5.5.

$$\mathbb{E}_x\left[f_{=1}^2(x)\right] = \frac{W^{=1}[F]}{\binom{\ell}{1}} + O(2^{4\ell-k}) = \frac{\eta}{\binom{\ell}{1}} + O(2^{4\ell-k}).$$

Proof. By Lemma 5.1 with $\varepsilon=0$ (since $F_{=1} \in J_{=1}$, Claim 2.19 and the fact that $\|f_{=1}\|_\infty = O(1)$ for a Boolean F, which is clear by the exact formula from Section 2.4.1), we get that

$$\mathbb{E}_L\left[F_{=1}[L]^2\right] = \binom{\ell}{1}\mathop{\mathbb{E}}_x\left[f_{=1}^2(x)\right] + O(2^{4\ell-k}).$$

The left hand side is $W^{=1}[F]=\eta$ by Claim 5.2. Rearranging finishes the proof. □

5.2 Proof of Theorem 3.3 – the main argument

The next step of the proof is to obtain information about $f_{=1}$ from the information we know about the fourth moment of $F_{=1}$. We write $F_{=1}$ as a sum of values of $f_{=1}$, and hence its fourth power can be written as a sum of products of values of $f_{=1}$ at four points. The heart of the proof of Theorem 3.3 goes by partitioning these products according to the linear dependencies of the participating points.

Lemma 5.6. *There is $c_0>0$, such that the following holds.*

Let $F\colon \binom{V}{\ell} \to \{0,1\}$ and denote $\eta = W^{=1}[F], \delta=\mu(F)$. Assume $\eta \geqslant 15\delta^{4/3}$ and $k \geqslant 6\ell + 2\log\frac{1}{\eta} + c_0$. Then at least one of the two holds:

1.

$$\mathbb{E}_x\left[f_{=1}^4(x)\right] \geqslant \frac{1}{2048\binom{\ell}{1}}\frac{\eta^5}{\delta^4}.$$

2.

$$\mathop{\mathbb{E}}_{x,y,z\in V}[f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] \geqslant \frac{1}{2048\begin{bmatrix}\ell\\1\end{bmatrix}^3}\frac{\eta^5}{\delta^4}.$$

Proof. By Claim 5.4, $\mathbb{E}_L[F_{=1}[L]^4]$ is at least $\frac{\eta^5}{512\delta^4}$.
Let us now, using the definition of $F_{=1}$, open the parenthesis in the computation of $\mathbb{E}_L[F_{=1}[L]^4]$. We partition the resulting summands according to their linear dependencies: in the sums below we only sum over vectors that are linearly independent, unless explicitly stated otherwise (e.g. we sum over x, y, z, w, u where $x \neq y$, $z \notin \mathsf{Span}(\{x, y\})$, $w \notin \mathsf{Span}(\{x, y, z\})$).

$$\begin{aligned}
\mathbb{E}_L\left[F_1[L]^4\right] &= \mathbb{E}_L\left[\left(\sum_{x\in L\setminus\{0\}} f_{=1}(x)\right)^4\right] \\
&= \mathop{\mathbb{E}}_{L}\left[\sum_{x\in L\setminus\{0\}} f_{=1}(x)^4\right] + \mathop{\mathbb{E}}_{L}\left[\sum_{x,y\in L\setminus\{0\}} f_{=1}^2(x)f_{=1}^2(y)\right] \\
&+ \mathop{\mathbb{E}}_{L}\left[\sum_{x,y\in L\setminus\{0\}} f_{=1}(x)f_{=1}(y)f_{=1}^2(x+y)\right] + \mathop{\mathbb{E}}_{L}\left[\sum_{\substack{x,y,z\in L\setminus\{0\}\\ u\in Span(\{x,y,z\})\setminus\{0,x,y,z\}}} f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(u)\right] \\
&+ \mathop{\mathbb{E}}_{L}\left[\sum_{x,y,z,w\in L\setminus\{0\}} f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(w)\right] + \mathop{\mathbb{E}}_{L}\left[\sum_{x,y\in L\setminus\{0\}} f_{=1}^3(x)f_{=1}(y)\right].
\end{aligned} \tag{4}$$

Denote the above expectations $A_1, ..., A_6$. Clearly by linearity of expectation $A_1 = \begin{bmatrix}\ell\\1\end{bmatrix}\mathbb{E}_x\left[f^4(x)\right]$. Below we bound A_2, A_3 by $O(\eta^2)$.

$$\begin{aligned}
|A_2| &= \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\mathop{\mathbb{E}}_{L}\left[\mathop{\mathbb{E}}_{x,y\in L\setminus\{0\}}\left[f_{=1}^2(x)f_{=1}^2(y)\right]\right] \leqslant \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\left(\mathop{\mathbb{E}}_{x,y\in V}\left[f_{=1}^2(x)f_{=1}^2(y)\right]+O(2^{-k})\right) \\
&= \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\left(\mathop{\mathbb{E}}_{x\in V}\left[f_{=1}^2(x)\right]^2+O(2^{-k})\right) \leqslant \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\left(\frac{\eta^2}{\begin{bmatrix}\ell\\1\end{bmatrix}^2}+O(2^{4\ell-k})\right) \\
&\leqslant \eta^2 + O(2^{6\ell-k}).
\end{aligned}$$

The second inequality is by Claim 5.5.

$$|A_3| = \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\left|\mathop{\mathbb{E}}_{x,y\in V}[(f_{=1}(x)f_{=1}(x+y))(f_{=1}(y)f_{=1}(x+y))]\right| + \begin{bmatrix}\ell\\1\end{bmatrix}^2 O(2^{-k}).$$

Apply Cauchy-Schwartz inequality on the first expectation to get

$$\begin{aligned}
|A_3| &= \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\sqrt{\mathop{\mathbb{E}}_{x,y\in V}[(f_{=1}(x)f_{=1}(x+y))^2]}\sqrt{\mathop{\mathbb{E}}_{x,y\in V}[(f_{=1}(y)f_{=1}(x+y))^2]} + \begin{bmatrix}\ell\\1\end{bmatrix}^2 O(2^{-k}) \\
&= \begin{bmatrix}\ell\\1\end{bmatrix}\left(\begin{bmatrix}\ell\\1\end{bmatrix}-1\right)\mathop{\mathbb{E}}_{x,y\in V}\left[f_{=1}^2(x)\right]^2 + \begin{bmatrix}\ell\\1\end{bmatrix}^2 O(2^{-k}),
\end{aligned}$$

which is $\leqslant \eta^2 + O(2^{6\ell-k})$ as before.

For A_4, the distribution of x, y, z is $O(2^{-k})$-close to uniform over V. Therefore

$$\begin{aligned} A_4 =& 3\begin{bmatrix}\ell\\1\end{bmatrix}^3 \mathop{\mathbb{E}}_{\substack{x,y,z\in V\\ u\in_R\{x+y,x+z,y+z\}}} [f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(u)] \\ &+ \begin{bmatrix}\ell\\1\end{bmatrix}^3 \mathop{\mathbb{E}}_{x,y,z\in V} [f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] + O\left(\begin{bmatrix}\ell\\1\end{bmatrix}^3 2^{-k}\right). \end{aligned}$$

The first term is 0 by independence: fix one of the choices for$u$, say $u = x + y$. Then z is independent of x, y, u, and hence

$$\mathop{\mathbb{E}}_{x,y,z\in V} [f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(u)] = \mathop{\mathbb{E}}_{x,y\in V} [f_{=1}(x)f_{=1}(y)f_{=1}(u)] \mathop{\mathbb{E}}_{z\in V} [f_{=1}(z)].$$

The last expression equals 0, since the average of $f_{=1}(z)$ is 0 by Lemma 2.19. Therefore,

$$A_4 = \begin{bmatrix}\ell\\1\end{bmatrix}^3 \mathop{\mathbb{E}}_{x,y,z\in V} [f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] + O\left(\begin{bmatrix}\ell\\1\end{bmatrix}^3 2^{-k}\right).$$

For A_5 note that the distribution of x, y, z, w is $O(2^{-k})$-close to uniform over V, and therefore $|A_5| = O(2^{4\ell-k})$. Similarly it is easy to see that A_6 is $O(2^{2\ell-k})$.

Plug everything into equation (4) to get

$$\begin{aligned} \begin{bmatrix}\ell\\1\end{bmatrix} \mathop{\mathbb{E}}_x \left[f_{=1}^4(x)\right] + \begin{bmatrix}\ell\\1\end{bmatrix}^3 \mathop{\mathbb{E}}_{x,y,z\in V} [f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] &\geqslant \mathbb{E}_L\left[Z[L]^4\right] - |A_2| - |A_3| - |A_5| - |A_6| \\ &\geqslant \frac{\eta^5}{512\delta^4} - 2\eta^2 - O\left(2^{6\ell-k}\right) \\ &\geqslant \frac{\eta^5}{1024\delta^4}, \end{aligned}$$

the last inequality is since $\eta \geqslant 15\delta^{4/3}$ and k is large enough in comparison to ℓ. Therefore are least one of the two is larger than $\frac{\eta^5}{2048\delta^4}$, and we are done. □

5.3 Linearity testing and zoom-outs

Lemma 5.6 in the previous section showed that a Boolean function on $G(V, \ell)$ having high level-1 mass must satisfy one of two conditions. Here we show that if the second case holds, then there exists a zoom-out increasing the average of F. More precisely:

Lemma 5.7. *Assume $\varepsilon \geqslant 4\delta^2\eta$, $k \geqslant 2\ell + \log\frac{1}{\delta} + c_0$, and*

$$\begin{bmatrix}\ell\\1\end{bmatrix}^3 \mathop{\mathbb{E}}_{x,y,z\in V} [f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] \geqslant \varepsilon.$$

Then there exists a hyperplane W such that $\mu_{W,\mathsf{out}}[F] \geqslant \delta + \sqrt{\frac{\varepsilon}{2\eta}}$.

Proof. A standard calculation shows that

$$\mathop{\mathbb{E}}_{x,y,z\in V}[f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] = \sum_S \widehat{f}^4_{=1}(S) \leqslant \max_S \widehat{f}^2_{=1}(S) \sum_S \widehat{f}^2_{=1}(S) = \frac{1.5\eta}{\binom{\ell}{1}} \max_{S\neq\phi} f^2_{=1}(S). \tag{5}$$

The last equality is by Parseval and the fact that by Claim 5.5 $\mathbb{E}_x\left[f^2_{=1}(x)\right] \leqslant \frac{1.5\eta}{\binom{\ell}{1}}$. Therefore there is $S^* \neq \emptyset$ (since $\widehat{f}_{=1}(\emptyset) = 0$) such that

$$\widehat{f}^2_{=1}(S^*) \geqslant \frac{\binom{\ell}{1}}{2\eta} \mathop{\mathbb{E}}_{x,y,z\in V}[f_{=1}(x)f_{=1}(y)f_{=1}(z)f_{=1}(x+y+z)] \geqslant \frac{\varepsilon}{1.5\eta\binom{\ell}{1}^2},$$

and so $\left|\widehat{f}_{=1}(S^*)\right| \geqslant \sqrt{\frac{\varepsilon}{1.5\eta}}\binom{\ell}{1}^{-1}$. Since by Claim 2.26, $\|f_{\approx 1} - f_{=1}\|_\infty \leqslant 2^{\ell+2-k}$, we conclude that

$$\left|\widehat{f}_{\approx 1}(S^*)\right| \geqslant \left|\widehat{f}_{=1}(S^*)\right| - \|f_{\approx 1} - f_{=1}\|_1 \geqslant \sqrt{\frac{\varepsilon}{2\eta}}\binom{\ell}{1}^{-1}$$

Apply Lemma 4.2 to conclude that

$$\frac{\binom{k-1}{\ell}}{\binom{k}{\ell} - \binom{k-1}{\ell}}\left|\mu_{W_{S^*,\mathsf{out}}} - \delta\right| \geqslant \sqrt{\frac{\varepsilon}{2\eta}}\binom{\ell}{1}^{-1}.$$

Multiplying by $\binom{\ell}{1}$ and using [7]

$$\frac{\binom{k-1}{\ell}\binom{\ell}{1}}{\binom{k}{\ell} - \binom{k-1}{\ell}} \leqslant 1,$$

we conclude that $\left|\mu_{W_{S^*,\mathsf{out}}} - \delta\right| \geqslant \sqrt{\frac{\varepsilon}{2\eta}}$. Finally, recall that by the hypothesis $\sqrt{\frac{\varepsilon}{2\eta}} > \delta$, and so it must be the case that $\left|\mu_{W_{S^*,\mathsf{out}}} - \delta\right| = \mu_{W_{S^*,\mathsf{out}}} - \delta$ (otherwise it is at most δ), and we are done. □

5.4 Putting it all together

Proof of Theorem 3.3. Let $F : \binom{V}{\ell} \to \{0,1\}$ be a function as given in Theorem 3.3, and apply Lemma 5.6 to it. If the first item holds, we have that $\mathbb{E}_x\left[f^2_{=1}(x)\right] \leqslant \frac{2\eta}{\binom{\ell}{1}}$ (by Claim 5.5), and $\mathbb{E}_x\left[f^4_{=1}(x)\right] \geqslant \frac{1}{2048\binom{\ell}{1}}\frac{\eta^5}{\delta^4}$. Therefore

$$\mathop{\mathbb{E}}_x\left[f^4_{=1}(x) - \frac{\eta^4}{8192\delta^4}f^2_{=1}(x)\right] \geqslant \frac{\eta^5}{4096\delta^4\binom{\ell}{1}},$$

and by an averaging argument (Clearly the function inside the expectation is upper bounded by $f^4_{=1}(x) \leqslant 2$), with probability at least $\frac{\eta^5}{8192\delta^4\binom{\ell}{1}}$ over random choice of x from V, we have that

$$f^4_{=1}(x) - \frac{\eta^4}{8192\delta^4}f^2_{=1}(x) > 0,$$

[7] Easy to conclude from Fact 2.11.

implying that $|f_{=1}(x)| \geqslant \frac{\eta^2}{91\cdot\delta^2}$. Thus, by Claim 2.26 $|f_{\approx 1}(x)| \geqslant |f_{=1}(x)| - \|f_{=1} - f_{\approx 1}\|_\infty \geqslant \frac{\eta^2}{100\delta^2}$. Recall that $f_{\approx 1}(x) = \mu_{\mathsf{Span}(x),\mathsf{in}}[F] - \delta \geqslant -\delta$, and since $\frac{\eta^2}{100\delta^2} > \delta$, it must be the case that $f_{\approx 1}(x) \geqslant \frac{\eta^2}{100\delta^2}$, implying $\mu_{\mathsf{Span}(x),\mathsf{in}}[F] \geqslant \delta + \frac{\eta^2}{100\delta^2}$.

If the second item holds, apply Lemma 5.7 with $\varepsilon = \frac{1}{2048\begin{bmatrix}\ell\\1\end{bmatrix}}\frac{\eta^5}{\delta^4}$ to conclude there is a hyperplane W with

$$\mu_{W,\mathsf{out}}[F] \geqslant \delta + \sqrt{\frac{\varepsilon}{2\eta}} \geqslant \delta + \frac{\eta^2}{100\delta^2}.$$

□

6 Improved Results for the First Level

In this we prove Theorem 3.5, restated below:

Theorem 3.5 (Restated) . *There exists $\ell_0 > 0$, such that for all $\ell > \ell_0$ the following holds.*

Let V be a k-dimensional vector space, $\varepsilon, \delta, \eta > 0$ such that $k \geqslant 30\ell^4 + \frac{100}{\varepsilon^3}$, $\sqrt{\frac{40}{\log(1/\delta)}} < \varepsilon \leqslant 1$.

Let $F\colon \begin{bmatrix}V\\\ell\end{bmatrix} \to \{0,1\}$, and assume $W^{=1}[F] = \eta$, $\mu(F) = \delta$. If $\eta \geqslant \delta^{2-\varepsilon}$, then at least one of the two must happen:

1. *There is x such that $\mu_{\mathsf{Span}(x),\mathsf{in}} \geqslant \delta + 2^{-\frac{20}{\varepsilon^2}}\frac{\eta^{1+\frac{\varepsilon}{4}}}{\delta}$.*
2. *There exists W of co-dimension 1, such that $\mu_{W,\mathsf{out}}[F] \geqslant \mu(F) + 2^{-\frac{20}{\varepsilon^2}}\frac{\eta^{1+\frac{\varepsilon}{4}}}{\delta}$.*

The proof follows the same lines of the proof of Theorem 3.3, however since a smaller weight on the first level is assumed, it requires studying higher moments of the first level component. Technically speaking, more types of sums appear in such moments and one needs to be able to control all of them. Fortunately there is a simple trick, namely Lemma 6.5, that allows handling all types of sums single-handedly.

6.1 Deviation of $F_{\approx r}$

The first part of the proof estimates higher moments of the first level component. For technical reason, we show this for the approximated first level component. Since this argument is applicable to component parts on all levels, we present it in a more general form.

Let $2m$ be an even integer, and let F be as in Theorem 3.5. Let $f_{\approx r}$,$F_{\approx r}$ be the functions from Definitions 2.23, 2.24, and denote $\eta = W^{=r}[F]$, $\delta = \mu(F)$ throughout this section.

Claim 6.1.

$$\mathop{\mathbb{E}}_{L}\left[F_{\approx r}[L]^2\right] = \eta + O(2^{13\ell^4 - \frac{1}{2}k}).$$

Proof. By the triangle inequality

$$\|F_{=r}\|_2 - \|F_{=r} - F_{\approx r}\|_2 \leqslant \|F_{\approx r}\|_2 \leqslant \|F_{=r}\|_2 + \|F_{=r} - F_{\approx r}\|_2.$$

Using Theorem 2.25, the fact that $\|F_{=r}\|_2^2 = \eta$, and squaring finishes the proof. □

Claim 6.2. *Suppose $k \geqslant 27\ell^4$, $\ell > \ell_0$. Then*

$$\mathop{\mathbb{E}}_{x}\left[f^2_{\approx r}(x)\right] = \frac{\eta}{\binom{\ell}{1}} + O(2^{13\ell^4 - \frac{1}{2}k}).$$

Proof. By Claim A.1 since F is Boolean, we have that $\|f_{\approx r}\|_\infty = O(2^{\ell^3})$. By Lemma A.2, for every L_{r-1}, the average of $f_{\approx r}$ on r-dimensional subspaces containing L_{r-1} is at most $\varepsilon = O(2^{10\ell^4 - k})$. Therefore, using Lemma 5.1 we get that

$$\mathop{\mathbb{E}}_{L}\left[F_{\approx r}[L]^2\right] = \binom{\ell}{1}\mathop{\mathbb{E}}_{x}\left[f^2_{\approx r}(x)\right] + O(2^{13\ell^4 - \frac{1}{2}k}).$$

By Claim 6.1, the left hand side is $\eta + O(2^{13\ell^4 - \frac{1}{2}k})$. Rearranging finishes the proof.

□

Lemma 6.3. *Suppose $k \geqslant 27\ell^4 + 2\log\frac{1}{\eta}$, $\ell > \ell_0$. Then*

$$\Pr_{L}\left[F_{\approx r}[L] \geqslant \frac{\eta}{4\delta}\right] \geqslant \frac{1}{4}\eta.$$

Proof. By the triangle inequality,

$$\|F - F_{\approx r}\|_2 \leqslant \|F - F_{=r}\|_2 + \|F_{=r} - F_{\approx r}\|_2 \leqslant \sqrt{\delta - \eta} + O(2^{13\ell^4 - \frac{1}{2}k}),$$

The last inequality is by Claim 5.2 and the fact that

$$\|F - F_{=r}\|_2^2 = \|F\|_2^2 - \|F_{=r}\|_2^2 = \delta - \eta.$$

Therefore

$$\mathop{\mathbb{E}}_{L}\left[(F - F_{\approx r})^2[L]\right] \leqslant \delta - \eta + O(2^{13\ell^4 - \frac{1}{2}k}) \leqslant \delta - \frac{3}{4}\eta.$$

From this point on, the proof follows a similar argument to the one in Lemma 5.3. Using Markov's inequality,

$$\Pr_{L}\left[(F - F_{\approx r})^2[L] \geqslant 1 - \frac{\eta}{2\delta}\right] \leqslant \frac{\delta - \frac{3}{4}\eta}{1 - \frac{\eta}{2\delta}} \leqslant \delta - \frac{1}{4}\eta,$$

and so with probability at least $1 - \delta + \frac{1}{4}\eta$ over the choice of L, $(F - F_{\approx r})^2[L] \leqslant 1 - \frac{\eta}{2\delta}$. Since with probability at least δ we have that $F[L] = 1$, we get that with probability at least $\frac{1}{4}\eta$ over the choice of L,

$$(1 - F_{\approx r}[L])^2 \leqslant 1 - \frac{\eta}{2\delta},$$

which, by $\sqrt{1-\varepsilon} \leqslant 1 - \frac{1}{2}\varepsilon$, implies that $|1 - F_{\approx r}[L]| \leqslant 1 - \frac{\eta}{4\delta}$. Opening the absolute value finishes the proof. □

The following lemma obtains a lower bound on high moments of the approximated first component:

Lemma 6.4. *Suppose $k \geqslant 27\ell^4 + 2\log\frac{1}{\eta}$, $\ell > \ell_0$. Then*

$$\mathop{\mathbb{E}}_{L}\left[F_{\approx r}[L]^{2m}\right] \geqslant \frac{\eta^{2m+1}}{2^{4m+2}\delta^{2m}}.$$

Proof. By Lemma 6.3,

$$\mathop{\mathbb{E}}_{L}\left[F_{\approx r}[L]^{2m}\right] \geqslant \Pr_{L}\left[F_{\approx r}[L] \geqslant \frac{\eta}{4\delta}\right]\left(\frac{\eta}{4\delta}\right)^{2m} \geqslant \frac{\eta^{2m+1}}{2^{4m+2}\delta^{2m}}.$$

□

6.2 A Fourier-Analytic lemma

A classical fourier-analytic computation shows that for any function $f\colon V \to \mathbb{R}$,

$$\left| \mathop{\mathbb{E}}_{x,y \in V} [f(x)f(y)f(x+y)] \right| \leqslant \|\widehat{f}\|_\infty \|f\|_2^2. \tag{6}$$

Estimates of the above type, namely expected values of product of f on linear combinations of randomly chosen points, have proven very useful (see for example [Rot53, Tre98, ST00, Hås01, HW03, HK05, KS13]). The following lemma is another generalization of (6).

Let $s \leqslant m$ be an integer, and let M be a binary $(2m-s) \times s$ matrix of rank $r \leqslant s$. For $x_1, ..., x_s \in V$, we denote $(y_1, .., y_{2m-s}) = M \cdot (x_1, x_2, ..., x_s)^{\mathrm{T}}$ - in words, y_i is the linear combination of x's, when the coefficient of x_j is $M[i,j]$. We write $\vec{y} = M \cdot \vec{x}$ in short.

Lemma 6.5. *Let $s \leqslant m$ be an integer, and M a binary $(2m-s) \times s$ matrix of rank $r \leqslant s$. Let $f\colon V \to \mathbb{R}$ be a function. Then*

$$\left| \mathop{\mathbb{E}}_{\substack{x_1,...,x_s \in V \\ \vec{y}=M\cdot\vec{x}}} [f(x_1)\cdots f(x_s)f(y_1)\cdots f(y_{2m-s})] \right| \leqslant \|f\|_\infty^{2m-s-r} \|\widehat{f}\|_\infty^{s-r} \|f\|_2^{2r}.$$

We defer the proof to the end of this section.

6.3 Proof of Theorem 3.5

Choose with foresight $m = \lceil \frac{2}{\varepsilon} \rceil$, and consider $\mathbb{E}_L \left[F_{\approx 1}[L]^{2m} \right]$. Our main task is to show an upper bound on the following quantity by expanding it out:

$$\mathbb{E}\left[F_1[L]^{2m}\right] = \mathop{\mathbb{E}}_{L} \left[\left(\sum_{x \in L \setminus \{0\}} f_{\approx 1}(x) \right)^{2m} \right].$$

Note that picking $2m$ terms from the sum can be equivalently done by first choosing s from 1 to $2m$ to be the number of linearly independent vectors, then choosing s linearly independent vectors from L and the rest $2m-s$ vectors are chosen from the span of the first. In other words, we can write

$$\begin{aligned}
\mathbb{E}\left[\tilde{F}_1[L]^{2m}\right] &= \mathop{\mathbb{E}}_{L} \left[\sum_{s=1}^{2m} \sum_{\substack{x_1,...,x_s \in L\setminus\{0\} \\ \text{linearly independent}}} f_{\approx 1}(x_1)\cdots f_{\approx 1}(x_s) \sum_{y_1,...,y_{2m-s} \in \mathsf{Span}(x_1,...,x_s)\setminus\{0\}} f_{\approx 1}(y_1)\cdots f_{\approx 1}(y_{2m-s}) \right] \\
&= \mathop{\mathbb{E}}_{L} \left[\sum_{s=1}^{2m} \alpha(s,m,\ell) \mathop{\mathbb{E}}_{\substack{x_1,...,x_s \in L\setminus\{0\} \\ \text{linearly independent} \\ y_1,...,y_{2m-s} \in \mathsf{Span}(x_1,...,x_s)\setminus\{0\}}} \left[f_{\approx 1}(x_1)\cdots f_{\approx 1}(x_s) \sum f_{\approx 1}(y_1)\cdots f_{\approx 1}(y_{2m-s}) \right] \right],
\end{aligned}$$

where $\alpha(s,m,\ell) = (2^\ell - 1)\cdots(2^\ell - 2^{s-1})(2^m-1)^{2m-s}$. Next, note that the distribution on $x_1,...,x_s$ is $O(s2^{-k})$-close to uniform, and the last expression is at most

$$\leqslant \sum_{s=1}^{2m} 2^{s\ell}2^{2ms} \left| \mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V \\ M(2m-s)\times s \text{ binary matrix} \\ \vec{y}=M\vec{x}}} [f_{\approx 1}(x_1)\cdots f_{\approx 1}(x_s)\cdot f_{\approx 1}(y_1)\cdots f_{\approx 1}(y_{2m-s})] \right| + O(2^{2m\ell}2^{-k}) \quad (7)$$

where we have used $\alpha(s,m,\ell) \leqslant 2^{s\ell}2^{2ms}$ and $\|f_{\approx 1}\|_\infty \leqslant 1$.

Fix s and a matrix M, and let $r = \mathsf{rank}(M)$. We consider the case where $s \leqslant m$ and $s > m$ separately. Denote

$$R \stackrel{def}{=} \max\left\{ \left|\mu_{\mathsf{Span}(x),\mathsf{in}} - \delta\right|, \left|\mu_{W,\mathsf{out}} - \delta\right| \;\middle|\; x \in V, W \subseteq V \text{ hyperplane} \right\}. \quad (8)$$

For $s \leqslant m$, apply Lemma 6.5 to obtain

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V \\ M(2m-s)\times s \text{ binary matrix} \\ \vec{y}=M\vec{x}}} [f_{\approx 1}(x_1)\cdots f_{\approx 1}(x_s)\cdot f_{\approx 1}(y_1)\cdots f_{\approx 1}(y_{2m-s})] \right| \\
&\leqslant \|f_{\approx 1}\|_\infty^{2m-s-r} \|\widehat{f_{\approx 1}}\|_\infty^{s-r} \|f_{\approx 1}\|_2^{2r} \\
&\leqslant R^{2m-s-r}\left(\begin{bmatrix}\ell\\1\end{bmatrix}^{-1} R\right)^{s-r}\left(\frac{\eta}{\begin{bmatrix}\ell\\1\end{bmatrix}}\right)^{r} + O\left(2^{16\ell^4-k}\right) \\
&\leqslant 2\cdot 2^{-\ell s} R^{2m-2r}\eta^r + O\left(2^{13\ell^4-\frac{1}{2}k}\right),
\end{aligned}$$

where the first inequality is by Lemma 6.5, the second inequality is by the definition of $f_{\approx 1}$, Lemma 4.2, and Claim 6.2.

For $s > m$ we apply the previous argument dually. More formally, we use the following lemma (the proof is deferred to the next section):

Lemma 6.6. *Let M be an $(2m-s)\times s$ matrix of rank r, $f\colon V \to \mathbb{R}$, and define $A_i = \{\, j \mid M[j,i] = 1\,\}$. Then*

$$\mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V \\ \vec{y}=M\vec{x}}} [f(x_1)\cdots f(x_s)\cdot f(y_1)\cdots f(y_{2m-s})] = \sum_{\substack{T_1,\ldots,T_{2m-s} \\ S_i = \bigoplus_{j\in A_i} T_j}} \widehat{f}(S_1)\cdots\widehat{f}(S_s)\cdot\widehat{f}(T_1)\cdots\widehat{f}(T_{2m-s}).$$

Using Lemma 6.6 we see that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V \\ \vec{y}=M\vec{x}}} [f_{\approx 1}(x_1)\cdots f_{\approx 1}(x_s)\cdot f_{\approx 1}(y_1)\cdots f_{\approx 1}(y_{2m-s})] \\
&\quad = \sum_{\substack{T_1,\ldots,T_{2m-s} \\ S_i = \bigoplus_{j\in A_i} T_j}} \widehat{f_{\approx 1}}(S_1)\cdots\widehat{f_{\approx 1}}(S_s)\cdot\widehat{f_{\approx 1}}(T_1)\cdots\widehat{f_{\approx 1}}(T_{2m-s}) \\
&\quad = 2^{(2m-s)k} \mathop{\mathbb{E}}_{\substack{T_1,\ldots,T_{2m-s} \\ S_i = \bigoplus_{j\in A_i} T_j}} \left[\widehat{f_{\approx 1}}(S_1)\cdots\widehat{f_{\approx 1}}(S_s)\cdot\widehat{f_{\approx 1}}(T_1)\cdots\widehat{f_{\approx 1}}(T_{2m-s})\right],
\end{aligned}$$

where $A_i = \{ j \mid M[j,i] = 1 \}$. Now we have $2m - s < m$, and we can use Lemma 6.5 (with $\widehat{f_{\approx 1}}$ instead of f and renaming of the parameters, the new "s" is $2m - s$) to get that the previous expression is upper-bounded by

$$2^{(2m-s)k} \|\widehat{f_{\approx 1}}\|_\infty^{2m-(2m-s)-r} \|\widehat{\widehat{f_{\approx 1}}}\|_\infty^{(2m-s)-r} \|\widehat{f_{\approx 1}}\|_2^{2r}. \tag{9}$$

Note that $\widehat{\widehat{f_{\approx 1}}} = 2^{-k} f_{\approx 1}$, and therefore

$$\|\widehat{f_{\approx 1}}\|_2^{2r} = 2^{-kr} \left(\sum_S \widehat{f_{\approx 1}}^2(S) \right)^r = 2^{-kr} \|f_{\approx 1}\|_2^{2r}, \tag{10}$$

where the last equality is by Parseval. We use (10) to get

$$\begin{aligned}
(9) &\leqslant 2^{(2m-s)k} \|\widehat{f_{\approx 1}}\|_\infty^{s-r} 2^{-(2m-s-r)k} \|f_{\approx 1}\|_\infty^{2m-s-r} 2^{-kr} \|f_{\approx 1}\|_2^{2r} \\
&= \|\widehat{f_{\approx 1}}\|_\infty^{s-r} \|f_{\approx 1}\|_\infty^{2m-s-r} \|f_{\approx 1}\|_2^{2r} \\
&\leqslant \left(\begin{bmatrix} \ell \\ 1 \end{bmatrix}^{-1} R \right)^{s-r} R^{2m-s-r} \left(\frac{\eta}{\begin{bmatrix} \ell \\ 1 \end{bmatrix}} \right)^r + O\left(2^{16\ell^4 - k} \right) \\
&\leqslant 2 \cdot 2^{-s\ell} R^{2m-2r} \eta^r + O\left(2^{13\ell^4 - \frac{1}{2}k} \right),
\end{aligned}$$

where in the third inequality we bounded $\|\widehat{f_{\approx 1}}\|_\infty$ by Lemma 4.2, $\|f_{\approx 1}\|_\infty$ by definitions, and $\|f_{\approx 1}\|_2$ by Claim 6.2.

We plug in both estimates (of $s \leqslant m$ and $s > m$) to (7) to get that

$$\begin{aligned}
\mathop{\mathbb{E}}_L \left[F_{\approx 1}[L]^{2m} \right] &\leqslant \sum_{s=1}^{2m} 2^{s\ell} 2^{2ms} \sum_{r=1}^{\min\{s, 2m-s\}} \left(2 \cdot 2^{-s\ell} R^{2m-2r} \eta^r + O\left(2^{13\ell^4 - \frac{1}{2}k} \right) \right) + O(2^{2m\ell} 2^{-k}) \\
&\leqslant \sum_{s=1}^{2m} \sum_{r=1}^{s} 2^{4m^2+1} R^{2m-2r} \eta^r + O\left(2^{14\ell^4 - \frac{1}{2}k} \right).
\end{aligned}$$

Therefore, there are s, r such that

$$2^{4m^2+1} R^{2m-2r} \eta^r \geqslant \frac{1}{4m^2} \left[\mathop{\mathbb{E}}_L \left[F_{\approx 1}[L]^{2m} \right] - O\left(2^{14\ell^4 - k} \right) \right] \geqslant \frac{\eta^{2m+1}}{2^{4m+5} m^2 \delta^{2m}},$$

where the last inequality is by Lemma 6.4, and the fact that the error term is at most half the main term for sufficiently large ℓ. Rearranging and using $m^2 \leqslant 2^m$ yields

$$R^{2m} \left(\frac{\eta}{R^2} \right)^r \geqslant \frac{\eta^{2m+1}}{2^{4m^2+5m+6} \delta^{2m}}. \tag{11}$$

Next, we claim that the parameter r in equation 11 is strictly smaller than m:

Proposition 6.7. $r < m$.

Proof. Assume toward contradiction $r = m$. We get $\eta^m \geqslant \frac{\eta^{2m+1}}{2^{4m^2+5m+6} \delta^{2m}}$, concluding that

$$\delta^{2-\varepsilon} \leqslant \eta \leqslant 2^{4m+6} \delta^{2\frac{m}{m+1}} \leqslant 2^{4m+6} \delta^{2-\frac{1}{m+1}}.$$

Rearranging and using $\delta^{-1/(m+1)} \leqslant \delta^{-\frac{1}{2}\varepsilon}$ yields

$$\delta^{-\frac{1}{2}\varepsilon} \leqslant 2^{4m+6},$$

taking logarithm, using the definition of m and rearranging implies

$$\varepsilon^2 \leqslant \frac{16+20\varepsilon}{\log 1/\delta} < \frac{40}{\log 1/\delta},$$

contradicting the condition on ε. □

If $\eta \leqslant R^2$ then the left hand side in (11) is at most R^{2m} and we conclude that

$$R^{2m} \geqslant \frac{\eta^{2m+1}}{2^{4m^2+5m+6}\delta^{2m}},$$

implying

$$R \geqslant 2^{-2m-6}\frac{\eta^{1+1/2m}}{\delta} \geqslant 2^{-\frac{8}{\varepsilon}-3}\frac{\eta^{1+\varepsilon/4}}{\delta}.$$

Else, since $r \leqslant m-1$ we get that

$$R^2 \geqslant 2^{-16m^2}\frac{\eta^{m+2}}{\delta^{2m}},$$

implying

$$R \geqslant 2^{-8m^2}\frac{\eta^{\frac{1}{2}m+1}}{\delta^m} \geqslant 2^{-\frac{80}{\varepsilon^2}}\frac{\eta}{\delta}.$$

the last inequality follows from the definition of m. Either way, we have

$$R \geqslant 2^{-\frac{80}{\varepsilon^2}}\frac{\eta^{1+\varepsilon/4}}{\delta}.$$

Using the lower bound we know on ε, η, it is easy to see that the last expression is greater than δ. Recalling the definition of R from (8) finishes the proof.

Auxiliary lemmas

Proof of Lemma 6.6. Expand out each $f(x_i), f(y_i)$ to its fourier representation to get

$$\mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V\\ \vec{y}=M\vec{x}}}[f(x_1)\cdots f(x_s)\cdot f(y_1)\cdots f(y_{2m-s})]$$

$$= \mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V\\ \vec{y}=M\vec{x}}}\left[\sum_{\substack{T_1,\ldots,T_{2m-s}\\ S_1,\ldots,S_s}} \widehat{f}(S_1)\chi_{S_1}(x_1)\cdots\widehat{f}(S_s)\chi_{S_s}(x_s)\cdot\widehat{f}(T_1)\chi_{T_1}(y_1)\cdots\widehat{f}(T_{2m-s})\chi_{T_{2m-s}}(y_{2m-s})\right].$$

Note that

$$\chi_{S_1}(x_1)\cdots\chi_{S_s}(x_s)\cdot\chi_{T_1}(y_1)\cdots\chi_{T_{2m-s}}(y_{2m-s}) = \chi_{S_1\oplus\bigoplus_{j\in A_1}T_j}(x_1)\cdots\chi_{S_s\oplus\bigoplus_{j\in A_s}T_j}(x_s).$$

Thus for a summand not to be 0, it must be the case that $S_i = \bigoplus_{j\in A_i}T_j$, and the proof is concluded. □

Lemma 6.5 (Restated) . *Let $s \leqslant m$ be an integer, and M a binary $(2m-s)\times s$ matrix of rank $r \leqslant s$. Let $f\colon V \to \mathbb{R}$ be a function. Then*

$$\left| \mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V\\ \vec{y}=M\cdot\vec{x}}} [f(x_1)\cdots f(x_s)f(y_1)\cdots f(y_{2m-s})] \right| \leqslant \|f\|_\infty^{2m-s-r}\|\widehat{f}\|_\infty^{s-r}\|f\|_2^{2r}.$$

Proof. Since the matrix M has rank r, there is a subset of r of the rows such that each other row is a linear combination of them. Such subset of the rows corresponds to a subset of the y's that form a basis for $\mathsf{Span}(y_1,...,y_{2m-r})$ - without loss of generality assume $y_1,..,y_r$ is a basis. Then each y_i for $i > r$ can be expressed as a linear combination of $y_1,...,y_r$. In particular, the function $f(y_1)\cdots f(y_{2m-s})$ is a function of $y_1,...,y_r$, which we denote by:

$$h(y_1,...,y_r) = f(y_1)\cdots f(y_{2m-s}).$$

Therefore the expectation we want to evaluate is

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V\\ \vec{y}=M\cdot\vec{x}}} [f(x_1)\cdots f(x_s)h(y_1,...,y_r)]\\
&= \mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V\\ \vec{y}=M\cdot\vec{x}}} \left[\sum_{S_1}\widehat{f}(S_1)\chi_{S_1}(x_1)\cdots\sum_{S_s}\widehat{f}(S_s)\chi_{S_s}(x_s)\sum_{T_1,\ldots,T_s}\widehat{h}(T_1,...,T_r)\chi_{T_1}(y_1)\cdots\chi_{T_r}(y_r)\right]\\
&\sum_{\substack{S_1,\ldots,S_s\\ T_1,\ldots,T_r}}\widehat{f}(S_1)\cdots\widehat{f}(S_s)\widehat{h}(T_1,...,T_r)\mathop{\mathbb{E}}_{\substack{x_1,\ldots,x_s\in V\\ \vec{y}=M\cdot\vec{x}}}[\chi_{S_1}(x_1)\cdots\chi_{S_s}(x_s)\cdot\chi_{T_1}(y_1)\cdots\chi_{T_s}(y_r)]
\end{aligned}$$

For every $i = 1,...,s$ consider

$$A_i = \{\, j \mid M[j,i] = 1\,\},$$

in words, the set of j such that x_i appears in the representation of y_j according to the x's. Then

$$\chi_{S_1}(x_1)\cdots\chi_{S_s}(x_s)\cdot\chi_{T_1}(y_1)\cdots\chi_{T_r}(y_r) = \prod_{i=1}^{s}\chi_{S_i\oplus\bigoplus_{j\in A_i}T_j}(x_i).$$

Thus our original expression equals

$$\begin{aligned}
&\sum_{\substack{S_1,\ldots,S_s\\ T_1,\ldots,T_r}}\widehat{f}(S_1)\cdots\widehat{f}(S_s)\widehat{h}(T_1,...,T_r)\mathop{\mathbb{E}}_{x_1,\ldots,x_s\in V}\left[\prod_{i=1}^{s}\chi_{S_i\oplus\bigoplus_{j\in A_i}T_j}(x_i)\right]\\
&= \sum_{\substack{S_1,\ldots,S_s\\ T_1,\ldots,T_r}}\widehat{f}(S_1)\cdots\widehat{f}(S_s)\widehat{h}(T_1,...,T_r)\prod_{i=1}^{s}\mathop{\mathbb{E}}_{x_i}\left[\chi_{S_i\oplus\bigoplus_{j\in A_i}T_j}(x_i)\right]\\
&= \sum_{\substack{T_1,\ldots,T_r\\ S_i=\bigoplus_{j\in A_i}T_j}}\widehat{f}(S_1)\cdots\widehat{f}(S_s)\widehat{h}(T_1,...,T_r).
\end{aligned}$$

Applying Cauchy-Schwartz on the last sum implies it is upper bounded by:

$$\leqslant \sqrt{\sum_{\substack{T_1,...,T_r \\ S_i=\bigoplus_{j\in A_i} T_j}} \widehat{f}^2(S_1)\cdots\widehat{f}^2(S_s)} \sqrt{\sum_{\substack{T_1,...,T_r \\ S_i=\bigoplus_{j\in A_i} T_j}} \widehat{h}^2(T_1,...,T_r)} \tag{12}$$

To estimate the first expectation, let us abuse notation and look at characters as $\{0,1\}$-valued vectors. This way, we can succinctly write that $(S_1,...,S_s) = M_s^{\mathrm{T}}(T_1,...,T_r)^{\mathrm{T}}$, where M_s is the $s\times r$ matrix whose rows are the first r rows of M. In particular, since M_s^{T}has rank r, we can choose a subset of r of its rows that are linearly independent - without loss of generality we assume the first r rows are such. Then

$$\begin{aligned}
\sum_{\substack{T_1,...,T_r \\ S_i=\bigoplus_{j\in A_i} T_j}} \widehat{f}^2(S_1)\cdots\widehat{f}^2(S_s) &\leqslant \|\widehat{f}^2\|_\infty^{s-r} \sum_{\substack{T_1,...,T_r \\ S_i=\bigoplus_{j\in A_i} T_j}} \widehat{f}^2(S_1)\cdots\widehat{f}^2(S_r) \\
&= \|\widehat{f}^2\|_\infty^{s-r} \sum_{S_1,...,S_r} \widehat{f}^2(S_1)\cdots\widehat{f}^2(S_r) \\
&= \|\widehat{f}^2\|_\infty^{s-r} \left(\sum_{S_1} \widehat{f}^2(S_1)\right)^r \\
&= \|\widehat{f}^2\|_\infty^{s-r} \mathop{\mathbb{E}}_x \left[f^2(x)\right]^r \\
&= \|\widehat{f}^2\|_\infty^{s-r} \|f\|_2^{2r}.
\end{aligned}$$

The first inequality is by taking out the maximal fourier coefficient, the first equality holds since the first r rows of M_S^{T} are linearly independent, the third equality is by Parseval.

To evaluate the second term in Equation (12) we use Parseval:

$$\begin{aligned}
\sum_{\substack{T_1,...,T_r \\ S_i=\bigoplus_{j\in A_i} T_j}} \widehat{h}^2(T_1,...,T_r) &= \mathop{\mathbb{E}}_{y_1,...,y_r} \left[h^2(y_1,...,y_r)\right] \\
&= \mathop{\mathbb{E}}_{y_1,...,y_r} \left[f^2(y_1)\cdots f^2(y_r)\cdots f^2(y_{2m-s})\right] \\
&\leqslant \|f^2\|_\infty^{2m-s-r} \mathop{\mathbb{E}}_{y_1,...,y_r} \left[f^2(y_1)\cdots f^2(y_r)\right] \\
&= \|f^2\|_\infty^{2m-s-r} \mathop{\mathbb{E}}_{y_1,...,y_r} \left[f^2(y_1)\right]^r \\
&= \|f^2\|_\infty^{2m-s-r} \|f\|_2^{2r}.
\end{aligned}$$

Plugging the last two estimate into Equation (12), we get that our original expression is upper bounded by

$$\sqrt{\|\widehat{f}^2\|_\infty^{s-r}\|f\|_2^{2r}\|f^2\|_\infty^{2m-s-r}\|f\|_2^{2r}} = \|f\|_\infty^{2m-s-r}\|\widehat{f}\|_\infty^{s-r}\|f\|_2^{2r}$$

□

7 Results for the Second Level

In this section we prove Theorem 3.7. The proof follows a similar strategy as in the proof of Theorem 3.3 – upper-bounding the fourth moment of the level 2 component of F – but it is technically more involved:

recall that the value of the level 2 function at L is a sum of values of a function defined over 2 dimensional spaces. Hence when computing the fourth moment one obtains various products of terms that behave differently, depending on the intersection pattern of the corresponding spaces. Since there are many intersection patterns between four 2-dimensional subspaces, simply enumerating over all of them, analyzing each pattern separately may be infeasible.

We currently do not have a fully systematic way of dealing with all types of summands at once – this issue seems to be the bottleneck in extending our results to all levels. For level 2 however we still manage to partition the terms into a relatively small number of classes, and then to bound each class separately.

Structure of Section 7. We begin in Section 7.1 by associating $f_{\approx 2}$ with a function over the hypercube, and establishing a connection between its fourier coefficients and zoom-outs – this is the level 2 analogue of Section 4. In Section 7.3 we define the key definitions that allow us to partition the terms into relatively small number of classes, present a set of example-terms that arise in the proof, and analyze them. Finally, in Section 7.4 we argue that those examples cover all types of terms (up to symmetries), and in Section 7.5 we conclude the proof of Theorem 3.7.

7.1 Fourier coefficients and zoom-outs on the second level

Let $F\colon \binom{V}{\ell} \to \mathbb{R}$ and consider $f_{\approx 2}\colon \binom{V}{2} \to \mathbb{R}$. Formally $f_{\approx 2}$ is a function on 2 dimensional subspaces, thus to discuss its fourier expansion we first have to extends its range to V^2.

Definition 7.1. *We identify $f_{\approx 2}$ from Definition 2.23 with $f_{\approx 2}\colon V^2 \to \mathbb{R}$ by*

$$f_{\approx 2}(x,y) = \begin{cases} f_{\approx 2}(\mathsf{Span}(x,y)) & x,y \text{ linearly independent}, \\ \mu - \mu_{\mathsf{Span}(x),\mathsf{in}} & x = y \neq 0 \text{ or } x \neq 0, y = 0 \\ \mu - \mu_{\mathsf{Span}(y),\mathsf{in}} & x = 0, y \neq 0, \\ 0 & x = y = 0. \end{cases}$$

Note that the two definitions agree whenever x, y span a 2-dimensional subspace.

This extension can be viewed to be natural by expanding out $\tilde{f}_2$ in terms on $\mu_{\circ,\mathsf{in}}$. We begin by noting few basic properties of $\tilde{f}_2$. We omit the self-evident proof.

Claim 7.2. *Let $F\colon \binom{V}{\ell} \to \mathbb{R}$, and $f_{\approx 2}\colon V^2 \to \mathbb{R}$ be from Definition 7.1.*

1. *For every x, $\mathbb{E}_y\left[f_{\approx 2}(x,y)\right] = 0$.*
2. *$f_{\approx 2}(x,y) = f_{\approx 2}(x, x \oplus y)$ for every x, y.*
3. *$f_{\approx 2}(x,y) = f_{\approx 2}(y,x)$ for every x, y.*

A few basic properties of $\widehat{f_{\approx 2}}$ that can be derived from the properties of $f_{\approx 2}$.

Claim 7.3. *Let $F\colon \binom{V}{\ell} \to \mathbb{R}$, and $f_{\approx 2}\colon V^2 \to \mathbb{R}$ be from Definition 7.1.*

1. *$\widehat{f_{\approx 2}}(S, \emptyset) = 0$ for every S.*
2. *For every S_1, S_2, $\widehat{f_{\approx 2}}(S_1, S_2) = \widehat{f_{\approx 2}}(S_1, S_1 \oplus S_2)$.*
3. *For every S_1, S_2, $\widehat{f_{\approx 2}}(S_1, S_2) = \widehat{f_{\approx 2}}(S_2, S_1)$.*

Proof. For the first item,

$$\widehat{f_{\approx 2}}(S,\emptyset) = \mathop{\mathbb{E}}_{x,y}[f_{\approx 2}(x,y)\chi_S(x)] = \mathop{\mathbb{E}}_{x}\left[\chi_S(x)\mathop{\mathbb{E}}_{y}[f_{\approx 2}(x,y)]\right] = \mathop{\mathbb{E}}_{x}[\chi_S(x)\cdot 0] = 0.$$

For the second item,

$$\begin{aligned}
\widehat{f_{\approx 2}}(S_1,S_2) &= \mathop{\mathbb{E}}_{x,y}[f_{\approx 2}(x,y)\chi_{S_1}(x)\chi_{S_2}(y)] \\
&= \mathop{\mathbb{E}}_{x,y}[f_{\approx 2}(x\oplus y,y)\chi_{S_1}(x)\chi_{S_2}(y)] \\
&= \mathop{\mathbb{E}}_{x,y}[f_{\approx 2}(x\oplus y,y)\chi_{S_1}(x\oplus y)\chi_{S_1\oplus S_2}(y)] \\
&= \mathop{\mathbb{E}}_{z,y}[f_{\approx 2}(z,y)\chi_{S_1}(z)\chi_{S_1\oplus S_2}(y)] \\
&= \widehat{f_{\approx 2}}(S_1,S_1\oplus S_2),
\end{aligned}$$

we used $f_{\approx 2}(x,y) = f_{\approx 2}(x\oplus y,y)$.

The third item similarly follows from $f_{\approx 2}(x,y) = f_{\approx 2}(y,x)$. □

The previous claim implies that $f_{\approx 2}$ is essentially a function on two-dimensional subspaces.

The following lemma establishes a connection between the fourier coefficients of $\tilde{f}_2$ and zoom-outs. The proof goes along similar lines the proof of Lemma 4.2, however, includes more tedious calculations.

Lemma 7.4. *Let $F\colon \binom{V}{\ell} \to \mathbb{R}$, $f_{\approx 2}\colon V^2 \to \mathbb{R}$ from Definition 2.23, and let $S_1, S_2 \subseteq [k]$ be distinct and non-empty. Then*

$$\begin{aligned}
\widehat{f_{\approx 2}}(S_1,S_2) =& c_2(k,\ell)\left[\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}}(F) - \mu(F) - (\mu_{W_{S_1\oplus S_2},\mathsf{out}}(F) + \mu_{W_{S_1},\mathsf{out}}(F) + \mu_{W_{S_2},\mathsf{out}}(F) - 3\mu(F))\right] \\
&+ O(2^{\ell-k}\|F\|_\infty),
\end{aligned}$$

where

$$c_2(k,\ell) = \frac{\binom{k-2}{\ell}}{\binom{k}{\ell} - 3\binom{k-1}{\ell} + 2\binom{k-1}{\ell-1}} = 2^{-2\ell} + O(2^{-3\ell}).$$

We defer the proof to Section E.1.

Remark 7.5. *With a little more effort it is possible to prove the following: there are constants $c_1(k,\ell), c_2(k,\ell)$ such that*

$$\widehat{f_{\approx 2}}(S_1,S_2) = c_1\left[\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}} - \mu(F)\right] - c_2\left[\mu_{W_{S_1},\mathsf{out}} + \mu_{W_{S_2},\mathsf{out}} + \mu_{W_{S_1\oplus S_2},\mathsf{out}} - 3\mu\right],$$

and $c_1 = 2^{-2\ell} + O(2^{-3\ell}), c_2 = 2^{-2\ell} + O(2^{-3\ell})$.

7.2 Basic claims and notations

Notations. Let $F_{\approx 2}, f_{\approx 2}$ be from Definition 2.24, 7.1. We will use the usual notations of $\delta = \mu(F), \eta = W^{=2}[F]$, and we assume F is $(2,\varepsilon)$ pseudo-random where $\varepsilon \geqslant \delta$. It will be convenient for us to prove 3.7 counter positively: assuming F is $(2,\varepsilon)$ pseudo-random, we derive an upper-bound on the weight of F on the second level.

For each $x \in V$, we choose an arbitrary complementing space M_x, namely a $k-1$ dimensional subspace of V such that $\mathsf{Span}(x) \oplus M_x = V$.

Definition 7.6. *For $F\colon \begin{bmatrix} V \\ \ell \end{bmatrix} \to \mathbb{R}$ and $x \in V$, we define $F_x\colon \begin{bmatrix} M_x \\ \ell-1 \end{bmatrix} \to \mathbb{R}$ by*

$$F_x[L'] = F[L' \oplus \mathsf{Span}(x)].$$

Definition 7.7. *For $W \subseteq V$, define $F_W\colon \begin{bmatrix} W \\ \ell \end{bmatrix} \to \mathbb{R}$ by*

$$F_W[L] = F[L].$$

The following claim asserts that $(2,\varepsilon)$ pseudo-randomness of F implies $(1, O(\varepsilon))$ pseudo-randomness of F, F_x, F_W for any $x \in V$ and hyperplane $W \subseteq V$.

Claim 7.8. *If F is $(2,\varepsilon)$ pseudo-random, then F is $(1,\varepsilon)$ pseudo-random, and for any $x \in V$ and hyperplane $W \subseteq V$, F_x, F_W are $(1, \varepsilon+\delta)$ pseudo-random.*

Proof. Let $x \in V$, and assume F_x is not $(1,\varepsilon+\delta)$ pseudo-random. Then there is x_2 or W_2 such that the density of F_{x,x_2} or F_{x,W_2} is at least the density of F_x plus $\delta+\varepsilon$ - assume for example that the first case holds (the other case is handled similarly). Then

$$\mu(F_{x,x_2}) \geqslant \mu(F_x) + \delta + \varepsilon \geqslant \delta + \varepsilon,$$

contradiction to the fact F is $(2,\varepsilon)$ pseudo-random. The proof for F_W is similar.

For F, assume it is not $(1,\varepsilon)$ pseudo-random. Then there is $x \in V$ ($W \subseteq$ hyperplane) such that the density of F_x (F_W) is at least $\delta+\varepsilon$ - assume for example that the first case holds (the other case is handled similarly). Clearly,

$$\mathop{\mathbb{E}}_{y \in M_x}[\mu(F_{x,y})] = \mu(F_x),$$

thus there is y such that $\mu(F_{x,y}) \geqslant \mu(F_x) \geqslant \delta+\varepsilon$, and contradiction to the fact F is $(2,\varepsilon)$ pseudo-random. □

We will often be interested about the implications of $(2,\varepsilon)$ pseudo-randomness of F on several measures that seem to be related to zoom-in/zoom-out of dimension one of F. This is the content of the following claim which will be repeatedly used later this section.

Claim 7.9. *Suppose F is $(2,\varepsilon)$ pseudo-random where $\varepsilon \geqslant \delta$, and define $h_x\colon V \to \mathbb{R}$ by $h_x(y) = f_{\approx 2}(x,y)$. Then*

1. $\|f_{\approx 2}\|_\infty \leqslant 4\varepsilon$.
2. $\|\widehat{f_{\approx 2}}\|_\infty \leqslant 4 \cdot 2^{-2\ell}(1+o(1))\varepsilon + O(2^{\ell-k})$.
3. $\|h_x\|_\infty \leqslant 8\varepsilon$.
4. *For any S_1,* $\sum\limits_{S_2} \widehat{f_{\approx 2}}(S_1,S_2)^2 \leqslant 900(1+o(1)) \cdot 2^{-3\ell}\varepsilon^{4/3} + O(2^{16\ell^4-k})$.
5. *For any x,* $\|\widehat{h_x}\|_\infty \leqslant 4 \cdot 2^{-\ell}(1+o(1))\varepsilon$.

Proof. Note that

$$f_{\approx 2}(x,y) = \mu_{\mathsf{Span}(x,y),\mathsf{in}}(F) - \delta + \delta - \mu_{\mathsf{Span}(x),\mathsf{in}}(F) + \delta - \mu_{\mathsf{Span}(y),\mathsf{in}}(F) + \delta - \mu_{\mathsf{Span}(x+y),\mathsf{in}}(F)$$

Each one of the four differences above is at most ε in absolute value by the $(2,\varepsilon)$ pseudo-randomness of F and the $(1,\varepsilon)$ pseudo-randomness of F (Claim 7.8).

For the second item, note that $\widehat{f_{\approx 2}}(S_1,S_2)$ is 0 if S_1, S_2 are linearly dependent (Claim 7.3); otherwise apply Lemma 7.4 to get

$$\begin{aligned}\widehat{f_{\approx 2}}(S_1,S_2) =& c_2(k,\ell)\left[\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}}(F) - \delta - (\mu_{W_{S_1\oplus S_2},\mathsf{out}}(F) + \mu_{W_{S_1},\mathsf{out}}(F) + \mu_{W_{S_2},\mathsf{out}}(F) - 3\delta)\right] \\ &+ O(2^{\ell-k}\|F\|_\infty),\end{aligned}$$

where $c_2(k,\ell)2^{-2\ell} + O(2^{-3\ell})$. Using the $(2,\varepsilon)$ pseudo-randomness of F and the $(1,\varepsilon)$ pseudo-randomness of F (Claim 7.8) we see that the densities of $F_{W_{S_1}}, F_{W_{S_2}}, F_{W_{S_1\oplus S_2}}, F_{W_{S_1}\cap W_{S_2}}$ is ε-close to δ, and so we get that.

$$\left|\widehat{f_{\approx 2}}(S_1,S_2)\right| \leqslant 4c_2(k,\ell)\varepsilon + O(2^{\ell-k}) = 4\cdot 2^{-2\ell}(1+o(1))\varepsilon + O(2^{\ell-k}).$$

The third item is immediate from the first.

For the fourth item, by Lemma 7.4 we have that

$$\begin{aligned}&\sum_{S_2}\widehat{f_{\approx 2}}^2(S_1,S_2) \\ &= c_2^2(k,\ell)\sum_{S_2}\left[\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}}(F) - \mu_{W_{S_1},\mathsf{out}}(F) + (2\mu(F) - \mu_{W_{S_1\oplus S_2},\mathsf{out}}(F) - \mu_{W_{S_2},\mathsf{out}}(F)) + O(2^{3\ell-k})\right]^2\end{aligned}$$

Use $(a_1+...+a_5)^2 \leqslant 5(a_1^2+...+a_5^2)$ to get that the above is upper-bounded by

$$\begin{aligned}5\cdot c_2(k,\ell)^2\Big[&\sum_{S_2}(\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}}(F) - \mu_{W_{S_1},\mathsf{out}}(F))^2 + \sum_{S_2}(\mu_{W_{S_1\oplus S_2},\mathsf{out}}(F) - \mu(F))^2 \\ &+ \sum_{S_2}(\mu_{W_{S_2},\mathsf{out}}(F) - \mu(F))^2 + \sum_{S_2}O(2^{6\ell-2k})\Big]\end{aligned} \tag{13}$$

We next estimate each sum separately. For the first sum, consider the function $G = F_{W_{S_1}}$, and let $g_{\approx 1}$ be its level 1 approximated point function [8].

By Lemma 4.2, $\widehat{g}_{\approx 1}(S_2) = c_1(k-1,\ell)(\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}}(F) - \mu_{W_{S_1},\mathsf{out}})$. Since the first sum counts each subspace twice ($W_{S_1}\cap W_{S_2}$ is counted for $S_2, S_1\oplus S_2$), we have that it equals

$$\begin{aligned}\sum_{S_2}(\mu_{W_{S_1}\cap W_{S_2},\mathsf{out}}(F) - \mu_{W_{S_1},\mathsf{out}})^2 &= \frac{2}{c_1(k-1,\ell)^2}\sum_{S_2}\widehat{g}^2(S_2) \\ &= \frac{2}{c_1(k-1,\ell)^2}\|g_{\approx 1}\|_2^2 \\ &\leqslant \frac{2}{c_1(k-1,\ell)^2}\left(\frac{W^{=1}[G]}{\begin{bmatrix}\ell\\1\end{bmatrix}} + O(2^{16\ell^4-k})\right).\end{aligned}$$

[8] I.e. the function for which $G_{\approx 1}[L] = \sum_{x\in L\setminus\{0\}} g_{\approx 1}(x)$ for all $L \subseteq W_{S_1}$.

The second equality is by Parseval, the third transition is by Claim 6.2. By Claim 7.8, G has density at most $\delta+\varepsilon \leqslant 2\varepsilon$ and is $(1,\varepsilon+\delta)$ pseudo-random, and in particular $(1,9\varepsilon^{2/3})$ pseudo-random. Hence by Corollary 3.4,

$$W^{=1}[G] \leqslant \sqrt{200\cdot 9\varepsilon^{2/3}}\mu(G) \leqslant 60\varepsilon^{4/3}.$$

Plugging this into the previous inequality, we get that the first sum in Expression (13) is at most

$$\frac{120}{\begin{bmatrix}\ell\\1\end{bmatrix}c_1(k-1,\ell)^2}\varepsilon^{4/3} + O(2^{2\ell+16\ell^4-k}).$$

The second and the third sum are equal and are estimated similarly; By Lemma 4.2 each one of them equals

$$\frac{1}{c_1(k,\ell)^2}\sum_{S_2}\widehat{f}^2_{\approx 1}(S_2) = \frac{1}{c_1(k,\ell)^2}\|f_{\approx 1}\|_2^2 \leqslant \left(\frac{W^{=1}[F]}{\begin{bmatrix}\ell\\1\end{bmatrix}} + O(2^{16\ell^4-k})\right),$$

in the first transition we used Parseval and in the second transition we used Claim 6.2. By Corollary 3.4, since F is $(1,\varepsilon)$ pseudo-random (Claim 7.8), it is in particular $(1,\frac{9}{2}\delta^{2/3})$ pseudo-random and we have that

$$W^{=1}[F] \leqslant \sqrt{200\cdot\frac{9}{2}\delta^{2/3}}\delta \leqslant 30\varepsilon^{4/3},$$

(we used $\delta \leqslant \varepsilon$) thus the second and third sum are bounded by

$$\frac{30}{\begin{bmatrix}\ell\\1\end{bmatrix}c_1(k-1,\ell)^2}\varepsilon^{4/3} + O(2^{2\ell+16\ell^4-k}).$$

The fourth sum is $O(2^{6\ell-k})$.

Combining all of the above, we get that the Expression in (13) is at most

$$5c_2(k,\ell)^2\cdot\frac{180}{\begin{bmatrix}\ell\\1\end{bmatrix}c_1(k-1,\ell)^2}\varepsilon^{4/3} + O(2^{16\ell^4-k}).$$

Since $c_2 = 2^{-2\ell}+O(2^{-3\ell})$, $c_1 = 2^{-\ell}+2^{-2\ell}$ and $\begin{bmatrix}\ell\\1\end{bmatrix} = 2^\ell(1+o(1))$, we have that the coefficient of $\varepsilon^{4/3}$ is $900(1+o(1))\cdot 2^{-3\ell}$, and we are done.

Finally, for the fifth item, denote $G = F_x$. Then for any $y \in M_x$,

$$\begin{aligned} h_x(y) &= \mu_{\mathsf{Span}(x,y),\mathsf{in}}(F) - \mu_{\mathsf{Span}(x),\mathsf{in}}(F) + \delta - \mu_{\mathsf{Span}(y),\mathsf{in}}(F) + \delta - \mu_{\mathsf{Span}(x+y),\mathsf{in}}(F)) \\ &= g_{\approx 1}(y) - f_{\approx 1}(y) - f_{\approx 1}(x+y). \end{aligned}$$

Therefore for any S_1 character on M_x,

$$\widehat{h_x}(S_1) = \widehat{g}_{\approx 1}(S_1) + - \mathop{\mathbb{E}}_{y\in W_{S_1}}[f_{\approx 1}(y)\chi_{S_1}(y)] - \mathop{\mathbb{E}}_{y\in W_{S_1}}[f_{\approx 1}(x+y)\chi_{S_1}(y)]. \tag{14}$$

We extend S_1 to $\tilde{S}_1$ on V so that $\chi_{\tilde{S}_1}(x) = 1$. Then

$$\mathop{\mathbb{E}}_{y\in W_{S_1}}[f_{\approx 1}(y)\chi_{S_1}(y)] + \mathop{\mathbb{E}}_{y\in W_{S_1}}[f_{\approx 1}(x+y)\chi_{S_1}(y)] = 2\mathop{\mathbb{E}}_{y\in V}\left[f_{\approx 1}(y)\chi_{\tilde{S}_1}(y)\right] = 2\widehat{f}_{\approx 1}(\tilde{S}_1).$$

Taking absolute value and using triangle inequality on equation (14), we get that

$$\left|\widehat{h_x}(S_1)\right| \leqslant |\widehat{g}_{\approx 1}(S_1)| + 2\left|\widehat{f}_{\approx 1}(\tilde{S}_1)\right|.$$

Finally we bound the first fourier coefficient on the right hand side by $2^{-\ell}(2+o(1))\varepsilon$ using the $(1,\delta+\varepsilon)$ pseudo-randomness of G and the second by $2^{-\ell}(1+o(1))\varepsilon$ using the $(1,\varepsilon)$-pseudo randomness of F (Both by Claim 7.8).We demonstrate for the first (the second is done similarly). By Lemma 4.2

$$|\widehat{g}_{\approx 1}(S_1)| = 2^{-\ell}(1+o(1))\left|\mu_{W_{S_1},\mathsf{out}}(G) - \mu(G)\right| \leqslant 2^{-\ell}(1+o(1))2\varepsilon.$$

□

Claim 7.10.

$$\|f_{\approx 2}\|_2^2 = \frac{\eta}{\begin{bmatrix}\ell\\2\end{bmatrix}} + O(2^{16\ell^4-k}) \leqslant \frac{10\eta}{2^{2\ell}} + O(2^{16\ell^4-k}).$$

Proof. Immediate by Lemma 5.1. □

7.3 Analysis of representative cases

In this section we unravel the fourth moment of $F_{\approx 2}$, and show how to group the different sums into relatively small number of classes. We then analyze some representatives of these sums in order to demonstrate the different arguments that are needed.

We begin with two important definitions that will be helpful throughout the rest of the section.

Definition 7.11. *Let $K, M, P, Q \subseteq V$ be subspaces. The total dimension of a term of the form $f_2(K)f_2(M)f_2(P)f_2(Q)$ is defined to be*

$$\mathsf{dim}(K \oplus M \oplus P \oplus Q).$$

Sometimes the spaces will be given in formal form with parameters; for this purpose we define the formal total dimension:

Definition 7.12. *Let $x_1, ..., x_8$ be formal variables, and let K, M, P, Q be subspaces of the formal subspace $\mathsf{Span}\{x_1, ..., x_8\}$. The formal total dimension of a term of the form $f_2(K)f_2(M)f_2(P)f_2(Q)$ is defined to be*

$$\mathsf{dim}(K \oplus M \oplus P \oplus Q),$$

when $K \oplus M \oplus P \oplus Q$ is treated as formal subspace of $\mathsf{Span}(x_1, \ldots, x_8)$.

The following lemma asserts that the fourth moment of $F_{\approx 2}$ can be written as linear combination of expectations of formal total dimension $d = 2, ..., 8$, where the coefficient for d is $\Theta_d(2^{d\ell})$.

Lemma 7.13. *There are $a_d(\ell)$ for $d = 0, ..., 8$ such that*

$$\mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\\ell\end{bmatrix}}\left[F_{\approx 2}[L]^4\right] = \sum_{d=0}^{8} a_d(\ell) \mathop{\mathbb{E}}_{\substack{K,M,P,Q\in\begin{bmatrix}V\\2\end{bmatrix}\\ \mathsf{dim}(K\oplus M\oplus P\oplus Q)=d}} [f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)].$$

Moreover, $a_d(\ell) = 0$ for $d = 0, 1$ and $|a_d(\ell)| \leqslant 2^{4d+d\ell}$ for any d.

Proof. By definition,

$$\mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}\left[F_{\approx 2}[L]^4\right] = \mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}\left[\left(\sum_{K\subseteq\begin{bmatrix}L\\ 2\end{bmatrix}} f_{\approx 2}(K)\right)^4\right]$$
$$= \mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}\left[\sum_{K,M,P,Q\subseteq\begin{bmatrix}L\\ 2\end{bmatrix}} f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)\right]$$

We partition the last sum according to $A = K \oplus M \oplus P \oplus Q$, which has dimension at least 2, to get

$$= \mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}\left[\sum_{d=2}^{8}\sum_{\substack{A\subseteq L\\ \dim(A)=d}}\sum_{\substack{K,M,P,Q\in\begin{bmatrix}A\\ 2\end{bmatrix}\\ K\oplus M\oplus P\oplus Q=A}} f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)\right]$$

Note that by symmetry, the size of the set

$$\{(K,M,P,Q) \mid \text{ all are subspaces, } K\oplus M\oplus P\oplus Q = A\}$$

depends only on $d = \dim(A)$; denote it by β_d. Then the last expectation is

$$= \mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}\left[\sum_{d=2}^{8}\begin{bmatrix}\ell\\ d\end{bmatrix}\mathop{\mathbb{E}}_{\substack{A\subseteq L\\ \dim(A)=d}}\left[\beta_d \mathop{\mathbb{E}}_{\substack{K,M,P,Q\in\begin{bmatrix}A\\ 2\end{bmatrix}\\ K\oplus M\oplus P\oplus Q=A}}[f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)]\right]\right]$$
$$= \sum_{d=2}^{8}\begin{bmatrix}\ell\\ d\end{bmatrix}\beta_d \mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}\left[\mathop{\mathbb{E}}_{\substack{A\subseteq L\\ \dim(A)=d}}\left[\mathop{\mathbb{E}}_{\substack{K,M,P,Q\in\begin{bmatrix}A\\ 2\end{bmatrix}\\ K\oplus M\oplus P\oplus Q=A}}[f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)]\right]\right],$$

where the last equality is by linearity of expectation. Note that for a fixed d, the 4-tuple K, M, P, Q is distributed uniformly over all 4-tuples of 2-dimensional subspaces of V such that their direct sum is of dimension d. Thus the last sum equals

$$\sum_{d=2}^{8}\begin{bmatrix}\ell\\ d\end{bmatrix}\beta_d \mathop{\mathbb{E}}_{\substack{K,M,P,Q\in\begin{bmatrix}V\\ 2\end{bmatrix}\\ \dim(K\oplus M\oplus P\oplus Q)=d}}[f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)].$$

Denoting $a_d(\ell) = \begin{bmatrix}\ell\\ d\end{bmatrix}\beta_d$ for $d \geqslant 2$ and 0 otherwise we conclude the lemma. The bound on $a_d(\ell)$ follows from the crude bounds $\begin{bmatrix}\ell\\ d\end{bmatrix} \leqslant 2^{d\ell}$ and $\beta_d \leqslant 2^{4d}$ which is easy to verify. □

Below we consider different types of sums and show how to obtain sufficient upper bounds on them. We write them in expectation notations since it is easier to work with.

7.3.1 Expectations that are essentially 0

The first type of expectations are those that can be argued to be essentially 0. The following observation is sufficient for us to identify such expectations. Roughly speaking, it asserts that if one of the spaces depends on a random vector y such that the direct sum of all 3 other spaces does not contain it, then the expectation is close to 0.

Lemma 7.14. *Let $x_1, ..., x_d$ be formal random vectors from V, and K, M, P, Q be formal vector spaces spanned by them. Define an event*

$$E = \{\, (K, M, P, Q) \mid \mathsf{dim}(K \oplus M \oplus P) \leqslant d-1, \mathsf{dim}(K \oplus M \oplus P \oplus Q) = d \}\,.$$

Then

$$2^{d\ell} \mathop{\mathbb{E}}_{\substack{x_1,\dots,x_d \in V \\ (K,M,P,Q)\in E}} [f_{\approx 2}(K) f_{\approx 2}(M) f_{\approx 2}(P) f_{\approx 2}(Q)] = O(2^{d\ell - k} \|f_{\approx 2}\|_\infty^4).$$

Proof.

$$\begin{aligned} &2^{d\ell} \mathop{\mathbb{E}}_{\substack{x_1,\dots,x_d \in V \\ (K,M,P,Q)\in E}} [f_{\approx 2}(K) f_{\approx 2}(M) f_{\approx 2}(P) f_{\approx 2}(Q)] \\ &= 2^{d\ell} \mathop{\mathbb{E}}_{\substack{x_1,\dots,x_d \in V \\ (K,M,P,Q)\in E}} \left[f_{\approx 2}(K) f_{\approx 2}(M) f_{\approx 2}(P) \mathop{\mathbb{E}}_{Q} \left[f_{\approx 2}(Q) \middle| K, M, P \right] \right] \end{aligned}$$

For any value of K, M, P, the induced distribution on $Q = \mathsf{Span}\{x, y\}$ is such that x has some distribution Λ: either $x \in K \oplus M \oplus P$ if $\mathsf{dim}(K \oplus M \oplus P) = d-1$, else uniform outside $K \oplus M \oplus P$. In both cases, y is uniform outside $K \oplus M \oplus P$. Thus

$$\mathop{\mathbb{E}}_{Q} \left[f_{\approx 2}(Q) \middle| K, M, P \right] = \mathop{\mathbb{E}}_{x \sim \Lambda} \left[\mathop{\mathbb{E}}_{y \notin K \oplus M \oplus P} [f_{\approx 2}(\mathsf{Span}\{x, y\})] \right].$$

For any fixed x the resulting distribution on y is $O(2^{-k})$ close to uniform over $y \neq x, 0$. Thus

$$\mathop{\mathbb{E}}_{y \notin K \oplus M \oplus P} [f_{\approx 2}(\mathsf{Span}\{x, y\})] = \mathop{\mathbb{E}}_{K \ni x} [f_{\approx 2}(\mathsf{Span}\{x, y\})] + O(2^{-k} \|f_{\approx 2}\|_\infty) = O(2^{-k} \|f_{\approx 2}\|_\infty).$$

□

7.3.2 Error terms

The second type of expectations are those that we bound by $O(\eta^2)$ using Cauchy-Schwartz inequality. An example is

$$\begin{bmatrix} \ell \\ 1 \end{bmatrix}^4 \mathop{\mathbb{E}}_{x_1,x_2,x_3,x_4,x_5,x_6} \left[f_{\approx 2}(\mathsf{Span}\{x_1, x_4\}) f_{\approx 2}(\mathsf{Span}\{x_2, x_5\}) f_{\approx 2}(\mathsf{Span}\{x_1 + x_2, x_4 + x_5\})^2 \right].$$

More generally,

Lemma 7.15. *Let $x_1, ..., x_4$ be formal random vectors from V, and K, M, P, Q be formal vector spaces spanned by them. Define an event*

$$E = \{ (K, M, P, Q) \mid K \cap M = P \cap Q = \{0\}, \mathsf{dim}(K \oplus M \oplus P \oplus Q) = 4\}$$

Then

$$2^{4\ell} \left| \mathop{\mathbb{E}}_{\substack{x_1,...,x_4 \in V \\ K,M,P,Q \in E}} [f_{\approx 2}(K), f_{\approx 2}(M), f_{\approx 2}(P), f_{\approx 2}(Q)] \right| \leqslant 100\eta^2 + O(2^{16\ell^4 - k}).$$

Proof. Applying Cauchy-Schwartz

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{\substack{x_1,...,x_4 \in V \\ K,M,P,Q \in E}} [f_{\approx 2}(K), f_{\approx 2}(M), f_{\approx 2}(P), f_{\approx 2}(Q)] \right| \\ &\leqslant \sqrt{\mathop{\mathbb{E}}_{\substack{x_1,...,x_4 \in V \\ K,M,P,Q \in E}} \left[{f_{\approx 2}}^2(K) {f_{\approx 2}}^2(M) \right]} \sqrt{\mathop{\mathbb{E}}_{\substack{x_1,...,x_4 \in V \\ K,M,P,Q \in E}} \left[{f_{\approx 2}}^2(P) {f_{\approx 2}}^2(Q) \right]} \\ &= \mathop{\mathbb{E}}_{\substack{x_1,...,x_4 \in V \\ K,M,P,Q \in E}} \left[{f_{\approx 2}}^2(K) {f_{\approx 2}}^2(M) \right], \end{aligned} \tag{15}$$

the last equality is because the distributions of $(K, M), (P, Q)$ are identical. Note that since $K \cap M = \{0\}$, the distribution of K, M is $O(2^{-k})$ close to choosing two independent 2-dimensional subspaces of V. Thus the last expression in 15 is

$$= \mathop{\mathbb{E}}_{K,M \in \binom{V}{2}} \left[{f_{\approx 2}}^2(K) {f_{\approx 2}}^2(M) \right] + O(2^{-k} \|f_{\approx 2}\|_\infty) = \|f_{\approx 2}\|_2^4 + + O(2^{-k} \|f_{\approx 2}\|_\infty).$$

Finally, $\|f_{\approx 2}\|_2^4 \leqslant \frac{100\eta^2}{2^{4\ell}} + O(2^{16\ell^4 - k})$ by Claim 7.10. □

7.3.3 Zoom-in 2 dimensions.

The only expression of this type is

$$2^{2\ell} \mathop{\mathbb{E}}_{K} \left[{f_{\approx 2}}^4(K) \right] \leqslant 2^{2\ell} \|f_{\approx 2}\|_2^2 \|f_{\approx 2}\|_\infty^2 \leqslant (10\eta) \cdot (4\varepsilon)^2 + O(2^{2\ell + 16\ell^4 - k}) = 160\eta\varepsilon^2 + O(2^{2\ell + 16\ell^4 - k}).$$

the last equality is by Claim 7.10 and Claim 7.9.

7.3.4 Reducing to first level via zoom in

There are some that can be handled by reducing to the first level. That is, to analyze it one has to rely on the fact that after zooming-in (or out) on one dimension, the resulting function is pseudo-random against zoom-in (or out) to dimension one. This, combined with Corollary 3.4 provides us an upper-bound on the level-1 mass of the zoomed-in function. To demonstrate the idea more concretely, consider the expression

$$2^{3\ell} \mathop{\mathbb{E}}_{x} \left[\mathop{\mathbb{E}}_{x \in K,M} \left[{f_{\approx 2}}^2(K) {f_{\approx 2}}^2(M) \right] \right]. \tag{16}$$

Define $g(x) = \mathbb{E}_{K \ni x}\left[{f_{\approx 2}}^2(K)\right]$. Then $\mathbb{E}_x\left[g(x)\right] = \mathbb{E}_K\left[{f_{\approx 2}}^2(K)\right] = \|f_{\approx 2}\|_2^2$. Therefore

$$\mathop{\mathbb{E}}_{x}\left[\mathop{\mathbb{E}}_{x \in K,M}\left[{f_{\approx 2}}^2(K) f_{\approx 2}{}^2(M)\right]\right] = \mathop{\mathbb{E}}_{x}\left[g^2(x)\right] \leqslant \|g\|_\infty \mathop{\mathbb{E}}_{x}[g(x)] = \|g\|_\infty \|f_{\approx 2}\|_2^2. \tag{17}$$

Let x^* be such $|g(x^*)| = \|g\|_\infty$. We have

$$\begin{aligned}
|g(x^*)| &= \mathop{\mathbb{E}}_{y \neq x^*,0}\left[{f_{\approx 2}}^2(\mathsf{Span}\{x^*, y\})\right] \\
&= \mathop{\mathbb{E}}_{y \neq x^*,0}\left[\left(\mu_{\mathsf{Span}\{x^*,y\},\mathsf{in}}(F) - \mu_{\mathsf{Span}\{x^*\},\mathsf{in}(F)} + \mu(F) - \mu_{\mathsf{Span}\{y\},\mathsf{in}}(F) + \mu(F) - \mu_{\mathsf{Span}\{x^*+y\},\mathsf{in}}(F)\right)^2\right] \\
&\leqslant 3 \mathop{\mathbb{E}}_{y \neq x^*,0}\left[(\mu_{\mathsf{Span}\{x^*,y\},\mathsf{in}}(F) - \mu_{\mathsf{Span}\{x^*\},\mathsf{in}}(F))^2 + (\mu_{\mathsf{Span}\{y\},\mathsf{in}}(F) - \mu(F))^2 + (\mu_{\mathsf{Span}\{x^*+y\},\mathsf{in}}(F) - \mu(F))^2,\right]
\end{aligned}$$

in the last inequality we used $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$. Note that $r(y) \stackrel{def}{=} \mu_{\mathsf{Span}\{x^*,y\},\mathsf{in}} - \mu_{\mathsf{Span}\{x^*\},\mathsf{in}}$ is the approximated point function of F_{x^*}. By Claim 6.2,

$$\mathop{\mathbb{E}}_{y \neq x^*,0}\left[r^2(y)\right] \leqslant \frac{1}{\begin{bmatrix}\ell\\1\end{bmatrix}} W^{=1}[F_{x^*}] + O(2^{16\ell^4 - k}\|F\|_\infty).$$

Similarly, $\mu_{\mathsf{Span}\{y\},\mathsf{in}} - \mu(F)$ is the approximated point function of F; since the distribution of y is $O(2^{-k})$ close to uniform, we have that:

$$\mathop{\mathbb{E}}_{y \neq x,0}\left[(\mu_{\mathsf{Span}\{y\},\mathsf{in}} - \mu(F))^2\right] \leqslant \frac{1}{\begin{bmatrix}\ell\\1\end{bmatrix}} W^{=1}[F] + O(2^{-k}\|F\|_\infty).$$

Therefore

$$\|g\|_\infty = |g(x^*)| \leqslant \frac{1}{\begin{bmatrix}\ell\\1\end{bmatrix}}(W^{=1}[F_{x^*}] + 2 \cdot W^{=1}[F]) + O(2^{16\ell^4 - k}).$$

We next upper bound the weight of F, F_{x^*} on their first level. By Claim 7.8, F is $(1, \varepsilon)$ pseudo-random and thus $(1, \frac{9}{2}\delta^{2/3})$ pseudo-random, and hence by Corollary 3.4 $W^{=1}[F]\sqrt{200 \cdot \frac{9}{2}\delta^{2/3}}\delta \leqslant 30\varepsilon^{4/3}$. Similarly, F_{x^*} has density at most 2ε and is $(1, \varepsilon + \delta)$ pseudo-random and in particular $(1, 9\varepsilon^{2/3})$ pseudo-random and therefore by Corollary 3.4 $W^{=1}[F_{x^*}] \leqslant \sqrt{200 \cdot 9\varepsilon^{2/3}}\mu(F_{x^*}) \leqslant 60\varepsilon^{4/3}$.

We thus obtain

$$\|g\|_\infty \leqslant \frac{9}{\begin{bmatrix}\ell\\1\end{bmatrix}}\varepsilon^{4/3} + O(2^{-k}) = 90(1 + o(1))2^{-\ell}\varepsilon^{4/3} + O(2^{16\ell^4 - k}).$$

Plugging this in Equation (17) and using Claim7.10 to estimate $\|f_{\approx 2}\|_2^2$, we get the expression in (16) is bounded above by

$$2^{3\ell}\left(90(1 + o(1))2^{-\ell}\varepsilon^{4/3} + O(2^{16\ell^4 - k})\right)\left(\frac{10\eta}{2^{2\ell}} + O(2^{16\ell^4 - k})\right) = 900(1 + o(1))\varepsilon^{4/3}\eta + O(2^{3\ell + 16\ell^4 - k}).$$

7.3.5 Ones requiring zoom-outs

An example is

$$2^{6\ell} \mathop{\mathbb{E}}_{x_1,x_2,x_3,x_4,x_5,x_6} [f_{\approx 2}(\mathsf{Span}\{x_1,x_4\}) f_{\approx 2}(\mathsf{Span}\{x_2,x_5\}) f_{\approx 2}(\mathsf{Span}\{x_3,x_6\}) f_{\approx 2}(\mathsf{Span}\{x_1+x_2+x_3, x_4+x_5+x_6\})].$$

A classical fourier-analytic computation shows that the above equals

$$2^{6\ell} \sum_{S_1,S_2} \widehat{f_{\approx 2}}^4(S_1,S_2).$$

Clearly by Parseval

$$\sum_{S_1,S_2} \widehat{f_{\approx 2}}^4(S_1,S_2) \leqslant \|\widehat{f_{\approx 2}}\|_\infty^2 \sum_{S_1,S_2} \widehat{f_{\approx 2}}^2(S_1,S_2) = \|\widehat{f_{\approx 2}}\|_\infty^2 \mathbb{E}\left[{f_{\approx 2}}^2\right].$$

Use Claims 7.9, 6.2 to get that the above is at most

$$2^{6\ell}\left(16(1+o(1))2^{-4\ell}\varepsilon^2 + O(2^{\ell-k})\right)\left(\frac{10\eta}{2^{2\ell}} + O(2^{16\ell^4-k})\right) = 160(1+o(1))\varepsilon^2\eta + O(2^{6\ell+16\ell^4-k}).$$

7.3.6 Reduce to first level via zoom-out

An example is

$$2^{5\ell} \mathop{\mathbb{E}}_{x_1,x_2,x_4,x_5,x_6} [f_{\approx 2}(\mathsf{Span}\{x_1,x_4\}) f_{\approx 2}(\mathsf{Span}\{x_1,x_5\}) f_{\approx 2}(\mathsf{Span}\{x_2,x_6\}) f_{\approx 2}(\mathsf{Span}\{x_2,x_4+x_5+x_6\})].$$

A standard fourier-analytic computation shows the above is

$$2^{5\ell} \sum_{S_1} \left(\sum_{S_2} \widehat{f_{\approx 2}}^2(S_1,S_2) \right)^2.$$

Clearly,

$$\begin{aligned}\sum_{S_1} \left(\sum_{S_2} \widehat{f_{\approx 2}}^2(S_1,S_2) \right)^2 &\leqslant \max_{S_2}\left\{ \sum_{S_2} \widehat{f_{\approx 2}}^2(S_1,S_2)\right\} \sum_{S_1}\sum_{S_2} \widehat{f_{\approx 2}}^2(S_1,S_2) \\ &= \max_{S_2}\left\{ \sum_{S_2} \widehat{f_{\approx 2}}^2(S_1,S_2)\right\} \|f_{\approx 2}\|_2^2,\end{aligned}$$

the last equality is by Parseval. Use Claims 7.9, 6.2 to get that the above is at most

$$2^{5\ell}\left(900(1+o(1))2^{-3\ell}\varepsilon^{4/3} + O(2^{16\ell^4-k})\right)\left(\frac{10\eta}{2^{2\ell}} + O(2^{16\ell^4-k})\right) = 9,000(1+o(1))\varepsilon^{4/3}\eta + O(2^{5\ell+16\ell^4-k}).$$

Expressions requiring combination of methods

Unfortunately there are two more arguments that we must present for the proof of Theorem 3.7 that require thinking about zooming in and zooming out simultaneously.

7.3.7 Zoom-in, Zoom-out Combination - Part 1

Consider the expectation

$$2^{4\ell} \mathop{\mathbb{E}}_{x,y,z,w} [f_{\approx 2}(\mathsf{Span}\{w,x\}) f_{\approx 2}(\mathsf{Span}\{w,y\}) f_{\approx 2}(\mathsf{Span}\{w,z\}) f_{\approx 2}(\mathsf{Span}\{w,x+y+z\})]. \tag{18}$$

Denote $g_w \colon V \to \mathbb{R}$, $g_w(x) = f(\mathsf{Span}\{w,x\})$. Then the above expression is

$$2^{4\ell} \mathop{\mathbb{E}}_{x,y,z,w} [g_w(x) g_w(y) g_w(z) g_w(x+y+z)] = 2^{4\ell} \mathop{\mathbb{E}}_{w} \left[\sum_S \widehat{g_w}^4(S) \right].$$

Note that $\mathbb{E}_w \left[\sum_S \widehat{g_w}^2(S) \right] = \mathbb{E}_w \left[\|g_w\|_2^2 \right] = \|f_{\approx 2}\|_2^2$, and therefore the last term is at most

$$\leqslant 2^{4\ell} \mathop{\mathbb{E}}_{w} \left[\sum_S \widehat{g_w}^2(S) \right] \|\widehat{g_w}\|_\infty^2 = 2^{4\ell} \|f_{\approx 2}\|_2^2 \|\widehat{g_w}\|_\infty^2$$

Use Claims 7.9, 6.2 to get that the expression in (19) is at most

$$2^{4\ell} \left(\frac{10\eta}{2^{2\ell}} + O(2^{16\ell^4 - k}) \right) \left((4 + o(1)) 2^{-2\ell} \varepsilon^2 + O(2^{16\ell^4 - k}) \right) = 40(1+o(1))\varepsilon^2 \eta + O(2^{4\ell + 16\ell^4 - k}).$$

7.3.8 Zoom-in, Zoom-out Combination - Part 2

The last type of expectation we shall examine is the following. In the previous term (Section 7.3.7), it was quite apparent that some combination of fourier argument along with "zoom-in" argument is needed (since one point - namely w, appeared in all spaces (suggesting zooming-in ideas), and on the rest corresponded to linearity testing - $x, y, z, x+y+z$). While in the below expectation it is much less apparent, we show that this combination can still be used to estimate it. We are currently unaware of a different analysis.

$$2^{4\ell} \mathop{\mathbb{E}}_{x,y,z,w} [f_{\approx 2}(\mathsf{Span}\{w,x\}) f_{\approx 2}(\mathsf{Span}\{w,y\}) f_{\approx 2}(\mathsf{Span}\{w,z\}) f_{\approx 2}(\mathsf{Span}\{x,y+z\})]. \tag{19}$$

Define $g_w(x) = f_{\approx 2}(w,x)$. Then the expression in (19) equals

$$\begin{aligned}
&2^{4\ell} \mathop{\mathbb{E}}_{x,y,z,w} [g_w(x) g_w(y) g_w(z) f_{\approx 2}(x, y+z)] \\
&= 2^{4\ell} \mathop{\mathbb{E}}_{x,y,z,w} \left[\sum_{S_1,S_2,S_3,S_4,S_5} \widehat{g_w}(S_1)\chi_{S_1}(x) \widehat{g_w}(S_2)\chi_{S_2}(y) \widehat{g_w}(S_3)\chi_{S_3}(z) \widehat{f_{\approx 2}}(S_4,S_5)\chi_{S_4}(x)\chi_{S_5}(y+z) \right] \\
&= 2^{4\ell} \mathop{\mathbb{E}}_{w} \left[\sum_{S_1,S_2} \widehat{g_w}(S_1) \widehat{g_w}^2(S_2) \widehat{f_{\approx 2}}(S_1,S_2) \right]
\end{aligned}$$

Apply Cauchy-Schwartz to upper bound the last expression by

$$\begin{aligned}
&\leqslant 2^{4\ell}\mathop{\mathbb{E}}_{w}\left[\sqrt{\sum_{S_1,S_2}\widehat{g_w}^2(S_1)\widehat{g_w}^4(S_2)}\sqrt{\sum_{S_1,S_2}\widehat{f_{\approx 2}}^2(S_1,S_2)}\right]\\
&\leqslant \binom{\ell}{1}^4\max_{w,S}|\widehat{g_w}(S)|\mathop{\mathbb{E}}_{w}\left[\sqrt{\sum_{S_1,S_2}\widehat{g_w}^2(S_1)\widehat{g_w}^2(S_2)}\|f_{\approx 2}\|_2\right]\\
&= 2^{4\ell}\max_{w,S}|\widehat{g_w}(S)|\|f\|_2\mathop{\mathbb{E}}_{w}\left[\|g_w\|_2^2\right].
\end{aligned}$$

Note that the last expectation is the expectation of ${f_{\approx 2}}^2$, namely $\|f_{\approx 2}\|_2^2$. Overall we have that the expression in (19) is at most

$$2^{4\ell}\|\widehat{g_{w^*}}\|_\infty\|f_{\approx 2}\|_2^3.$$

Use Claims 7.9, 6.2 to upper bound the last expression by

$$2^{4\ell}\left((4+o(1))2^{-\ell}\varepsilon+O(2^{16\ell^4-k})\right)\left(\frac{10\eta}{2^{2\ell}}+O(2^{16\ell^4-k})\right)^{1.5}\leqslant 160(1+o(1))\varepsilon\eta^{1.5}+O(2^{4\ell+16\ell^4-k}).$$

7.4 Summary - the above cover all arising terms

We shall now examine carefully all type of expressions that arise when evaluating

$$\mathop{\mathbb{E}}_{L}\left[\left(\sum_{K\subseteq L}f_{\approx 2}(L_2)\right)^4\right]$$

by opening the parenthesis. As stated earlier, each type of sum corresponds to the value of $f_{\approx 2}$ on four 2-dimensional subspaces K, M, P, Q. We shall divide the expressions according to the total dimension defined earlier, namely

$$\mathsf{dim}(K\oplus M\oplus P\oplus Q).$$

Throughout this section, we shall ignore terms of the order $O(2^{O(\ell)-k})$ that arise as negligible error terms.

7.4.1 Total dimension 2

The only summand of this type is $2^{2\ell}\mathbb{E}_{L_2}\left[f^4(L_2)\right]$, which is bounded in Section 7.3.3 by $160\eta\varepsilon^2$.

7.4.2 Total dimension 3

We can think of summands with total dimension 3 in the following way. First, we pick L_3 a 3-dimensional subspace, and then we pick subspaces L_2, L_2', L_2'', L_2''' whose direct sum is L_3. Note that not all subspaces can be identical, as otherwise we would have total dimension 2 Therefore, such sums can be written as

$$2^{3\ell}\mathop{\mathbb{E}}_{\substack{x,y,z\in V\\ L_2,L_2'\subseteq\mathsf{Span}\{x,y,z\}}}\left[f_{\approx 2}(\mathsf{Span}\{x,y\})f_{\approx 2}(\mathsf{Span}\{x,z\})f_{\approx 2}(L_2)f_{\approx 2}(L_2')\right].$$

There are actually two subcases for this sum. Either one of the two dimensional subspaces repeats three times, in which case the sum can be written in the form

$$2^{3\ell} \mathop{\mathbb{E}}_{x,y,z\in V} \left[{f_{\approx 2}}^3(\mathsf{Span}\{x,y\}) f_{\approx 2}(\mathsf{Span}\{x,z\}) \right] = O(2^{3\ell-k}),$$

where the equality is by Lemma 7.14. Else, each subspaces repeats at most twice. In this case we can partition the 4-spaces into 2 pairs such that of distinct subspaces. Notice that the spaces in each such pair must have intersection of dimension exactly 1. Assume $\{L_2, L_2'\}, \{L_2'', L_2'''\}$ is such partition. Then we have by Cauchy-Schwartz

$$\begin{aligned} 2^{3\ell}\mathbb{E}\left[f_{\approx 2}(L_2) f_{\approx 2}(L_2') f_{\approx 2}(L_2'') f_{\approx 2}(L_2'') \right] &\leqslant 2^{3\ell} \sqrt{\mathbb{E}\left[{f_{\approx 2}}^2(L_2) {f_{\approx 2}}^2(L_2') \right]} \sqrt{\mathbb{E}\left[{f_{\approx 2}}^2(L_2'') {f_{\approx 2}}^2(L_2''') \right]} \\ &= 2^{3\ell} \mathop{\mathbb{E}}_{x\in V} \left[\mathop{\mathbb{E}}_{x\in L_2, L_2'} \left[{f_{\approx 2}}^2(L_2) {f_{\approx 2}}^2(L_2') \right] \right]. \end{aligned}$$

The last term was bounded in Section 7.3.4 by $900(1+o(1))\varepsilon^{4/3}\eta$.

Total dimension higher than 4

For the rest of this section, we shall deal with expectations with total dimension at least 4. It turns out that expectation with total dimensions higher than 4 can be treated quite easily once terms with total dimension less than 4 have been treated by a duality trick which we present below. Suppose we have total dimension d. Roughly speaking, the idea is to view such expectations in the fourier spectrum, where their total dimension becomes $8-d$. In this case we simply adapt one of the previous arguments from the "less than 4 total dimension" cases. Formally, this trick is encapsulated in Lemma 7.16, which is in the spirit of Lemma 6.6 with essentially identical proofs.

Finally in the end of this section, we deal with expectations with total dimension exactly 4. Those turn out to require a more detailed case-analysis type examination.

Lemma 7.16. *Let $g\colon V^r \to \mathbb{R}$ and let M be an $(mr)\times d$ matrix of rank d. For vectors $x_1,...,x_d \in V$, consider the mr vectors $y_1,...,y_{mr}$ defined by $\vec{y} = M\vec{x}$ (i.e. $y_j = M[j,1]x_1 + ... + M[j,d]x_d$ for all j). Define $A_j = \{\, i \mid M[i,j]=1\}$ for $j=1,...,d$. Then*

$$\mathop{\mathbb{E}}_{\substack{x_1,...,x_d\in_R V\\ \vec{y}=M\vec{x}}} \left[\prod_{i=0}^{m-1} g(y_{ir+1},\ldots,y_{(i+1)r}) \right] = \sum_{\substack{S_1,...,S_{mr}\\ \forall j \bigoplus_{i\in A_j} S_i = \emptyset}} \prod_{i=0}^{m-1} \widehat{g}(S_{ir+1},\ldots,s_{(i+1)r})$$

Proof. Expanding each term according to its fourier expansion, we get that the left hand side equals

$$\begin{aligned} &\mathop{\mathbb{E}}_{\substack{x_1,...,x_d\in_R V\\ \vec{y}=M\vec{x}}} \left[\prod_{i=0}^{m-1} \left[\sum_{S_{ir+1},\ldots,S_{(i+1)r}} \widehat{g}(S_{ir+1},\ldots,S_{(i+1)r}) \prod_{j=1}^{r} \chi_{S_{ir+j}}(y_{ir+j}) \right] \right] \\ &= \mathop{\mathbb{E}}_{\substack{x_1,...,x_d\in_R V\\ \vec{y}=M\vec{x}}} \left[\sum_{S_1,\ldots,S_{mr}} \prod_{i=0}^{m-1} \widehat{g}(S_{ir+1},\ldots,S_{(i+1)r}) \prod_{j=1}^{r} \chi_{S_{ir+j}}(y_{ir+j}) \right] \\ &= \sum_{S_1,\ldots,S_{mr}} \prod_{i=0}^{m-1} \widehat{g}(S_{ir+1},\ldots,S_{(i+1)r}) \mathop{\mathbb{E}}_{\substack{x_1,...,x_d\in_R V\\ \vec{y}=M\vec{x}}} \left[\prod_{j=1}^{mr} \chi_{S_j}(y_j) \right], \end{aligned}$$

the last equality is by linearity of expectation. Consider the product of characters inside the expectation in the expression above. Recalling that $\vec{y} = M\vec{x}$, we have by multiplicativity of the characters and the definition of A that

$$\prod_{j=1}^{mr} \chi_{S_j}(y_j) = \prod_{j=1}^{d} \chi_{\bigoplus_{i \in A_j} S_i}(x_j),$$

since the y_i's that contribute a factor of $\chi_{S_i}(x_j)$ are precisely $i \in A_j$. Thus, the expectation is non-zero if and only if for all j, $\bigoplus_{i \in A_j} S_i = \emptyset$, in which case it is 1, finishing the proof. □

The main point of the above lemma is that a product that has total dimension d out of potential mr translates into a product with $mr - d$ total dimension in the fourier domain, since there are d independent linear equations in the right hand side of Lemma 7.16 [9]

7.4.3 Total dimension 5

We can expectations of this type by taking $x_1, ..., x_5 \in V$ formally, considering $y_1, ..., y_8 \in \mathsf{Span}(\{x_1, ..., x_5\})$ that have formal total dimension 5 and looking at

$$2^{5\ell}\mathbb{E}\left[f_{\approx 2}(\mathsf{Span}\{y_1, y_2\}) \cdots f_{\approx 2}(\mathsf{Span}\{y_7, y_8\})\right].$$

As discussed earlier, we would like to view this expression in the fourier domain. To do that, consider M the coefficient matrix of the y's according to the basis x. Apply Lemma 7.16 (with $m = 4, r = 2, d = 5$) to get a sum of fourier coefficients of total dimension 3 there:

$$2^{5\ell} \sum_{\substack{S_1,...,S_8 \\ \mathsf{dim}(\{S_1,...,S_8\}=3)}} \widehat{f_{\approx 2}}(S_1, S_2) \cdots \widehat{f_{\approx 2}}(S_7, S_8)$$

,we identified a set S_1 with its indicator vector). Repeating the argument in Section 7.4.2 we either get that the sum is an essentially 0 expression or

$$2^{5\ell} \sum_{S_1} \left(\sum_{S_2} \widehat{f_{\approx 2}}^2(S_1, S_2) \right)^2,$$

which we bounded in Section 7.3.6 by $9,000\varepsilon^{4/3}\eta$.

7.4.4 Total dimension 6

Moving to the fourier domain like in the previous section, we see that the total dimension on the fourier domain is 2. Therefore, it is either an essentially 0 expression (e.g. if something of the form $\widehat{f_{\approx 2}}(S_1, S_1)$) or else it is

$$\begin{bmatrix} \ell \\ 1 \end{bmatrix}^6 \sum_{S_1, S_2} \widehat{f_{\approx 2}}^4(S_1, S_2),$$

which was upper bounded in Section 7.3.5 by $160\varepsilon^2\eta$.

[9] The equations are linearly independent since they are defined by the transpose of M, that has rank d.

7.4.5 Total dimension $7, 8$

Those are essentially 0 by Lemma 7.14.

7.4.6 Total dimension 4

This case has several type of sums, which we enumerate over using symmetries. Consider the general form

$$2^{4\ell}\mathbb{E}\left[f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)\right].$$

If there is a space appearing 3 times, then the sum must be of the type

$$2^{4\ell}\mathop{\mathbb{E}}_{x,y,z,w}\left[{f_{\approx 2}}^3(\mathsf{Span}\{x,y\})f_{\approx 2}(\mathsf{Span}\{z,w\})\right],$$

which is essentially 0 by Lemma 7.14.

If there is a space appearing twice, say $K = M$, then either both P, Q intersect K on $\{0\}$, in which case we have a bound of $100\eta^2$ from Lemma 7.15. Else at least one of them, say P, intersects K on dim $\geqslant 1$. In this case $\mathsf{dim}(K \oplus P \oplus P) \leqslant 3$ and the term is essentially 0 by Lemma 7.14.

Next consider the main case, in which each space appears at most once; we further divide by looking at the point that appears in the maximal amount spaces denote this point by w, and branch according to the number of spaces it appears in. If it appears in all 4 spaces, we have an expression of the form

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4\mathop{\mathbb{E}}_{x,y,z,w}[f_{\approx 2}(\mathsf{Span}\{w,x\})f_{\approx 2}(\mathsf{Span}\{w,y\})f_{\approx 2}(\mathsf{Span}\{w,z\})f_{\approx 2}(\mathsf{Span}\{w,\cdot\})],$$

where $\cdot$ can be any linear combination of x, y, z. If it is not $x + y + z$, we have a bound of $100\eta^2$ by Lemma 7.15. If it is $x + y + z$, we get a bound of $10\varepsilon^2\eta$ by Section 7.3.7.

Let us now assume w appears in 3 spaces. Then the general form of our expression is

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4\mathop{\mathbb{E}}_{x,y,z,w}[f_{\approx 2}(\mathsf{Span}\{w,x\})f_{\approx 2}(\mathsf{Span}\{w,y\})f_{\approx 2}(\mathsf{Span}\{w,*\})f_{\approx 2}(\mathsf{Span}\{\cdot,\cdot\})],$$

If $*$ is x, y or $x + y$, we have essentially 0 by Lemma 7.14. Else it is an independent point, and our general form is

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4\mathop{\mathbb{E}}_{x,y,z,w}[f_{\approx 2}(\mathsf{Span}\{w,x\})f_{\approx 2}(\mathsf{Span}\{w,y\})f_{\approx 2}(\mathsf{Span}\{w,z\})f_{\approx 2}(\mathsf{Span}\{*,\cdot\})].$$

There are several cases here, but they are all the same up to switching names [10] ; they are all equivalent to:

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4\mathop{\mathbb{E}}_{x,y,z,w}[f(\mathsf{Span}\{w,x\})f(\mathsf{Span}\{w,y\})f(\mathsf{Span}\{w,z\})f(\mathsf{Span}\{x,y+z\})].$$

which was bounded by $160\varepsilon\eta^{1.5}$ in Section 7.3.8.

[10] The following pairs are name-interchangeable: $(x, w + x), (y, w + y), (z, w + z)$.

Finally we consider the case in which w appears in two spaces (this is the last case, since if every two spaces intersect in $\{0\}$ we have a bound of $100\eta^2$ by Lemma 7.15). Our general form is

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4 \mathop{\mathbb{E}}_{x,y,z,w}[f_{\approx 2}(\mathsf{Span}\{w,x\})f_{\approx 2}(\mathsf{Span}\{w,y\})f_{\approx 2}(\mathsf{Span}\{*,*\})f_{\approx 2}(\mathsf{Span}\{\cdot,\cdot\})].$$

Both of blank spaces need to depend on an independent point of w,x,y since otherwise the term is essentially 0 by Lemma 7.14; the other point must be from $\mathsf{Span}\{w,x,y\}$.

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4 \mathop{\mathbb{E}}_{x,y,z,w}[f(\mathsf{Span}\{w,x\})f(\mathsf{Span}\{w,y\})f(\mathsf{Span}\{z,*\})f(\mathsf{Span}\{\cdot,\cdot\})].$$

If $*$ is not y,x or w then the third space intersects both the first ones in $\{0\}$, and since the fourth depends on z it must intersect one of the first spaces in $\{0\}$, and we have a bound of $100\eta^2$ by Lemma 7.15. Since $*$ cannot be w, it must be x or y - let us assume by symmetry it is x:

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4 \mathop{\mathbb{E}}_{x,y,z,w}[f_{\approx 2}(\mathsf{Span}\{w,x\})f_{\approx 2}(\mathsf{Span}\{w,y\})f_{\approx 2}(\mathsf{Span}\{z,x\})f_{\approx 2}(\mathsf{Span}\{\cdot,\cdot\})].$$

If the last space intersects the first one on $\{0\}$ we have a bound of $100\eta^2$ by Lemma 7.15. Thus we may assume it contains w,x or $w+x$. Note that this point cannot be w or x, as otherwise there would be a point in three spaces:

$$\begin{bmatrix}\ell\\1\end{bmatrix}^4 \mathop{\mathbb{E}}_{x,y,z,w}[f_{\approx 2}(\mathsf{Span}\{w,x\})f_{\approx 2}(\mathsf{Span}\{w,y\})f_{\approx 2}(\mathsf{Span}\{z,x\})f_{\approx 2}(\mathsf{Span}\{*,w+x\})].$$

$*$ has to depend on z, so it can be $z, z+y, z+w+y, z+w$. The cases $*=z, z+w$ the term is essentially 0 by Lemma 7.14 (isolating the second space). In the other two cases we have a bound of $100\eta^2$ by Lemma 7.15 (grouping the first with the fourth and the second with the third).

7.5 Proof of Theorem 3.7

In this section we prove a counter-positive version of Theorem 3.7, stated below. Theorem 3.7 can be derived from it rather easily - the (rather self-evident) proof is deferred to Section E.2.

Theorem 7.17. *Let V be a k-dimensional vector space, $10 < \ell < k$ an integer and $\delta > 0$ such that $k \geqslant 27\ell^4 + 10\log\frac{1}{\delta} + 10$. Let $F\colon \begin{bmatrix}V\\\ell\end{bmatrix} \to \{0,1\}$, and assume $\mu(F) = \delta$.*

If $\varepsilon \geqslant \delta$ and F is $(2,\varepsilon)$ pseudo-random, then

$$W^{=2}[F] \leqslant 2^{17}\delta\varepsilon^{1/3}.$$

Proof. Denote $\eta = W^{=2}[F], \delta = \mu(F)$. By Lemma 6.4 we have that

$$\mathop{\mathbb{E}}_{L}\left[F_{\approx 2}[L]^4\right] \geqslant \frac{\eta^5}{2^{10}\delta^4}. \tag{20}$$

On the other hand, by Lemma 7.13,

$$\mathop{\mathbb{E}}_{L\in\begin{bmatrix}V\\\ell\end{bmatrix}}\left[F_{\approx 2}[L]^4\right] = \sum_{d=2}^{8} a_d(\ell)2^{d\ell} \mathop{\mathbb{E}}_{\substack{K,M,P,Q\in\begin{bmatrix}V\\\ell\end{bmatrix}\\ \dim(K\oplus M\oplus P\oplus Q)=d}}[f_{\approx 2}(K)f_{\approx 2}(M)f_{\approx 2}(P)f_{\approx 2}(Q)].$$

where $|a_d(\ell)| \leqslant 2^{4d}$ for all d.

We apply the bounds we got from Section 7.4 [11] to get

$$\begin{aligned}\mathop{\mathbb{E}}_{L\in\binom{V}{\ell}}\left[F_{\approx 2}[L]^4\right] &\leqslant \max_{d=2,\ldots,8} 2^{4d}(160\eta\varepsilon^2 + 900\eta\varepsilon^{4/3} + 1,200\eta\varepsilon^{1.5} + 9,000\eta\varepsilon^{4/3} + 160\eta\varepsilon^2 + 100\eta^2 + 10\eta\varepsilon^2) + o(1)\\ &\leqslant 2^{40}\eta^2 + 2^{50}\varepsilon^{4/3}\eta. \qquad (21)\end{aligned}$$

Putting (20) and (21) together we have

$$\frac{\eta^5}{2^{10}\delta^4} \leqslant 2^{40}\eta^2 + 2^{50}\varepsilon^{4/3}\eta.$$

Hence one of the terms in the right hand side must be at least half of the left hand side. If the first one is such term, rearranging yields

$$\eta \leqslant 2^{17}\delta^{4/3}.$$

Otherwise the second summand is at least half the left hand side, and rearranging implies

$$\eta \leqslant 2^{15}\delta\varepsilon^{1/3}.$$

Since $\varepsilon \geqslant \delta$, we have $\eta \leqslant 2^{17}\delta\varepsilon^{1/3}$ in either case and we are done. □

References

[ABS15] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. *J. ACM*, 62(5):42, 2015.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.

[BCN12] A.E. Brouwer, A.M. Cohen, and A. Neumaier. *Distance-Regular Graphs*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Springer Berlin Heidelberg, 2012.

[CMM06] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 205–214, 2006.

[DKK+16] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? *Electronic Colloquium on Computational Complexity (ECCC)*, 23:198, 2016.

[DS05] Irit Dinur and Samuel Safra. On the Hardness of Approximating Minimum Vertex Cover. *Annals of Mathematics*, 162(1):439–485, 2005.

[11] Strictly speaking, the expectation computed for each d is $O(2^{-k})$-close to matching expression we computed in Section 7.4 since with except probability $O(2^{-k})$, the spaces there have total dimension d.

[FGL+96] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, March 1996.

[Fil16a] Yuval Filmus. Friedgut-Kalai-Naor theorem for slices of the boolean cube. *Chicago J. Theor. Comput. Sci.*, 2016, 2016.

[Fil16b] Yuval Filmus. An orthogonal basis for functions over a slice of the boolean hypercube. *Electr. J. Comb.*, 23(1):P1.23, 2016.

[FKMW16] Yuval Filmus, Guy Kindler, Elchanan Mossel, and Karl Wimmer. Invariance principle on the slice. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 15:1–15:10, 2016.

[FM16] Yuval Filmus and Elchanan Mossel. Harmonicity and invariance on slices of the boolean cube. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 16:1–16:13, 2016.

[FW86] Peter Frankl and Richard M. Wilson. The erdös-ko-rado theorem for vector spaces. *J. Comb. Theory, Ser. A*, 43(2):228–236, 1986.

[Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[HK05] Johan Håstad and Subhash Khot. Query efficient PCPs with perfect completeness. *Theory of Computing*, 1(1):119–148, 2005.

[HW03] Johan Håstad and Avi Wigderson. Simple analysis of graph tests for linearity and PCP. *Random Struct. Algorithms*, 22(2):139–160, 2003.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002*, page 25, 2002.

[KKMO07] 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, April 2007.

[KM16] Subhash Khot and Dana Moshkovitz. Candidate hard unique game. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 63–76, 2016.

[KMS16] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games and grassmann graphs. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:124, 2016.

[Kol11] Alexandra Kolla. Spectral algorithms for unique games. *Computational Complexity*, 20(2):177–206, 2011.

[KR08] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. *J. Comput. Syst. Sci.*, 74(3):335–349, May 2008.

[KS13] Subhash Khot and Muli Safra. A two-prover one-round game with strong soundness. *Theory of Computing*, 9(28):863–887, 2013.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 245–254, New York, NY, USA, 2008. ACM.

[Rot53] Klaus Friedrich Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.

[Sri12] Murali K. Srinivasan. A positive combinatorial formula for the complexity of the q-analog of the n-cube. *Electr. J. Comb.*, 19(2):34, 2012.

[Sri14] Murali K. Srinivasan. The goldman-rota identity and the grassmann scheme. *Electr. J. Comb.*, 21(1):37, 2014.

[ST00] Alex Samorodnitsky and Luca Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, OR, USA*, pages 191–199, 2000.

[Tre98] Luca Trevisan. Recycling queries in PCPs and in linearity tests (extended abstract). In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 299–308, 1998.

[Tre08] Luca Trevisan. Approximation algorithms for unique games. *Theory of Computing*, 4(1):111–128, 2008.

Appendix

A Proof of Theorem 2.25

We begin with a crude upper bound on the functions $f_{\approx i}$ from 2.24.

Claim A.1. *For every i, $\|f_{\approx i}\|_\infty \leqslant 2^{i^3}\|F\|_\infty$.*

Proof. Induction on i. For $i = 0, 1$ this is obvious. Let $i \geqslant 2$ and assume for $j < i$, then for every $L_i \in \binom{V}{i}$,

$$\begin{aligned}|f_{\approx i}(L_i)| &= \left|\mu_{L_i,\mathsf{in}}(F) - \sum_{j=0}^{i-1}\sum_{L_j\subseteq L_i} f_{\approx j}(L_j)\right| \\ &\leqslant \|F\|_\infty + \sum_{j=0}^{i-1} 2^{i^2}\|f_{\approx j}\|_\infty.\end{aligned}$$

The last inequality is by the triangle inequality. Use the induction hypothesis to get

$$|f_{\approx i}(L_i)| \leqslant \|F\|_\infty + \sum_{j=0}^{i-1} 2^{i^2}2^{j^3}\|F\|_\infty \leqslant (1 + i2^{i^2+(i-1)^3})\|F\|_\infty \leqslant 2^{i^3}\|F\|_\infty,$$

the last inequality follows by $1 + i2^{i^2+(i-1)^3} \leqslant 2^{i^3}$ for every $i \geqslant 2$. □

Lemma A.2. *For every $L_{i-1} \in \binom{V}{i-1}$,*

$$\left|\mathop{\mathbb{E}}_{L_i\supseteq L_{i-1}}[f_{\approx i}(L_i)]\right| \leqslant 2^{10i^4-k}\|F\|_\infty.$$

Proof. For each $i = 1, ..., \ell - 1$, let ξ_i be the maximum over $L_{i-1} \in \binom{V}{i-1}$ of $\left|\mathbb{E}_{L_i\supseteq L_{i-1}}[f_{\approx i}(L_i)]\right|$. We will upper bound ξ_i inductively.

For $i = 1$, Claim 2.7 yields that $\mathbb{E}_{x\in V\setminus\{0\}}\left[\tilde{f}_1(x)\right] = 0$, i.e. $\xi_1 = 0$. Let $i > 1$, assume the statement for all $j < i$ and prove for i. Let L_{i-1} be the one obtaining ξ_i.

$$\begin{aligned}\xi_i = \left|\mathop{\mathbb{E}}_{\substack{L_i\in\binom{V}{i}\\ L_{i-1}\subseteq L_i}}[f_{\approx i}(L_i)]\right| &= \left|\mathop{\mathbb{E}}_{\substack{L_i\in\binom{V}{i}\\ L_{i-1}\subseteq L_i}}\left[\mu_{L_i,\mathsf{in}}(F) - \sum_{j=0}^{i-1}\sum_{L_j\subseteq L_i} f_{\approx j}(L_j)\right]\right| \\ &= \left|\mu_{L_{i-1},\mathsf{in}}(F) - \mathop{\mathbb{E}}_{\substack{L_i\in\binom{V}{i}\\ L_{i-1}\subseteq L_i}}\left[\sum_{j=0}^{i-1}\sum_{L_j\subseteq L_i} f_{\approx j}(L_j)\right]\right|\end{aligned}$$

The second equality is by linearity of expectation and Claim 2.7. We divide the inner sum to L_j contained in L_{i-1} and those that are not. Using linearity of expectation, the contribution from those contained in L_{i-1} is

$$\sum_{j=0}^{i-1}\sum_{L_j\subseteq L_{i-1}} f_{\approx j}(L_j) = f_{\approx i-1}(L_{i-1}) + \sum_{j=0}^{i-2}\sum_{L_j\subseteq L_{i-1}} f_{\approx j}(L_j) = \mu_{L_{i-1},\mathsf{in}}(F),$$

the last equality is by the definition of $f_{\approx i-1}$. Therefore

$$\xi_i = \left| \mathop{\mathbb{E}}_{\substack{L_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ L_{i-1} \subseteq L_i}} \left[\sum_{j=0}^{i-1} \sum_{\substack{L_j \subseteq L_i \\ L_j \not\subseteq L_{i-1}}} f_{\approx j}(L_j) \right] \right|$$

$$= \left| \mathop{\mathbb{E}}_{\substack{L_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ L_{i-1} \subseteq L_i}} \left[\sum_{j=0}^{i-1} \sum_{r=0}^{j-1} \begin{bmatrix} i-1 \\ r \end{bmatrix} \begin{bmatrix} i-r \\ j-r \end{bmatrix} \mathop{\mathbb{E}}_{\substack{L_j \subseteq L_i \\ \dim(L_j \cap L_{i-1}) = r}} [f_{\approx j}(L_j)] \right] \right|.$$

We next interchange order of summations and note that for fixed j, r L_j is a random j-dimensional subspace intersecting L_{i-1} in dimension r to get that

$$\xi_i = \left| \sum_{j=0}^{i-1} \sum_{r=0}^{j-1} \begin{bmatrix} i-1 \\ r \end{bmatrix} \begin{bmatrix} i-r \\ j-r \end{bmatrix} \mathop{\mathbb{E}}_{\substack{L_j \in \begin{bmatrix} V \\ j \end{bmatrix} \\ \dim(L_j \cap L_{i-1}) = r}} [f_{\approx j}(L_j)] \right|.$$

Fix j, r and consider the inner expectation.

$$\mathop{\mathbb{E}}_{\substack{L_j \in \begin{bmatrix} V \\ j \end{bmatrix} \\ \dim(L_j \cap L_{i-1}) = r}} [f_{\approx j}(L_j)] = \frac{1}{\begin{bmatrix} i-1 \\ r \end{bmatrix}} \sum_{R_r \subseteq L_{i-1}} \mathop{\mathbb{E}}_{\substack{L_j \in \begin{bmatrix} V \\ j \end{bmatrix} \\ L_j \cap L_{i-1} = R_r}} [f_{\approx j}(L_j)]$$

for every $R_r \subseteq L_{i-1}$, the distribution over L_j that intersect L_{i-1} on R_r is $j2^{i-k}$-close to the distribution of L_j such that $R_r \subseteq L_j$. Therefore

$$\left| \mathop{\mathbb{E}}_{\substack{L_j \in \begin{bmatrix} V \\ j \end{bmatrix} \\ L_j \cap L_{i-1} = R_r}} [f_{\approx j}(L_j)] \right| \leqslant \mu \left| \mathop{\mathbb{E}}_{\substack{L_j \in \begin{bmatrix} V \\ j \end{bmatrix} \\ L_j \supseteq R_r}} [f_{\approx j}(L_j)] \right| + j2^{i-k} \|f_{\approx j}\|_\infty = |\mu_{R_r, \mathsf{in}}(f_{\approx j})| + j2^{i-k} \|f_{\approx j}\|_\infty$$

which by Claim A.1 and definition of ξ_j, is bounded by $\xi_j + j2^{i-k}2^{j^3}\|F\|_\infty$. Plugging it back in, we get the same bound on

$$\left| \mathop{\mathbb{E}}_{\substack{L_j \in \begin{bmatrix} V \\ j \end{bmatrix} \\ \dim(L_j \cap L_{i-1}) = r}} [f_{\approx j}(L_j)] \right|,$$

and hence by the triangle inequality,

$$\xi_i \leqslant \sum_{j=0}^{i-1} \sum_{r=0}^{j-1} \begin{bmatrix} i-1 \\ r \end{bmatrix} \begin{bmatrix} i-r \\ j-r \end{bmatrix} (\xi_j + j2^{i-k}2^{j^3}\|F\|_\infty)$$

$$\leqslant i^2 2^{2i^2} (\max_{j=1,\ldots,i-1} \xi_j + 2^{i-k}2^{i^3}\|F\|_\infty).$$

Applying the induction hypothesis, we get that the last expression is at most

$$\begin{aligned}&\leqslant i^2 2^{2i^2}(2^{10(i-1)^4-k}\|F\|_\infty + 2^{i^3+i-k}\|F\|_\infty)\\&\leqslant 2^{4i^2}(2^{10(i-1)^4+1-k}\|F\|_\infty)\\&\leqslant 2^{10i^4-k}\|F\|_\infty.\end{aligned}$$

□

We use ξ_i defined in the proof of the previous lemma for the next lemma as well. First, note that for every $j \leqslant i-1$ and every $L_j \in \binom{V}{j}$,

$$\left|\mathop{\mathbb{E}}_{L_i \supseteq L_j}[f_{=i}(L_i)]\right| = \left|\mathop{\mathbb{E}}_{L_{i-1}\supseteq L_j}\left[\mathop{\mathbb{E}}_{L_i \supseteq L_{i-1}}[f_{=i}(L_i)]\right]\right| \leqslant \mathop{\mathbb{E}}_{L_{i-1}\supseteq L_j}\left[\left|\mathop{\mathbb{E}}_{L_i \supseteq L_{i-1}}[f_{=i}(L_i)]\right|\right] \leqslant \xi_i.$$

Lemma A.3. *Let $L_{i-1} \in \binom{V}{i-1}$ and let $F_{\approx i}$ be from Definition 2.24. Then*

$$\left|\mathop{\mathbb{E}}_{L \supseteq L_{i-1}}[F_{\approx i}[L]]\right| \leqslant 2^{20\ell^4-k}\|F\|_\infty.$$

Proof. Let ξ_i be from Lemma A.2. Use Claim 2.18 to get that

$$\mathop{\mathbb{E}}_{L \supseteq L_{i-1}}[F_{\approx i}[L]] = \alpha_{i-1}\cdot \mu_{L_{i-1},\mathsf{in}}(f_{\approx i}) + \sum_{r=0}^{i-2}\alpha_r \sum_{R_r \subseteq L_{i-1}} \mu_{R_r,\mathsf{in}}(f_{\approx i}),$$

where all α_r are upper bounded by $2^{7\ell^2}$ in absolute value. Then by the triangle inequality, the definition of ξ_i and the estimate $\binom{i-1}{r} \leqslant 2^{i^2}$,

$$\left|\mathop{\mathbb{E}}_{L \supseteq L_{i-1}}[F_{\approx i}[L]]\right| \leqslant |\alpha_{i-1}|\,\xi_i + \sum_{r=0}^{i-2}|\alpha_r|\,2^{i^2}\xi_i$$

Next we use the bounds we have on the α_r's and ξ_i to get the last expression is at most

$$\leqslant i2^{i^2}2^{7\ell^2}\xi_i \leqslant 2^{9\ell^2}2^{10i^4-k}\|F\|_\infty \leqslant 2^{20\ell^4-k}\|F\|_\infty.$$

□

Define

$$\tilde{G}_i[L] \stackrel{def}{=} (F - F_{\approx i} - F_{\approx i-1} - ... - F_{\approx 0})[L].$$

We show that $\tilde{G}_i$ is nearly perpendicular to $J_{\leqslant i}$ in the following sense:

Lemma A.4. *Suppose $k \geqslant 5\ell^2+1$, and let $i \leqslant \ell$, $L_i \in \binom{V}{i}$. Then*

$$\left|\mathop{\mathbb{E}}_{L \supseteq L_i}\left[\tilde{G}_i[L]\right]\right| \leqslant 2^{20\ell^4-k}\|F\|_\infty$$

Proof. By definition,

$$\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\tilde{G}_i[L]\right] = \mathop{\mathbb{E}}_{L\supseteq L_i}\left[F[L] - \sum_{j=0}^{i} F_{\approx j}[L]\right] = \mu_{L_i,\mathsf{in}}(F) - \sum_{j=0}^{i}\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\sum_{L_j\subseteq L} f_{\approx j}(L_j)\right], \tag{22}$$

the last equality is by linearity of expectation and the definition of $F_{\approx j}$. Consider the last sum, and for each j divide the inner sum into L_j that are contained in L_i and those that are not. The contribution from $L_j \subseteq L_i$ is

$$\begin{aligned}
\sum_{j=0}^{i}\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\sum_{L_j\subseteq L\cap L_i} f_{\approx j}(L_j)\right] &= \sum_{j=0}^{i}\begin{bmatrix} i \\ j\end{bmatrix}\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\mathop{\mathbb{E}}_{L_j\subseteq L\cap L_i}[f_{\approx j}(L_j)]\right] \\
&= \sum_{j=0}^{i}\begin{bmatrix} i \\ j\end{bmatrix}\mathop{\mathbb{E}}_{L_j\subseteq L_i}[f_{\approx j}(L_j)] \\
&= \sum_{j=0}^{i}\sum_{L_j\subseteq L_i} f_{\approx j}(L_j) \\
&= f_{\approx i}(L_i) + \sum_{j=0}^{i-1}\sum_{L_j\subseteq L_i} f_{\approx j}(L_j) \\
&= \mu_{L_i,\mathsf{in}}(F),
\end{aligned}$$

where we noted L_j is a randomly chosen j-dimensional subspace of L_i in the second expression, and used the definition of $f_{\approx i}$ in the last equality.

Plugging this into Equation (22), we get that

$$\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\tilde{G}_i[L]\right] = -\sum_{j=0}^{i}\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\sum_{\substack{L_j\subseteq L\\ L_j\not\subseteq L_i}} f_{\approx j}(L_j)\right].$$

Using standard manipulations (as in the proof of Lemma A.2), the last sum equals

$$-\sum_{j=0}^{i}\sum_{r=1}^{j-1}\begin{bmatrix} i \\ r\end{bmatrix}\begin{bmatrix} \ell-i \\ j-r\end{bmatrix}\begin{bmatrix} i \\ r\end{bmatrix}\mathop{\mathbb{E}}_{R_r\subseteq L_i}\left[\mathop{\mathbb{E}}_{\substack{L_j\in\begin{bmatrix}V\\ j\end{bmatrix}\\ L_j\cap L_i = R_r}}[f_{\approx j}(L_j)]\right]$$

Fix $j \leqslant i, r \leqslant j-1$ and R_r, and consider the inner expectation. Then the distribution induced on L_j is $r2^{i-k}$ close to uniform over $L_j \supseteq R_r$. Thus the absolute value of this expectation is at most

$$\begin{aligned}
\leqslant \left|\mathop{\mathbb{E}}_{L_j\supseteq R_r}[f_{\approx j}(L_j)]\right| + 2r2^{i-k}\|f_{\approx j}\|_\infty &= |\mu_{R_r,\mathsf{in}}(f_{\approx j})| + 2r2^{i-k}\|f_{\approx j}\|_\infty \\
&\leqslant 2^{10j^4-k}\|F\|_\infty + 2i2^{i-k}2^{i^3}\|F\|_\infty \\
&\leqslant 2^{14i^4-k}\|F\|_\infty,
\end{aligned}$$

the second inequality is by Lemma A.2 and Claim A.1. Thus by the triangle inequality,

$$\left|\mathop{\mathbb{E}}_{L\supseteq L_i}\left[\tilde{G}_i[L]\right]\right| \leqslant \sum_{j=0}^{i}\sum_{r=1}^{j-1}\begin{bmatrix}i\\r\end{bmatrix}\begin{bmatrix}\ell-i\\j-r\end{bmatrix}2^{14i^4-k}\|F\|_\infty \leqslant i^2 2^{i^2+\ell^2}2^{14i^4-k}\|F\|_\infty \leqslant 2^{20\ell^4-k}\|F\|_\infty.$$

□

Before we show prove 2.25, we require rough bounds $\|f_{=i}\|_\infty$ in terms of $\|F\|_\infty$, provided by the following claim.

Claim A.5. *For every $i \in 0, 1, ..., \ell$,*

$$\|f_{=i}\|_\infty \leqslant 2^{2i+i\ell^2}\|F\|_\infty.$$

Proof. Induction on i. For $i = 0, 1$ this is clear by the exact formulas in Section 2.4.1. Let $i \geqslant 2$, assume for all $j < i$ and prove for i. Then for every $j < i$, since $F_{=j}[L] = \sum_{L_j\subseteq L} f_{=j}(L_j)$, we have that

$$\|F_{=j}\|_\infty \leqslant \begin{bmatrix}\ell\\j\end{bmatrix}\|f_{=j}\|_\infty \leqslant 2^{2j+(j+1)\ell^2}\|F\|_\infty,$$

the last inequality is by the induction hypothesis.

Let $L_i \in \begin{bmatrix}V\\i\end{bmatrix}$ be a space obtaining $|f_{=i}(L_i)| = \|f_{=i}\|_\infty$. Since $F - F_{=i} - ... - F_{=0}$ is perpendicular to $J_{\leqslant i}$,

$$\mathop{\mathbb{E}}_{L\supseteq L_i}[(F - F_{=i} - ... - F_{=0})[L]] = 0,$$

and therefore by triangle inequality

$$\left|\mathop{\mathbb{E}}_{L\supseteq L_i}[F_{=i}[L]]\right| \leqslant |\mu_{L_i,\mathsf{in}}| + \sum_{j=0}^{i-1}\left|\mathop{\mathbb{E}}_{L\supseteq L_i}[F_{=j}[L]]\right| \leqslant \|F\|_\infty + \sum_{j=0}^{i-1}\|F_{=j}\|_\infty \leqslant \|F\|_\infty + \sum_{j=0}^{i-1}2^{2j+(j+1)\ell^2}\|F\|_\infty. \tag{23}$$

On the other hand,

$$\begin{aligned}\mathop{\mathbb{E}}_{L\supseteq L_i}[F_{=i}[L]] &= \mathop{\mathbb{E}}_{L\supseteq L_i}\left[\sum_{L'_i\subseteq L} f_{=i}(L'_i)\right] = f_{=i}(L_i) + \mathop{\mathbb{E}}_{L\supseteq L_i}\left[\sum_{L'_i\subseteq L, L'_i\neq L_i} f_{=i}(L'_i)\right]\\ &= f_{=i}(L_i) + \mathop{\mathbb{E}}_{L\supseteq L_i}\left[\sum_{r=0}^{i-1}\begin{bmatrix}i\\r\end{bmatrix}\begin{bmatrix}\ell-r\\i-r\end{bmatrix}\mathop{\mathbb{E}}_{R_r\subseteq L_i}\left[\mathop{\mathbb{E}}_{\substack{L'_i\subseteq L\\ L'_i\cap L_i = R_r}}\left[f_{=i}(L'_i)\right]\right]\right].\end{aligned}$$

Rearranging and noting that for fixed r, R_r, L'_i is distributed uniformly among the i-dimensional subspaces intersecting L_i in R_r, we get that

$$\mathop{\mathbb{E}}_{L\supseteq L_i}[F_i[L]] = f_{=i}(L_i) + \sum_{r=0}^{i-1}\begin{bmatrix}i\\r\end{bmatrix}\begin{bmatrix}\ell-r\\i-r\end{bmatrix}\mathop{\mathbb{E}}_{R_r\subseteq L_i}\left[\mathop{\mathbb{E}}_{\substack{L'_i\in\begin{bmatrix}V\\i\end{bmatrix}\\ L'_i\cap L_i = R_r}}\left[f_{=i}(L'_i)\right]\right].$$

For a fixed $r \leqslant i-1$, R_r, the distribution of L_i' is $i2^{r-k}$ close to uniform over all i-dimensional subspaces containing R_r, and hence

$$\left| \mathop{\mathbb{E}}_{\substack{L_i' \in \binom{V}{i} \\ L_i' \cap L_i = R_r}} \left[f_{=i}(L_i') \right] \right| \leqslant \left| \mathop{\mathbb{E}}_{L_i' \supseteq R_r} \left[f_{=i}(L_i') \right] \right| + i2^{r-k} \|f_{=i}\|_\infty = i2^{r-k} \|f_{=i}\|_\infty,$$

in the last equality we used the fact the inner expectation is 0 by Lemma 2.19 ($F_{=i} \in J_{=i}$). Therefore by the triangle inequality,

$$\begin{aligned}
\left| \mathop{\mathbb{E}}_{L \supseteq L_i} [F_i[L]] \right| &\geqslant |f_{=i}(L_i)| - \mathop{\mathbb{E}}_{L \supseteq L_i} \left[\sum_{r=0}^{i-1} \begin{bmatrix} i \\ r \end{bmatrix} \begin{bmatrix} \ell - r \\ i - r \end{bmatrix} \mathop{\mathbb{E}}_{R_r \subseteq L_i} \left[\left| \mathop{\mathbb{E}}_{\substack{L_i' \subseteq L \\ L_i' \cap L_i = R_r}} \left[f_{=i}(L_i') \right] \right| \right] \right] \\
&\geqslant \|f_{=i}\|_\infty - i \cdot 2^{i^2} 2^{\ell^2} \cdot i \cdot 2^{r-k} \|f_{=i}\|_\infty \\
&\geqslant \frac{1}{2} \|f_{=i}\|_\infty
\end{aligned}$$

Combining this with Equation (23) we conclude that

$$\begin{aligned}
\|f_{=i}\|_\infty \leqslant 2\|F\|_\infty + 2 \sum_{j=0}^{i-1} 2^{2j+(j+1)\ell^2} \|F\|_\infty &\leqslant 2\|F\|_\infty + \cdot 2^{i\ell^2} \|F\|_\infty \sum_{j=0}^{i-1} 2^{2j} \\
&\leqslant 2\|F\|_\infty + 2 \cdot 2^{i\ell^2} \|F\|_\infty \frac{2^{2i}}{3} \\
&\leqslant 2^{2i+i\ell^2} \|F\|_\infty.
\end{aligned}$$

□

Proposition A.6. *Let $i \leqslant \ell$ be integers, $H[L]$, $F[L] = \sum_{L_i \subseteq L} f_i(L_i)$ be a functions, and assume $|\mathbb{E}_{L \supseteq L_i} [H[L]]| \leqslant \varepsilon$ for every $L_i \in \binom{V}{i}$. Then*

$$|\langle H, F \rangle| \leqslant 2^{i\ell} \|f_i\|_\infty \varepsilon.$$

Proof.

$$\begin{aligned}
\langle H, F \rangle = \mathop{\mathbb{E}}_{L} [H[L]F[L]] &= \mathop{\mathbb{E}}_{L} \left[\sum_{L_i \subseteq L} H[L] f_i(L_i) \right] \\
&= \mathop{\mathbb{E}}_{L} \left[\begin{bmatrix} \ell \\ i \end{bmatrix} \mathop{\mathbb{E}}_{L_i \subseteq L} [H[L] f_i(L_i)] \right] \\
&= \begin{bmatrix} \ell \\ i \end{bmatrix} \mathop{\mathbb{E}}_{L_i \in \binom{V}{i}} \left[f_i(L_i) \mathop{\mathbb{E}}_{L \supseteq L_i} [H[L]] \right].
\end{aligned}$$

The proof is concluded by using the triangle inequality and the crude estimate $\begin{bmatrix} \ell \\ i \end{bmatrix} \leqslant 2^{i\ell}$. □

A.1 Proof of Theorem 2.25

Theorem 2.25 (Restated) . *Assume $k \geqslant 5\ell^2 + 1$, let V be a k-dimensional vector space over $\mathbb{F}_2$. Let $F\colon \begin{bmatrix} V \\ \ell \end{bmatrix} \to \mathbb{R}$, $0 \leqslant i \leqslant \ell$, let F_i be the projection of F onto $J_{=i}$ and let $F_{\approx i}$ be from Definition 2.24. Then*

$$\|F_{=i} - F_{\approx i}\|_2^2 \leqslant 2^{26\ell^4 - k}\|F\|_\infty^2.$$

Proof. For $i = 0, 1$ the theorem is obvious by the exact formula we have seen for $F_{=0}, F_{=1}$.

Recall that we have defined

$$\tilde{G}_i[L] = (F - F_{\approx i} - \cdots - F_{\approx 0})[L],$$

and define

$$G_i[L] = (F - F_{=i} - ... - F_{=0})[L].$$

Note that G_i is perpendicular to $J_{\leqslant i}$. Apply Proposition A.6 to $\langle \tilde{G}_i, F_{\approx i}\rangle$ ($H = \tilde{G}_i$, using Claims A.1,A.4), and to $\langle \tilde{G}_i, F_{=i}\rangle$ ($H = \tilde{G}_i$, using Claim A.5, A.4) to get that both terms are bounded by $\varepsilon \stackrel{def}{=} 2^{25\ell^4 - k}\|F\|_\infty^2$.

On the other hand, since F_i is perpendicular to $J_{\leqslant i-1}$ and $F_{\approx 0} + ... + F_{\approx i-1} \in J_{\leqslant i-1}$,

$$\langle \tilde{G}_i, F_{=i}\rangle = \langle F, F_{=i}\rangle - \langle F_{\approx i}, F_{=i}\rangle = \langle F_{=i}, F_{=i}\rangle - \langle F_{\approx i}, F_{=i}\rangle = \|F_{=i}\|_2^2 - \langle F_{\approx i}, F_{=i}\rangle,$$

the second equality is by orthogonality. Thus

$$\left|\|F_{=i}\|_2^2 - \langle F_{\approx i}, F_{=i}\rangle\right| \leqslant \varepsilon \tag{24}$$

Additionally,

$$\langle \tilde{G}_i, F_{\approx i}\rangle = \langle F, F_{\approx i}\rangle - \langle F_{\approx i}, F_{\approx i}\rangle - \sum_{j=0}^{i-1} \langle F_{\approx j}, F_{\approx i}\rangle,$$

thus

$$\left|\langle F, F_{\approx i}\rangle - \|F_{\approx i}\|_2^2\right| \leqslant \left|\langle \tilde{G}_i, F_{\approx i}\rangle\right| + \sum_{j=0}^{i-1} |\langle F_{\approx j}, F_{\approx i}\rangle| \leqslant i\varepsilon, \tag{25}$$

We have used $|\langle F_{\approx i}, F_{\approx j}\rangle| \leqslant \varepsilon$ that follows by Proposition A.6 using Claims A.3 and A.1.

Since G_i is perpendicular to $J_{\leqslant i}$ and $F_{\approx i} \in J_{\leqslant i}$ we have

$$0 = \langle G_i, F_{\approx i}\rangle = \langle F, F_{\approx i}\rangle - \langle F_{=i}, F_{\approx i}\rangle - \sum_{j=0}^{i-1} \langle F_{=j}, F_{\approx i}\rangle,$$

Thus as before

$$|\langle F, F_{\approx i}\rangle - \langle F_{=i}, F_{\approx i}\rangle| \leqslant \sum_{j=0}^{i-1} |\langle F_{=j}, F_{\approx i}\rangle| \leqslant (i-1)\varepsilon. \tag{26}$$

We have used $|\langle F_{\approx i}, F_{=j}\rangle| \leqslant \varepsilon$ that follows by Proposition A.6 using Claims A.3 and A.5.

Combining Equation (24),(25), (26) we finish the proof:

$$\begin{aligned}
\|F_{=i} - F_{\approx i}\|_2^2 &= \|F_{=i}\|_2^2 - 2\langle F_{=i}, F_{\approx i}\rangle + \|F_{\approx i}\|_2^2 \\
&= (\|F_{=i}\|_2^2 - \langle F_{\approx i}, F_{=i}\rangle) + (\langle F, F_{\approx i}\rangle - \langle F_{=i}, F_{\approx i}\rangle) + (\|F_{\approx i}\|_2^2 - \langle F, F_{\approx i}\rangle) \\
&\leqslant 2i\varepsilon \leqslant 2^{26\ell^4 - k}\|F\|_\infty^2.
\end{aligned}$$

□

B Proof of Claim 2.18

B.1 Auxiliary propositions

For an integer $d \geqslant 0$, define a sequence $\{a_d(m)\}_{m\in\mathbb{Z}}$ by $a_d(0) = 1$, $a_d(1) = 1 - \begin{bmatrix} d+1 \\ d \end{bmatrix}$, and inductively

$$a_d(m+1) = 1 - \sum_{n=0}^{m} \begin{bmatrix} d+m+1 \\ d+n \end{bmatrix} a_d(n)$$

We also define $a_d(-1) = a_d(-2) = \cdots = 0$.

To prove Claim 2.18, we require two simple propositions. The first of them is an inclusion-exclusion type statement.

Proposition B.1. *Let $0 \leqslant d \leqslant j < i$ be integers and $f\colon \begin{bmatrix} V \\ i \end{bmatrix} \to \mathbb{R}$ be a function. Then for every $L_j \in \begin{bmatrix} V \\ j \end{bmatrix}$,*

$$\sum_{\substack{R_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ \dim(R_i \cap L_j) \geqslant d}} f(R_i) = \sum_{r=d}^{j} a_d(r-d) \sum_{R_r \subseteq L_j} \sum_{R_i \supseteq R_r} f(R_i).$$

Proof. Let us count the number of times an R_i appears in the sum in the right hand side. Fix R_i such that $\dim(R_i \cap L_j) = d+m$, for $m \geqslant 0$. Note that R_i appears in the inner sums corresponding to $r = d, ..., d+m$, and is counted $\begin{bmatrix} d+m \\ r \end{bmatrix}$ times respectively in each sum (since this is the number of R_r from L_j that are subspaces of R_i). Therefore the coefficient of $f(R_i)$ on the right hand side is

$$\sum_{r=d}^{d+m} \begin{bmatrix} d+m \\ r \end{bmatrix} a_d(r-d) = a_d(m) + \sum_{n=0}^{m-1} \begin{bmatrix} d+m \\ d+n \end{bmatrix} a_d(n) = 1,$$

the last equality is by the definition of the sequence $a_d(n)$. □

Proposition B.2. *Let $0 \leqslant d \leqslant j < i$ be integers and $f\colon \begin{bmatrix} V \\ i \end{bmatrix} \to \mathbb{R}$ be a function. Then*

$$\sum_{\substack{R_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ \dim(R_i \cap L_j) = d}} f(R_i) = \sum_{r=d}^{j} \left(a_d(r-d) - a_{d+1}(r-d-1)\right) \begin{bmatrix} k-r \\ i-r \end{bmatrix} \sum_{R_r \subseteq L_j} \mathop{\mathbb{E}}_{R_i \supseteq R_r} [f(R_i)].$$

Proof. It holds that

$$\sum_{\substack{R_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ \dim(R_i \cap L_j) = d}} f(R_i) = \sum_{\substack{R_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ \dim(R_i \cap L_j) \geqslant d}} f(R_i) - \sum_{\substack{R_i \in \begin{bmatrix} V \\ i \end{bmatrix} \\ \dim(R_i \cap L_j) \geqslant d+1}} f(R_i).$$

Use Proposition B.1 on each one of the sums, we get that their difference equals

$$\begin{aligned}
&\sum_{r=d}^{j} a_d(r-d) \sum_{R_r \subseteq L_j} \sum_{R_i \supseteq R_r} f(R_i) - \sum_{r=d+1}^{j} a_{d+1}(r-d-1) \sum_{R_r \subseteq L_j} \sum_{R_i \supseteq R_r} f(R_i) \\
&= \sum_{r=d}^{j} \left(a_d(r-d) - a_{d+1}(r-d-1)\right) \sum_{R_r \subseteq L_j} \sum_{R_i \supseteq R_r} f(R_i) \\
&= \sum_{r=d}^{j} \left(a_d(r-d) - a_{d+1}(r-d-1)\right) \begin{bmatrix} k-r \\ i-r \end{bmatrix} \sum_{R_r \subseteq L_j} \mathop{\mathbb{E}}_{R_i \supseteq R_r} [f(R_i)].
\end{aligned}$$

□

The third and final proposition is a crude upper-bound on the sequence $a_d(m)$.

Proposition B.3. *For every $d \geqslant 0$ and m,*

$$|a_d(m)| \leqslant |m+1|\, 2^{m(d+m)}.$$

Proof. Fix d, the proof is by induction on m (clear for negative m's). For $n = 1, 2$, the claim is obvious by the definition of $a_d(m)$. Let $m \geqslant 3$, assume for all $n \leqslant m$ and prove for $m+1$. By definition and the triangle inequality,

$$\begin{aligned} |a_d(m+1)| &= \left| 1 - \sum_{n=0}^{m} \begin{bmatrix} d+m+1 \\ d+n \end{bmatrix} a_d(n) \right| \\ &\leqslant 1 + \sum_{n=0}^{m} \begin{bmatrix} d+m+1 \\ d+n \end{bmatrix} |a_d(n)|. \end{aligned}$$

We apply the induction hypothesis and the crude estimate $\begin{bmatrix} t \\ p \end{bmatrix} \leqslant 2^{t\cdot(t-p)}$ to get

$$\begin{aligned} |a_d(m+1)| &\leqslant 1 + \sum_{n=0}^{m} 2^{(d+m+1)(m+1-n)}(n+1)2^{n(d+n)} \\ &= 1 + 2^{(d+m+1)(m+1)} \sum_{n=0}^{m} (n+1)2^{n(n-m-1)} \\ &\leqslant 1 + 2^{(d+m+1)(m+1)} \sum_{n=0}^{m} (n+1)2^{-n} \\ &\leqslant 1 + 2^{(d+m+1)(m+1)}(m+1) \sum_{n=0}^{m} 2^{-n} \\ &\leqslant 1 + (m+1)2^{(d+m+1)(m+1)} \\ &\leqslant (m+2)2^{(d+m+1)(m+1)}. \end{aligned}$$

□

Claim 2.18 (Restated) . *Suppose $0 \leqslant j < i \leqslant \ell$, $k \geqslant 7\ell^2 + 1$. There exists $\beta_0, ..., \beta_j \in \mathbb{R}$ such that the following holds. For every $F \in J_{\leqslant i}$ given by $F[L] = \sum_{R_i \subseteq L} f(R_i)$ and $L_j \in \begin{bmatrix} V \\ j \end{bmatrix}$,*

$$\mathop{\mathbb{E}}_{L \supseteq L_j}[F[L]] = \left(\begin{bmatrix} \ell - j \\ i - j \end{bmatrix} + \beta_j \right) \cdot \mu_{L_j,\mathsf{in}}(f) + \sum_{r=0}^{j-1} \beta_r \sum_{R_r \subseteq L_j} \mu_{R_r,\mathsf{in}}(f).$$

Additionally, the β's have the following properties:

- *for $r = 0, ..., j-1$, $|\beta_r| \leqslant 2^{6\ell^2}$.*
- *$|\beta_j| \leqslant 2^{7\ell^2 - k}$.*

In particular, the coefficient of $\mu_{L_j,\mathsf{in}}(f)$ is not 0.

Proof. By definition

$$\begin{aligned}
\mathop{\mathbb{E}}_{L\supseteq L_j}[F[L]] &= \mathop{\mathbb{E}}_{L\supseteq L_j}\left[\sum_{R_i\subseteq L} f(R_i)\right] \\
&= \mathop{\mathbb{E}}_{L\supseteq L_j}\left[\sum_{d=0}^{j}\sum_{\substack{R_i\subseteq L\\ \dim(R_i\cap L_j)=d}} f(R_i)\right] \\
&= \mathop{\mathbb{E}}_{L\supseteq L_j}\left[\sum_{d=0}^{j}\begin{bmatrix} j\\ d\end{bmatrix}\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}\mathop{\mathbb{E}}_{\substack{R_i\subseteq L\\ \dim(R_i\cap L_j)=d}}[f(R_i)]\right]. \qquad (27)
\end{aligned}$$

We have used the fact that there are $\begin{bmatrix} j\\ d\end{bmatrix}\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}$ i-dimensional subspaces of a given ℓ-dimensional space L that intersect L_j in dimension d. Using linearity of expectation, we interchange expectation and the sum to get that the expression in (27) equals

$$\begin{aligned}
&\sum_{d=0}^{j}\begin{bmatrix} j\\ d\end{bmatrix}\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}\mathop{\mathbb{E}}_{\substack{R_i\in\begin{bmatrix}V\\ i\end{bmatrix}\\ \dim(R_i\cap L_j)=d}}[f(R_i)] \\
&= \begin{bmatrix}\ell-j\\ i-j\end{bmatrix}\mathop{\mathbb{E}}_{R_i\supseteq L_j}[f(R_i)] + \sum_{d=0}^{j-1}\begin{bmatrix} j\\ d\end{bmatrix}\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}\mathop{\mathbb{E}}_{\substack{R_i\in\begin{bmatrix}V\\ i\end{bmatrix}\\ \dim(R_i\cap L_j)=d}}[f(R_i)].
\end{aligned}$$

Next we rework the expectation inside the sum. Fix $d\leqslant j-1$, then

$$\begin{aligned}
\mathop{\mathbb{E}}_{\substack{R_i\in\begin{bmatrix}V\\ i\end{bmatrix}\\ \dim(R_i\cap L_j)=d}}[f(R_i)] &= \frac{1}{\begin{bmatrix} j\\ d\end{bmatrix}\begin{bmatrix} k-j\\ i-d\end{bmatrix}}\sum_{\substack{R_i\in\begin{bmatrix}V\\ i\end{bmatrix}\\ \dim(R_i\cap L_j)=d}} f(R_i) \\
&= \frac{1}{\begin{bmatrix} j\\ d\end{bmatrix}\begin{bmatrix} k-j\\ i-d\end{bmatrix}}\sum_{r=d}^{j}(a_d(r-d)-a_{d+1}(r-d-1))\begin{bmatrix} k-r\\ i-r\end{bmatrix}\sum_{R_r\subseteq L_{i-1}}\mathop{\mathbb{E}}_{R_i\supseteq R_r}[f(R_i)],
\end{aligned}$$

the second equality is by Proposition B.2. Plugging the previous two equalities into Equation 27 yields

$$\mathop{\mathbb{E}}_{L\supseteq L_j}[F[L]] = \begin{bmatrix}\ell-j\\ i-j\end{bmatrix}\mu_{L_j,\mathsf{in}}(f) + \sum_{d=0}^{j-1}\frac{\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}}{\begin{bmatrix} k-j\\ i-d\end{bmatrix}}\sum_{r=d}^{j}(a_d(r-d)-a_{d+1}(r-d-1))\begin{bmatrix} k-r\\ i-r\end{bmatrix}\sum_{R_r\subseteq L_{i-1}}\mu_{R_r,\mathsf{in}}(f).$$

Interchanging the order of summation on d, r, we get that the last expression equals

$$\begin{aligned}
&= \begin{bmatrix}\ell-j\\ i-j\end{bmatrix}\mu_{L_j,\mathsf{in}}(f) + \sum_{r=0}^{j}\begin{bmatrix} k-r\\ i-r\end{bmatrix}\sum_{d=0}^{\min(r,j-1)}\frac{\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}}{\begin{bmatrix} k-j\\ i-d\end{bmatrix}}(a_d(r-d)-a_{d+1}(r-d-1))\sum_{R_r\subseteq L_j}\mu_{R_r,\mathsf{in}}(f) \\
&= \begin{bmatrix}\ell-j\\ i-j\end{bmatrix}\mu_{L_{i-1},\mathsf{in}}(f) + \sum_{r=0}^{j}\beta_r\sum_{R_r\subseteq L_{i-1}}\mu_{R_r,\mathsf{in}}(f),
\end{aligned}$$

as desired. We next estimate the coefficients. Let $r = 0, 1, ..., j-1$, then

$$|\beta_r| = \left| \begin{bmatrix} k-r \\ i-r \end{bmatrix} \sum_{d=0}^{\min(r,j-1)} \frac{\begin{bmatrix} \ell-j \\ i-d \end{bmatrix}}{\begin{bmatrix} k-j \\ i-d \end{bmatrix}} \left(a_d(r-d) - a_{d+1}(r-d-1)\right) \right|$$

Note that since $i - d \geqslant i - r$ for each j in the sum, and the Gaussian coefficients are increasing in that interval, we have that $\begin{bmatrix} k-j \\ i-d \end{bmatrix} \geqslant \begin{bmatrix} k-j \\ i-r \end{bmatrix}$. Further note that by the triangle inequality and Proposition B.3, $\max_{d=0,\dots,j-1} \left|a_d(r-d) - a_{d+1}(r-d-1)\right| \leqslant 2(r+1) \cdot 2^{r^2} \leqslant 2^{2r^2+1}$, we get that

$$|\beta_r| \leqslant 2^{2r^2+1} \frac{\begin{bmatrix} k-r \\ i-r \end{bmatrix}}{\begin{bmatrix} k-i+1 \\ i-r \end{bmatrix}} \sum_{d=0}^{j-1} \begin{bmatrix} \ell-j \\ i-d \end{bmatrix}.$$

The sum can be bounded by

$$\sum_{d=0}^{j-1} \begin{bmatrix} \ell-j \\ i-d \end{bmatrix} \leqslant j \max_{d=0,\dots,j-1} \begin{bmatrix} \ell-j \\ i-d \end{bmatrix} \leqslant i 2^{\ell^2} \leqslant 2^{2\ell^2}.$$

The ratio of the Gaussian coefficients can be bounded by:

$$\begin{aligned} \frac{\begin{bmatrix} k-r \\ i-r \end{bmatrix}}{\begin{bmatrix} k-j \\ i-r \end{bmatrix}} &= \frac{(2^{k-r}-1)\cdots(2^{k-r}-2^{i-r-1})(2^{i-r}-1)\cdots(2^{i-r}-2^{i-r-1})}{(2^{k-j}-1)\cdots(2^{k-j}-2^{i-r-1})(2^{k-i+1}-2^{i-r-1})(2^{i-r}-1)\cdots(2^{i-r}-2^{i-r-1})} \\ &= \prod_{m=0}^{i-r-1} \frac{2^{k-r}-2^m}{2^{k-j}-2^m} \leqslant \prod_{m=0}^{i-r-1} \frac{2^{k-r}}{2^{k-j-1}} = 2^{(j+1-r)(i-r)} \leqslant 2^{\ell^2}. \end{aligned}$$

Combining the above, we conclude that $\beta_r \leqslant 2^{2r^2+1} 2^{2\ell^2} 2^{\ell^2} \leqslant 2^{6\ell^2}$.

Finally we estimate the coefficient of $\mu_{L_j,\mathsf{in}}(f)$. Clearly this coefficient is $\begin{bmatrix} \ell-i+1 \\ i-j \end{bmatrix} + \beta_j$.

$$\begin{aligned} |\beta_j| &= \left| \begin{bmatrix} k-j \\ i-j \end{bmatrix} \sum_{d=0}^{j-1} \frac{\begin{bmatrix} \ell-j \\ i-d \end{bmatrix}}{\begin{bmatrix} k-j \\ i-d \end{bmatrix}} \left(a_d(r-d) - a_{d+1}(r-d-1)\right) \right| \\ &\leqslant \frac{\begin{bmatrix} k-j \\ i-j \end{bmatrix}}{\begin{bmatrix} k-j \\ i-j+1 \end{bmatrix}} \sum_{d=0}^{j-1} \begin{bmatrix} \ell-j \\ i-d \end{bmatrix} \left|a_d(r-d) - a_{d+1}(r-d-1)\right| \end{aligned}$$

The maximal difference between the a's is bounded the same way as before by $2^{3\ell^2}$. Additionally, a basic computation shows that

$$\begin{aligned} \frac{\begin{bmatrix} k-j \\ i-j \end{bmatrix}}{\begin{bmatrix} k-j \\ i-j+1 \end{bmatrix}} &= \frac{1}{(2^{k-j}-2^{i-j})} \frac{(2^{i-j+1}-1)\cdots(2^{i-j+1}-2^{i-j})}{(2^{i-j}-1)\cdots(2^{i-j}-2^{i-j-1})} \\ &= \frac{1}{(2^{k-j}-2^{i-j})} 2^{i-j-1}(2^{i-j+1}-1) \\ &= \frac{1}{2^{k-i+1}-2}(2^{i-j+1}-1) \\ &\leqslant \frac{1}{2^{k-i}}(2^i-1) \\ &\leqslant 2^{2\ell-k}. \end{aligned}$$

Combining the above yields

$$\begin{aligned}|\beta_j| &\leqslant 2^{3\ell^2}2^{2\ell-k}\sum_{d=0}^{j-1}\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}\\ &\leqslant 2^{3\ell^2}2^{2\ell-k}j\max_{d=0,\ldots,j-1}\begin{bmatrix}\ell-j\\ i-d\end{bmatrix}\\ &\leqslant 2^{3\ell^2}2^{2\ell-k}2^{\ell}2^{\ell^2}\\ &\leqslant 2^{7\ell^2-k}.\end{aligned}$$

□

C Proof of Theorem 3.6

Below we construct the S randomly.

Let $\{X_u\}_{u\in V\setminus\{0\}}$ be independent uniform $\{-1,1\}$-valued random variables. Let $Y_2,...,Y_{2^\ell}$ be additional uniform $\{-1,1\}$ independent random variables (those will only be used in the analysis). Define a random variable

$$g[L]=\frac{1}{\sqrt{2^\ell-1}}\sum_{u\in L\setminus\{0\}}X_u.$$

To construct the set S, add each $L\in\begin{bmatrix}V\\ \ell\end{bmatrix}$ such that $g[L]\geqslant c\log(1/\delta)$ for constant c such that $\delta\leqslant \Pr_L\left[g[L]\geqslant c\log(1/\delta)\right]\leqslant 2\delta$ (there exists such constant since $g[L]$ is close to a standard Gaussian random variable).

The density of S is $\Theta(\delta)$ with probability $1-o(1)$. For each L, let Q_L be the indicator random variable which is one iff $L\in S$. Let μ be the expectation of Q_L (which we know is between δ and 2δ). We prove that

$$\Pr\left[\left|\sum_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}Q_L-\begin{bmatrix}k\\ \ell\end{bmatrix}\mu\right|\leqslant 2^{-\ell}\begin{bmatrix}k\\ \ell\end{bmatrix}\right]\geqslant 1-o(1).$$

Given that, it is clear we have $\mu(S)=\mu+O(2^{-\ell})=\Theta(\delta)$ with probability $1-o(1)$.

Proposition C.1.

$$\Pr\left[\left|\sum_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}Q_L-\mu\right|\leqslant 2^{-\ell}\begin{bmatrix}k\\ \ell\end{bmatrix}\right]\geqslant 1-O(2^{10\ell-2k}).$$

Proof. We use the fourth moment method. Note that all Q_L have identical expectation μ.

Let us estimate the variance of the sum of the Q_L's.

$$\mathbb{E}\left[\left(\sum_{L\in\begin{bmatrix}V\\ \ell\end{bmatrix}}Q_L-\mu\right)^4\right]=\mathbb{E}\left[\sum_{L,M,P,N\in\begin{bmatrix}V\\ \ell\end{bmatrix}}(Q_L-\mu)(Q_M-\mu)(Q_P-\mu)(Q_N-\mu)\right].$$

Note that if there is a space out of $\{L, M, P, N\}$ that intersects all others in $\{0\}$, then the expectation is 0. For any quadruples, this expectation is at most 1. Therefore the variance above is at most the number of quadruples such that for any space intersects at least one other space non-trivially. The number of such quadruples is at most

$$\begin{bmatrix} k \\ \ell \end{bmatrix} \cdot \begin{bmatrix} \ell \\ 1 \end{bmatrix} \cdot 3 \begin{bmatrix} k-1 \\ \ell-1 \end{bmatrix} \cdot \begin{bmatrix} k \\ \ell \end{bmatrix} \cdot 3 \begin{bmatrix} \ell \\ 1 \end{bmatrix} \cdot \begin{bmatrix} k-1 \\ \ell-1 \end{bmatrix}$$

The first factor chooses L, the second factor chooses non-trivial intersection, the third factor chooses the rest of the space to intersect L - say M. The fourth factor chooses P, and the fourth factor chooses non-trivial intersection of N with one of L, M, N and the fifth completes the choice of N.

To get some crude estimate of it, note that

$$\frac{\begin{bmatrix} k-1 \\ \ell-1 \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix}} \leqslant \frac{2^{2\ell}}{2^k - 2^{\ell-1}} \leqslant 2^{2\ell+1-k}.$$

Hence the number of quadruples is at most

$$\begin{bmatrix} k \\ \ell \end{bmatrix}^4 2^{6\ell+6-2k}.$$

Therefore by Markov's inequality

$$\Pr\left[\left|\sum_{L \in \begin{bmatrix} V \\ \ell \end{bmatrix}} Q_L - \mu \begin{bmatrix} k \\ \ell \end{bmatrix}\right| \geqslant 2^{-\ell} \begin{bmatrix} k \\ \ell \end{bmatrix}\right] = \Pr\left[\left|\sum_{L \in \begin{bmatrix} V \\ \ell \end{bmatrix}} Q_L - \mu \begin{bmatrix} k \\ \ell \end{bmatrix}\right|^4 \geqslant 2^{-4\ell} \begin{bmatrix} k \\ \ell \end{bmatrix}^4\right] \leqslant O(2^{10\ell-2k}).$$

□

Zooms of S. For any hyperplane W, one can repeat the argument of the proof of Proposition C.1 to show that $\mu(S_W) = \mu + O(2^{-\ell})$ with probability $O(2^{10\ell-2k})$. Hence by union bound, zoom-outs into one dimension do not increase the density of S by more that $O(2^{-\ell})$ with probability $1 - O(2^{10\ell-k})$.

Fix a non-zero $v \in V$. We analyze the expectation of $\mu(S_v) - \mu(S)$ conditioned on $X_v = 1$. A similar analysis works conditioned on $X_v = -1$.

Denote $m = \lceil c\sqrt{\log(1/\delta)}\sqrt{2^\ell - 1} \rceil$. Let L be a subspace containing v. Then the probability L conditioned on $X_v = 1$ is in S is

$$\Pr\left[\frac{1}{\sqrt{2^\ell - 1}}(1 + Y_3 + ... + Y_{2^\ell}) \geqslant c\sqrt{\log(1/\delta)}\right] = \Pr\left[Y_3 + ... + Y_{2^\ell} \geqslant m - 1\right],$$

hence this is the expected density of S_v.

Similarly, the expected density of $\mu(S)$ is

$$\Pr\left[Y_2 + Y_3 + ... + Y_{2^\ell} \geqslant m\right] = \frac{1}{2}\Pr\left[Y_3 + ... + Y_{2^\ell} \geqslant m - 1\right] + \frac{1}{2}\Pr\left[Y_3 + ... + Y_{2^\ell} \geqslant m + 1\right].$$

Let us assume m is even– the proof is similar otherwise. Then

$$\mathbb{E}\left[\mu(S_v)\right] - \mu = \frac{1}{2}\Pr\left[m - 1 \leqslant Y_3 + ... + Y_{2^\ell} \leqslant m\right] = \Theta(2^{-2^\ell}) \binom{2^\ell - 2}{2^{\ell-1} - 1 + m/2}.$$

Proposition C.2. $\mathbb{E}\left[\mu(S_v) \mid X_v = 1\right] - \mu \leqslant \Omega(2^{-\ell/2})$.

Proof. Using the standard fact

$$\sum_{i=m/2}^{n/2} \binom{n}{n/2+i} \geqslant \Omega\left(\frac{n}{m}\right) \binom{n}{n/2+m/2}$$

we get that by the inequality preceding the proposition that

$$\begin{aligned}
\mathbb{E}\left[\mu(S_v)\right] - \mu &\leqslant 2^{-\ell/2}\sqrt{\log(1/\delta)}O(2^{-2^\ell}) \sum_{i=m/2}^{2^{\ell-1}-1} \binom{2^\ell - 2}{2^{\ell-1}-1+i} \\
&= O(2^{-\ell/2}\sqrt{\log(1/\delta)}) \sum_{i=m/2}^{2^{\ell-1}-1} \binom{2^\ell - 1}{2^{\ell-1}+i} \\
&= O(2^{-\ell/2}\sqrt{\log(1/\delta)})\Pr\left[Y_2 + ... + Y_{2^\ell} \geqslant m\right] \\
&= O(2^{-\ell/2}\sqrt{\log(1/\delta)}\delta).
\end{aligned}$$

□

Proposition C.3. $\mathbb{E}\left[\mu(S_v) \mid X_v = 1\right] - \mu \geqslant \Omega(2^{-\ell/2}\delta)$.

Proof. Using the standard fact

$$\sum_{i=m/2}^{n/2} \binom{n}{n/2+i} \leqslant \sqrt{n}\binom{n}{n/2+m/2}$$

we get that by the inequality preceding the propositions that

$$\begin{aligned}
\mathbb{E}\left[\mu(S_v)\right] - \mu &\geqslant 2^{-\ell/2}\Omega(2^{-2^\ell}) \sum_{i=m/2}^{2^{\ell-1}-1} \binom{2^\ell - 2}{2^{\ell-1}-1+i} \\
&= \Omega(2^{-\ell/2})2^{-2^\ell} \sum_{i=m/2}^{2^{\ell-1}-1} \binom{2^\ell - 1}{2^{\ell-1}+i} \\
&= \Omega(2^{-\ell/2})\Pr\left[Y_2 + ... + Y_{2^\ell} \geqslant m\right] \\
&= \Omega(2^{-\ell/2}\delta).
\end{aligned}$$

□

From both propositions we see that conditioned on $X_v = 1$, the expected density of S_v is $\mu + O(2^{-\ell/2})$. Repeating the argument of C.1 it is easy to prove that with probability $1 - O(2^{10\ell-2k})$ the density of S_v is $\mu + O(2^{-\ell/2})$ conditioned on $X_v = 1$.

The case where we condition on $X_v = -1$ is similar.

By a union bound the probability for all v, the density of S_v is $\mu + O(2^{-\ell/2})$ is $1 - O(2^{10\ell-k})$.

Finally, taking a union bound over zoom-out and zoom-in we see S is $(1, O(2^{-\ell/2}))$ pseudo-random with probability $1 - O(2^{10\ell-k})$.

S has weight $\Omega(\delta^2)$ on the first level with probability $\Omega(\delta^2)$. Recall that the weight of S on the first level equals (Claim 6.2)

$$\mathbb{E}\left[W^{=1}[S]\right]\begin{bmatrix}\ell\\1\end{bmatrix}\mathop{\mathbb{E}}_{v}\left[(\mu(S_v)-\mu(S))^2\right]-o(1).$$

Fix $v_i \in V$. Then

$$\begin{aligned}\mathop{\mathbb{E}}_{X_{v_1},\ldots,X_{v_{2^k}}}\left[(\mu(S_v)-\mu(S))^2\right] &\geqslant \mathop{\mathbb{E}}_{X_{v_i}}\left[\mathop{\mathbb{E}}_{X_{v_j},j\neq i}\left[\mu(S_v)-\mu(S)\right]^2\right]\\ &= \Pr\left[X_{v_i}=1\right]\Omega(2^{-\ell}\delta^2)\\ &= \Omega(2^{-\ell}\delta^2).\end{aligned}$$

The first inequality is by $\mathbb{E}\left[Z^2\right] \geqslant \mathbb{E}\left[Z\right]^2$. The second inequality one is by Proposition C.3 and the fact that $\mu(S) = \mu + O(2^{-\ell})$ with probability $1-o(1)$ conditioned on $X_{v_i}=1$ (this is true since we proved $\mu(S)=\mu+O(2^{-\ell})$ with probability $1-o(1)$ and we condition on event of probability $\frac{1}{2}$).

Therefore, since $\begin{bmatrix}\ell\\1\end{bmatrix} = 2^\ell - 1$, we get that

$$\mathbb{E}\left[W^{=1}[S]\right] \geqslant \Omega(\delta^2).$$

Since $W^{=1}[S] \leqslant \mu(S) \leqslant 1$, we have by an averaging argument that $W^{=1}[S] \geqslant c'\delta^2$ with probability $\Omega(\delta^2)$ (for explicit c').

Wrapping things up. Let G_1 be the event $\mu(S) = \Theta(\delta)$, G_2 the event S is $(1, O(2^{-\ell/2}))$ pseudo-random and G_3 the event $W^{=1}[S] \geqslant c'\delta^2$ for the an explicit constant. Then in the above we have seen $\Pr[G_1], \Pr[G_2] \geqslant 1-o(1)$ while $\Pr[G_3] \geqslant \frac{1}{4}\delta^2$, and therefore

$$\Pr\left[G_1 \cap G_2 \cap G_3\right] \geqslant \Omega(\delta^2) - o(1) > 0,$$

and in particular there exists a choice of S as desired.

D Proof of Theorem 1.3

Fact D.1. *Let $\delta > 0$, k, ℓ be integers and V a k-dimensional space over $\mathbb{F}_2$. For sufficiently large k, ℓ, if S is a set of vertices in $G(V,\ell)$ of density δ, then $\Phi(S) \geqslant \frac{1}{2} - \frac{1}{2}\delta$.*

Proof. Let F be the indicator function of S. Then

$$\Phi(S) = 1 - \frac{1}{\delta}\langle F, A_{G(V,\ell)}F\rangle. \tag{28}$$

Writing the spectral decomposition $F = F_{=0} + \ldots + F_{=\ell}$ and plugging it into the inner product, we see that

$$\langle F, A_{G(V,\ell)}F\rangle = \sum_{i=0}^{\ell}\lambda_i\langle F_{=i}, F_{=i}\rangle \leqslant \langle F_{=0}, F_{=0}\rangle + \lambda_1\sum_{i=1}^{\ell}\langle F_{=i}, F_{=i}\rangle.$$

In the last inequality we used $\lambda_0 = 1$ and $\lambda_i \leqslant \lambda_1$ for all $i \geqslant 1$ (Fact 2.15). Notice that since $F_{=0} \equiv \delta$ and $\lambda_1 \leqslant \frac{1}{2}$ (Fact 2.14), we get that

$$\langle F, A_{G(V,\ell)} F \rangle \leqslant \delta^2 + \frac{1}{2}(\sum_{i=1}^{\ell} \langle F_{=i}, F_{=i} \rangle) = \delta^2 + \frac{1}{2}(\sum_{i=0}^{\ell} \langle F_{=i}, F_{=i} \rangle - \delta^2) = \delta^2 + \frac{1}{2}(\|F\|_2^2 - \delta^2).$$

The last equality is by Parseval. Since $\|F\|_2^2 = \delta$, we conclude that

$$\langle F, A_{G(V,\ell)} F \rangle \leqslant \frac{1}{2}(\delta^2 + \delta).$$

Plugging this into Equation 28 finishes the proof. □

Theorem 1.3 (Restated) . *For every $\delta > \frac{1}{2}$ there exists $\varepsilon > 0$ such that the following holds for sufficiently large k, ℓ. If $\mathcal{F}$ is a labeling of $G(V, \ell)$ by linear functions that has δ-consistency in the Grassmann test, then there exists a linear function $\mathcal{H} \colon V \to \mathbb{F}_2$ such that*

$$\Pr_L [\mathcal{F}[L] \equiv \mathcal{H}|_L] \geqslant \varepsilon.$$

Proof. Define $\varepsilon > 0$ by $\delta = \frac{1}{2} + \varepsilon$, and fix $b = \lfloor \ell/10 \rfloor$. For any $B \in \begin{bmatrix} V \\ b \end{bmatrix}$, define

$$S[B] = \left\{ L \in \begin{bmatrix} V \\ \ell \end{bmatrix} \;\middle|\; B \subseteq L \right\}.$$

Note that the following distribution over edges is uniform: sample $B \in \begin{bmatrix} V \\ b \end{bmatrix}$, then sample $L, L' \in S[B]$ conditioned on $\dim(L \cap L') = \ell - 1$. Therefore, the expected fraction of edges satisfied by $\mathcal{F}$ inside $S[B]$ is δ. By an averaging argument, there is a set $\mathcal{B} \subseteq \begin{bmatrix} V \\ b \end{bmatrix}$ of relative size at least $\frac{1}{2}\varepsilon$, such that at least $\frac{1}{2} + \frac{1}{2}\varepsilon$ fraction of the edges are satisfied inside $S[B]$ for each $B \in \mathcal{B}$.

Fix $B \in \mathcal{B}$. Partition $S[B]$ into 2^b parts, according to the value $\mathcal{F}[L]|_B$. Namely, let $g_1, ..., g_{2^b}$ be all 2^b linear functions $g_i \colon B \to \{0, 1\}$, and define

$$P_i[B] = \{ L \in S[B] \mid \mathcal{F}[L]|_B \equiv g_i \}.$$

We claim there exists i such that $P_i[B]$ contains at least ε fraction of the spaces in $S[B]$. Assume towards contradiction this is not the case. Note that the induced subgraph on $S[B]$ is isomorphic to $G(W, \ell - b)$ for some W of dimension $k - b$. Since the density of each $P_i[B]$ is strictly smaller than ε in this induced subgraph, we have $\Phi(P_i[B]) > \frac{1}{2} - \frac{1}{2}\varepsilon$ by Fact D.1.

Note that any edge going outside of $P_i[B]$ is not satisfied by $\mathcal{F}$ (since the labels of the endpoints do not agree on B). Since $P_i[B]$ for $i = 1, ..., 2^b$ cover the subgraph, we conclude that all edges go out of one of them. Combining all of the above, we see that less than $\frac{1}{2} + \frac{1}{2}\varepsilon$ fraction of the edges in $S[B]$ are satisfied, and contradiction.

Therefore, there exists i_B such that $|P_{i_B}[B]| \geqslant \varepsilon |S[B]|$. Define the assignment $\mathcal{P}$ on $\begin{bmatrix} V \\ b \end{bmatrix}$ that assigns each space a linear function on it, by $\mathcal{P}[B] = g_{i_B}$ for $B \in \mathcal{B}$ and else arbitrarily. Then clearly

$$\Pr_{\substack{B \in \begin{bmatrix} V \\ b \end{bmatrix}, L \in \begin{bmatrix} V \\ \ell \end{bmatrix} \\ B \subseteq L}} [\mathcal{F}[L]|_B \equiv \mathcal{P}[B]] \geqslant \Pr[B \in \mathcal{B}] \Pr_{\substack{B \in \mathcal{B}, L \in \begin{bmatrix} V \\ \ell \end{bmatrix} \\ B \subseteq L}} [\mathcal{F}[L]|_B \equiv \mathcal{P}[B]] \geqslant \frac{1}{2}\varepsilon \cdot \varepsilon = \frac{1}{2}\varepsilon^2.$$

By [KMS16, Theorem D.1] (for large enough ℓ), we conclude there exists a linear function $g \colon V \to \{0, 1\}$, such that

$$\Pr_{L \in \begin{bmatrix} V \\ \ell \end{bmatrix}} [\mathcal{F}[L] \equiv g|_L] \geqslant \frac{\varepsilon^6}{2^{12}},$$

as desired. □

E Missing Proofs

E.1 Proof of Lemma 7.4

By definition

$$\begin{aligned}\widehat{f}_{\approx 2}(S_1,S_2) &= \mathop{\mathbb{E}}_{x,y}\left[f_{\approx 2}(x,y)\chi_{S_1}(x)\chi_{S_2}(y)\right]\\ &= \frac{1}{4}\left(\mathop{\mathbb{E}}_{\substack{x\in W_{S_1}\\ y\in W_{S_2}}}[f_{\approx 2}(x,y)] - \mathop{\mathbb{E}}_{\substack{x\notin W_{S_1}\\ y\in W_{S_2}}}[f_{\approx 2}(x,y))] - \mathop{\mathbb{E}}_{\substack{x\in W_{S_1}\\ y\notin W_{S_2}}}[f_{\approx 2}(x,y)] + \mathop{\mathbb{E}}_{\substack{x\notin W_{S_1}\\ y\notin W_{S_2}}}[f_{\approx 2}(x,y)]\right).\end{aligned}$$

Using the second item in Claim 7.2, we conclude that the sum of any two consecutive expectations is 0; hence

$$\widehat{f}_{\approx 2}(S_1,S_2) = \mathop{\mathbb{E}}_{\substack{x\notin W_{S_1}\\ y\notin W_{S_2}}}[f_{\approx 2}(x,y)].$$

Denote

$$U \stackrel{def}{=} \{(x,y) \mid x\notin W_{S_1}, y\notin W_{S_2}\}.$$

Since the probability x, y are linearly dependent is at most $O(2^{-k})$, we have that

$$\mathop{\mathbb{E}}_{\substack{x\notin W_{S_1}\\ y\notin W_{S_2}}}[f_{\approx 2}(x,y)] = \mathop{\mathbb{E}}_{(x,y)\in_R U}\left[\mu_{\mathsf{Span}(x,y),\mathsf{in}} - \mu_{\mathsf{Span}(x),\mathsf{in}} - \mu_{\mathsf{Span}(y),\mathsf{in}} - \mu_{\mathsf{Span}(x+y),\mathsf{in}} + 2\mu\right] + O(2^{-k}\|F\|_\infty).$$

We now use linearity of expectation.

Proposition E.1. *The distribution of $x+y$ is uniform over V when $(x,y)\in_R U$.*

Proof. Sample $(x,y)\in_R U$ With probability $\frac{1}{2}$, $x\in W_{S_2}$, in which case it is easy to see $x+y$ is distributed uniformly outside W_{S_2}. With probability $\frac{1}{2}$, $x\notin W_{S_2}$ in which case $x+y$ is distributed uniformly over W_{S_2}. □

Thus, the expectation of $\mu_{\mathsf{Span}(x+y),\mathsf{in}}$ is μ, and the last expectation equals

$$\mathop{\mathbb{E}}_{(x,y)\in_R U}[f_{\approx 2}(x,y)] = \mathop{\mathbb{E}}_{(x,y)\in_R U}\left[(\mu_{\mathsf{Span}(x,y),\mathsf{in}} - \mu) - (\mu_{\mathsf{Span}(x),\mathsf{in}} + \mu_{\mathsf{Span}(y),\mathsf{in}} - 2\mu)\right], \tag{29}$$

which is $\mathbb{E}_{(x,y)\in_R U}\left[(\mu_{\mathsf{Span}(x,y),\mathsf{in}} - \mu) - (f_{\approx 1}(x) + f_{\approx 1}(y))\right]$ by the definition of $f_{\approx 1}$.

By the proof of Lemma 4.2,

$$-\mathop{\mathbb{E}}_{x\notin W_{S_1}}[f_{\approx 1}(x)] = \frac{\begin{bmatrix}k-1\\ \ell\end{bmatrix}}{\begin{bmatrix}k\\ \ell\end{bmatrix} - \begin{bmatrix}k-1\\ \ell\end{bmatrix}}(\mu_{W_{S_1},\mathsf{out}} - \mu),$$

and thus the above combined mean that

$$\widehat{f}_{\approx 2}(S_1,S_2) = \mathop{\mathbb{E}}_{(x,y)\in_R U}\left[\mu_{\mathsf{Span}(x,y),\mathsf{in}} - \mu\right] + \frac{\begin{bmatrix}k-1\\ \ell\end{bmatrix}}{\begin{bmatrix}k\\ \ell\end{bmatrix} - \begin{bmatrix}k-1\\ \ell\end{bmatrix}}(\mu_{W_{S_1},\mathsf{out}} + \mu_{W_{S_2},\mathsf{out}} - 2\mu) + O(2^{-k}\|F\|_\infty). \tag{30}$$

We compute the first expectation. Define $U' = \{\,(x,y) \in U \mid x \neq y\}$. Since the probability $x = y$ when $(x,y) \in U$ is $O(2^{-k})$, we have that

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\in_R U}\left[\mu_{\mathsf{Span}(x,y),\mathsf{in}} - \mu\right] &= \mathop{\mathbb{E}}_{(x,y)\in_R U'}\left[\mu_{\mathsf{Span}(x,y),\mathsf{in}} - \mu\right] + O(2^{-k}\|F\|_\infty) \\ &= \mathop{\mathbb{E}}_{(x,y)\in_R U'}\left[\mathop{\mathbb{E}}_{L\supseteq \mathsf{Span}(x,y)}\left[F[L] - \mu\right]\right] + O(2^{-k}\|F\|_\infty)\end{aligned}$$

We now consider the resulting distribution on L. First, note that each L supported on that distribution is not contained in W_{S_1} and in W_{S_2}. In particular, $\mathsf{dim}(L \cap W_{S_1} \cap W_{S_2})$ can be $\ell - 1$, in which case $L \subseteq W_{S_1 \oplus S_2}, L \not\subseteq W_{S_1}$ [12], or else $\mathsf{dim}(L \cap W_{S_1} \cap W_{S_2}) = \ell - 2$. We now proceed to calculate the precise weight given to each L in the last expectation; clearly, it is proportional to the number of tuples (x,y) in U' so that $x, y \in L$. This number is ($\overline{W} \stackrel{def}{=} V \setminus W$)

$$\begin{aligned}&\left(\left|L \cap \overline{W_{S_1}}\right| - \left|L \cap \overline{W_{S_1}} \cap \overline{W_{S_2}}\right|\right)\left|L \cap \overline{W_{S_2}}\right| \\ &+ \left|L \cap \overline{W_{S_1}} \cap \overline{W_{S_2}}\right|\left(\left|L \cap \overline{W_{S_2}}\right| - 1\right),\end{aligned}$$

the first line corresponds to choosing $x \in W_{S_2} \setminus W_{S_1}$, and the second corresponds to choosing x outside both W_{S_1}, W_{S_2} (this reduces the options for y by 1). Thus, the weight of L is proportional to

$$\rho(L) = \begin{cases} 2^{2\ell-2} - 2^{\ell-1} & L \subseteq W_{S_1 \oplus S_2}, L \not\subseteq W_{S_1}, \\ 2^{2\ell-2} - 2^{\ell-2} & \mathsf{dim}(L \cap W_{S_1} \cap W_{S_2}) = \ell - 2, \\ 0 & \text{else.}\end{cases}$$

And to get the weight of L, $\rho(L)$ should be divided by

$$W = \left|U'\right| \begin{bmatrix} k-2 \\ \ell-2 \end{bmatrix} = (2^{2k-2} - 2^{k-2}) \begin{bmatrix} k-2 \\ \ell-2 \end{bmatrix}.$$

Thus,

$$\mathop{\mathbb{E}}_{\substack{(x,y)\in_R U \\ L\supseteq \mathsf{Span}(x,y)}}\left[F[L] - \mu\right] = \frac{1}{W}\sum_{L : w(L) > 0} \rho(L)(F[L] - \mu).$$

We manipulate the last sum. We count all L-spaces $2^{2\ell-2} - 2^{\ell-2}$ times and subtract L that were overcounted due to that to get

$$\begin{aligned}&= \frac{2^{2\ell-2} - 2^{\ell-2}}{W}\sum_{L}(F[L] - \mu) - \frac{2^{2\ell-2} - 2^{\ell-2}}{W}\sum_{L \subseteq W_{S_1} \cap W_{S_2}}(F[L] - \mu) - \frac{2^{\ell-2}}{W}\sum_{\substack{L \subseteq W_{S_1 \oplus S_2} \\ L \not\subseteq W_{S_1} \cap W_{S_2}}}(F[L] - \mu) \\ &- \frac{2^{2\ell-2} - 2^{\ell-2}}{W}\sum_{\substack{L \subseteq W_{S_1} \\ L \not\subseteq W_{S_1} \cap W_{S_2}}}(F[L] - \mu) - \frac{2^{2\ell-2} - 2^{\ell-2}}{W}\sum_{\substack{L \subseteq W_{S_1} \\ L \not\subseteq W_{S_1} \cap W_{S_2}}}(F[L] - \mu).\end{aligned}$$

[12] In this case there must be $W \supseteq W_{S_1} \cap W_{S_2}$ of co-dimension 1 containing L, but there are only 3 co-dimensional 1 spaces containing $W_{S_1} \cap W_{S_2}$, namely $W_{S_1}, W_{S_2}, W_{S_1 \oplus S_2}$.

The first sum is 0 since μ is the average of F over all L. For the 3 last sums, remove the restriction $L \not\subseteq W_{S_1} \cap W_{S_2}$ and compensate for that in the coefficient of the second sum to get

$$\begin{aligned}
&= \frac{2^{2\ell-2}}{W} \sum_{L \subseteq W_{S_1} \cap W_{S_2}} (F[L] - \mu) - \frac{2^{\ell-2}}{W} \sum_{L \subseteq W_{S_1 \oplus S_2}} (F[L] - \mu) \\
&- \frac{2^{2\ell-2} - 2^{\ell-2}}{W} \sum_{L \subseteq W_{S_1}} (F[L] - \mu) - \frac{2^{2\ell-2} - 2^{\ell-2}}{W} \sum_{L \subseteq W_{S_1}} (F[L] - \mu).
\end{aligned}$$

We turn those into expectations to get

$$\begin{aligned}
&= \frac{2^{2\ell-2} \begin{bmatrix} k-2 \\ \ell \end{bmatrix}}{W} (\mu_{W_{S_1} \cap W_{S_2}, \mathsf{out}} - \mu) \\
&- \frac{2^{\ell-2} \begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{W} (\mu_{W_{S_1 \oplus S_2}, \mathsf{out}} + \mu_{W_{S_1}, \mathsf{out}} + \mu_{W_{S_2}, \mathsf{out}} - 3\mu) \\
&- \frac{(2^{2\ell-2} - 2^{\ell-1}) \begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{W} (\mu_{W_{S_1}, \mathsf{out}} + \mu_{W_{S_2}, \mathsf{out}} - 2\mu).
\end{aligned}$$

Plugging this into Equation (30) we conclude that

$$\begin{aligned}
\widehat{f}_{\approx 2}(S_1, S_2) = &\frac{2^{2\ell-2} \begin{bmatrix} k-2 \\ \ell \end{bmatrix}}{W} (\mu_{W_{S_1} \cap W_{S_2}, \mathsf{out}} - \mu) \\
&- \frac{2^{\ell-2} \begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{W} (\mu_{W_{S_1 \oplus S_2}, \mathsf{out}} + \mu_{W_{S_1}, \mathsf{out}} + \mu_{W_{S_2}, \mathsf{out}} - 3\mu) \\
&+ \left(\frac{\begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix} - \begin{bmatrix} k-1 \\ \ell \end{bmatrix}} - \frac{(2^{2\ell-2} - 2^{\ell-1}) \begin{bmatrix} k-1 \\ \ell \end{bmatrix}}{W} \right) (\mu_{W_{S_1}, \mathsf{out}} + \mu_{W_{S_2}, \mathsf{out}} - 2\mu) + O(2^{-k} \|F\|_\infty)
\end{aligned}$$

The last step in the proof is to estimate the coefficients above. Plugging in W and doing basic manipulations with Gaussian coefficients shows that the coefficients in the third line is $O(2^{\ell-k})$, that the difference between the first two is $O(2^{\ell-k})$, and that the second is $O(2^{\ell-k})$-close to $c_2(k,\ell) = \frac{\begin{bmatrix} k-2 \\ \ell \end{bmatrix}}{\begin{bmatrix} k \\ \ell \end{bmatrix} - 3\begin{bmatrix} k-1 \\ \ell \end{bmatrix} + 2\begin{bmatrix} k-1 \\ \ell-1 \end{bmatrix}}$. Hence we get

$$\widehat{f}_{\approx 2}(S_1, S_2) = c_2 \left[\mu_{W_{S_1} \cap W_{S_2}, \mathsf{out}} - \mu - (\mu_{W_{S_1 \oplus S_2}, \mathsf{out}} + \mu_{W_{S_1}, \mathsf{out}} + \mu_{W_{S_2}, \mathsf{out}} - 3\mu) \right] + O(2^{\ell-k} \|F\|_\infty)$$

E.2 Proof that Theorem 7.17 implies Theorem 3.7

Suppose F satisfying all the conditions, and let $\varepsilon \geqslant \delta$ be infimum of all ε such that F is $(2, \varepsilon)$ pseudo-random. We show that

$$\varepsilon \geqslant 2^{-51} \left(\frac{\eta}{\delta} \right)^3.$$

Suppose we showed that. Then either $\varepsilon = \delta$ and thus $\delta \geqslant 2^{-51} \left(\frac{\eta}{\delta} \right)^3$ - which cannot happen by the assumption on η. Otherwise $\varepsilon > \delta$ and therefore F is not $(2, 2^{-52} \left(\frac{\eta}{\delta} \right)^3)$ pseudo-random (minimality of ε).

We now show the lower bound on ε. By Theorem 7.17, we have that

$$\eta = W^{=2}[F] \leqslant 2^{17} \delta \varepsilon^{1/3},$$

rearranging yields leads to the desired lower bound on ε.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il