ON THE SECOND EIGENVALUE OF HYPERGRAPHS

Joel Friedman
Avi Wigderson

CS-TR-232-89

November 1989

On the Second Eigenvalue of Hypergraphs

Joel Friedman*
Dept. of Comp. Sci.
Princeton University
Princeton, NJ 08544

Avi Wigderson
Dept. of Comp. Sci.
Hebrew University
Jerusalem 91904
ISRAEL

November 7, 1989

1 The Second Eigenvalue of a Hypergraph

In this section we will first define the second eigenvalue of 3-uniform hypergraph, and then discuss the general notion, as it applies to other uniform hypergraphs and graphs. To motivate our definition, notice that if G is an undirected d-regular graph, i.e. each vertex has degree d, then the second largest eigenvalue in absolute value, λ_2, of G's adjacency matrix, A, satisfies

$$|\lambda_2| = \left\| A - \frac{d}{n} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \right\|_{L^2(V)},$$

where $n = |V|$.

Let $G = (V, E)$ be a 3-uniform hypergraph; i.e. E is a subset of subsets of V of size 3. We consider the space, $L^2(V)$, of real valued functions on V with the usual inner product; let $e_1, \ldots, e_n$ be the standard basis for $L^2(V)$, where e_i takes the value 1 on the i-th vertex of V and 0 elsewhere. It is natural to construct from G a trilinear form, that is a map from τ, mapping triples of vectors in $L^2(V)$ to $\mathbf{R}$, namely

$$\tau\left(\sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \beta_j e_j, \sum_{k=1}^{n} \gamma_k e_k\right) = \sum_{i,j,k} \alpha_i \beta_j \gamma_k \tau_{i,j,k}$$

*The author wishes to acknowledge the National Science Foundation for supporting this research in part under Grant CCR–8858788, and the Office of Naval Research under Grant N00014–87–K–0467.

where

$$\tau_{i,j,k} = \begin{cases} 1 & \text{if } \{i,j,k\} \in E \\ 0 & \text{otherwise} \end{cases}$$

Let $\vec{1}$ denote the all 1's vector, and let $\vec{1} \otimes \vec{1} \otimes \vec{1}$ denote the all 1's trilinear form (i.e. the form ν defined as τ but with $\nu_{i,j,k} = 1$ for all i, j, k). We define τ to be d-regular if

$$\sigma = \tau - \frac{d}{n} \vec{1} \otimes \vec{1} \otimes \vec{1}$$

has the property that

$$\sigma(\vec{1}, u, v) = \sigma(u, \vec{1}, v) = \sigma(u, v, \vec{1}) = 0$$

for all u and v. In this case we define the *second eigenvalue* of G (and τ) to be

$$\lambda_2 = \|\sigma\|_{L^2(V)} = \sup_{\|u\|=\|v\|=\|w\|=1} |\sigma(u,v,w)|.$$

We pause for a few remarks. First of all, for simplicity we have abused some of the usual tensor conventions; in particular, we will often identify spaces and elements of spaces with their duals implicitly when the distinction serves no purpose and/or needlessly complicates the notation or discussion. For example, the all 1's trilinear form is really the dual of $\vec{1} \otimes \vec{1} \otimes \vec{1}$ in the usual conventions.

Second, just as ordinary directed graphs can have multiple edges and self-loops, we can accomodate such notions here. If an edge $\{i, j, k\}$ occurs with multiplicity m, then the six corresponding entries of τ will have an m instead of a 1. Also, an edge which is a self-loop, $\{i, i, i\}$, should contribute 6 to $\tau_{i,i,i}$, on the principle that the contribution should be determined so that each edge contributes a total of 6 to all the entries of τ. Similarly an edge, $\{i, i, j\}$ should contribute 2 to each of $\tau_{i,i,j}$, $\tau_{i,j,i}$, and $\tau_{j,i,i}$.

Thirdly, we can also handle the notion of directed edges. We say that a directed 3-uniform hypergraph is a hypergraph where each edge $\{i, j, k\}$ has a specified order; equivilently, E is a subset of $V \times V \times V$. Now the trilinear form τ is defined by $\tau_{i,j,k}$ being the multiplicity of the edge (i, j, k). From a directed hypergraph we can form an undirected hypergraph, i.e. the usual notion of hypergraph, by forgetting about the ordering on the edges. The new trilinear form is the symmetrized version of the old one. This also explains why undirected degenerate edges such as $\{i, i, i\}$ should contribute 6 to $\tau_{i,i,i}$ in the usual (undirected) notion of hypergraph. Identifying a hypergraph with its trilinear form, we can think of undirected hypergraphs as special cases of directed hypergraphs.

Fourth, we define for any trilinear form on $L^2(V)$, μ, its first eigenvalue to be its norm with respect to $L^2(V)$, i.e.

$$\lambda_1(\mu) = \|\mu\|_{L^2(V)} = \sup_{\|u\|=\|v\|=\|w\|=1} |\mu(u,v,w)|.$$

If u, v, w is a triple achieving the above sup, we shall call (u, v, w) or $u \times v \times w$ an *eigenvector*. We can also define d-regular for a general trilinear form as above, and therefore extend the notion of second eigenvalue to directed hypergraphs. It is easy to check that an undirected hypergraph is d-regular if for every vertices i, j, there are exactly d vertices k for which $\{i, j, k\}$ is an edge (with the case $i = j$ included); similarly, a directed hypergraph is d-regular if for every vertices i, j there are d edges of each of the forms $(i, j, \cdot)$, $(i, \cdot, j)$, and $(\cdot, i, j)$.

Fifth, it becomes clear how we want to define the second eigenvalue in the more general case. For any sets, $V_1, \ldots, V_k$, we define the first eigenvalue of a k-linear form mapping $V_1 \times \cdots \times V_k$ to $\mathbf{R}$ as its norm with respect to the norms $L^2(V_i)$. A k-linear form, τ, is d-regular if

$$\sigma = \tau - \frac{d}{n} \vec{1}_{V_1} \otimes \cdots \otimes \vec{1}_{V_k},$$

with n being the largest among the $|V_i|$'s, satisfies

$$\sigma(v_1, \ldots, v_k) = 0$$

whenever some v_i is $\vec{1}_{V_i}$. In this case we define the second eigenvalue of τ to be $\lambda_1(\sigma)$. Notice that for a directed graph, with adjacency matrix A, which is d-regular, it's second eigenvalue in the new sense is the square root of the (classical) second largest eigenvalue of AA^T.

Sixth, the first and second eigenvalue of multilinear forms as defined above are always non-negative numbers. While much of the literature on graphs considers the second eigenvalue as the eigenvalue with the second largest absolute value, sometimes one simply considers the second largest (positive) eigenvalue. Here the distinction is lost.

Lastly, we could define the eigenvalues and/or norms of multilinear forms with respect to other norms on the space of functions on V, such as $L^p(V)$ for any $1 \leq p \leq \infty$. In the applications we have in mind, it does not seem to help to consider the other norms; and, in fact, one can say more about the $L^2(V)$ norms than for other values of p.

We return to the discussion of 3-uniform hypergraphs. There is a lot of structure for graphs that we don't know how to carry over to 3-uniform hypergraphs. For example, we don't know how multiply two hypergraphs. However, a fair amount of the theory for graphs can be generalized to 3-uniform hypergraphs. We will prove:

Theorem 1.1 *The first eigenvalue of a d-regular 3-uniform hypergraph is $dn^{1/2}$; i.e. the associated trilinear form τ is maximized on $(\vec{1}, \vec{1}, \vec{1})$.*

Theorem 1.2 *A symmetric trilinear form τ has a unit vector v such that $\|\tau\| = |\tau(v, v, v)|$.*

Theorem 1.3 *Every 3-uniform, d-regular hypergraph has second eigenvalue $\geq \sqrt{d(n-d)/n}$.*

Next we'd like to prove a theorem about random hypergraphs. For graphs there are probability spaces of random d-regular graphs that are easy to work with. For hypergraphs the situation is more complicated. It becomes convenient to weaken the requirement of d-regularity. For any trilinear form, τ, we define its *second eigenvalue with respect to d-regularity* to be

$$\lambda_{2,d}(\tau) \equiv \|\tau - \frac{d}{n}\vec{1} \otimes \vec{1} \otimes \vec{1}\|.$$

Theorem 1.4 *For any C there is a C' such that for any n and $d \geq C' \log n$, a random 3-uniform hypergraph on n vertices with dn^2 edges chosen randomly has second eigenvalue $\leq C'(\log n)^{3/2}\sqrt{d}$ with probability $\geq 1 - n^{-C\min(\sqrt{d}, n^{1/6})}$.*

All of the above theorems have generalizations to t-uniform hypergraphs. Theorems 1.1, 1.2, and 1.3 hold as is for t-uniform hypergraphs with $t > 3$. The analog of theorem 1.4 for $t \geq 3$ is:

Theorem 1.5 *For any n, $d \geq Ct \log n$, a random t-uniform hypergraph on n vertices with dn^{t-1} edges chosen randomly has second eigenvalue $\leq (C \log n)^{t/2}\sqrt{d}$ with probability at least $1 - n^{-C' min(\sqrt{d}, n^{1/(t+3)})/t}$.*

We remark that presumably with more work, one could replace $t/2$ in the above exponent of $\log n$ in the eigenvalue estimate with $(t-2)/2$ (see the end of section 6). In our applications the $\log n$ factors are unimportant.

We do not know, at present, how to explicitly construct hypergraphs to match the bounds in theorem 1.5. However, it is easy to construct graphs with λ_1 and λ_2 separated. Let $G = (V, E)$ be any graph which is d-regular, and let $\widetilde{G} = (V, V^{t-2} \times E)$. It is easy to see that $\lambda_2(\widetilde{G}) \leq n^{(t-2)/2}\lambda_2(G)$.

To achieve smaller λ_2, we generalize the notion of Cayley graphs to hypergraphs. Let G be a finite group, and H a symmetric set of generators. Then the t-uniform Cayley hypergraph on G and H, has vertex set G and edges

$$E = \{(x_1, \ldots, x_t) : x_1 \cdots x_t \in H\}.$$

We can calculate the eigenvalues of the t-uniform Cayley hypergraph from those of the Cayley graph and a knowledge of the decomposition of $L^2(G)$ under the right regular representation. While most explicit constructions of graphs with essentially optimal λ_2 are Cayley graphs, λ_2 of Cayley hypergraphs are far from optimal:

Theorem 1.6 *If $|G| = n$ and $|H| = d$, then the t-uniform Cayley hypergraph on G and H has $\lambda_2 \geq n^{(t-2)/4}\sqrt{d(n-d)/n}$.*

Friedman has recently (see [Fri89]) constructed Cayley hypergraphs that nearly achieve the above lower bound— let AFFINE(p) be the group of affine transformations on $\mathbf{Z}/p\mathbf{Z}$, and $H = \{r^2x \pm r) : r \in (\mathbf{Z}/p\mathbf{Z})^*\}$. Then $n = p(p-1)$ and $d = 2(p-1)$. For the associated hypergraph we have:

Theorem 1.7 ([Fri89]) *For the above hypergraph,* $\lambda_2 = 2p^{(t-1)/2}$.

In sections 2 and 3 we discuss an application of the notion of eigenvalue of a hypergraph, based on a hope of finding explicit constructions to achieve the bounds in theorem 1.4. In section 4 we continue the technical discussion, giving some simple properties of hypergraphs. In section 5 we discribe the "Cayley hypergraphs," and give some bounds and mention one construction. In section 6 we prove theorem 1.4 and its generalization, theorem 1.5.

2 Dispersers and Randomness

Let $\Sigma = \{0,1\}$. For a language $L \in \Sigma^*$, let $L_n = L \cap \Sigma^n$. For a function $\alpha: N \to [0,1]$, we define the class $RATE(\alpha)$ by $L \in RATE(\alpha)$ iff for all (large) n,

$$\frac{\log |L_n|}{n} \leq \alpha(n).$$

Let $k, m: N \to N$ be functions, with $k(n), m(n) = n^{O(1)}$. We shall be interested in mappings $\mu : \Sigma^* \to (\Sigma^*)^k$ that for each n, $\mu : \Sigma^{m(n)} \to (\Sigma^n)^k$. For a language L and a mapping μ, let

$$\mu^{-1}(L) = \{y : \mu(y)_i \in L \text{ for all } 1 \leq i \leq k\}.$$

We say that μ is *efficient* if μ can be computed in determinsitic poly-time.

We say that μ is a (universal) (α, β) *rate disperser* if $\forall L \in RATE(\alpha)$, $\mu^{-1}(L) \in RATE(\beta)$. (Note that a diperser acts as the "inverse" of compression, i.e. strives to make languages sparser.)

There are several ways in which dispersers are intimately realted to randomized computing, and they all stem from a fundemantal observation of Sipser [Sip86]. Think of L as the language of non-witnesses of some RP algorithm[1]. Then an efficient disperser μ defines a new RP algorithm, with a language of non-witnesses $\mu^{-1}(L)$, which is much sparser.

Typically, $L \in RATE(1 - 1/n)$ (i.e. error probability bounded by $1/2$). Sipser [Sip86] and Santha [San86] also gave a probabilistic argument showing that most mappings μ are excellent dispersers.

Theorem 2.1 ([Sip86,San86]) $\forall \epsilon > 0$ *there exist* $(1 - 1/n, n^{\epsilon - 1})$ *dispersers.*

With our restriction that $m(n), k(n) = n^{O(1)}$, [CW89] observed that the above theorem is the best possible.

Theorem 2.2 ([CW89]) *There are no* $(1 - 1/n, n^{o(1)-1})$ *dispersers.*

The three motivations for *constructing* efficient dispersers, and the state of art regarding them are summarized below (see also [CW89]).

[1] It is natural, and discussed at length in [CW89,IZ89] to consider dispersers for BPP. We do not pursue it here

2.1 Determinsitic Amplification

The first explicit construction of a disperser is implicit in [AKS87].

Theorem 2.3 ([AKS87]) *For every $k = k(n)$, there is an efficient $(1 - 1/n, 1 - k/m)$ disperser, with $m(n) = n + O(k(n))$. (In particular, for some $\epsilon > 0$, efficient $(1 - 1/n, 1 - \epsilon)$ dispersers exist.)*

By the discussion above, this immediately implies:

Corollary 2.4 ([CW89]) *The error probability of any RP algorithm can be d ecreased from 1/2 to 2^{-k} at the cost of only $O(k)$ additional random bits.*

The disperser construction of [AKS87] is based on random walks on expander graphs. More elementary, though weaker, dispersers were later constructed in [IZ89].

2.2 Randomness vs. Determinism or Time vs. Space

That efficient dispersers can relate these two fundamental questions of computer science, is the main issue of Sipser's paper [Sip86].

Theorem 2.5 ([Sip86]) *Assume that for some $\epsilon > 0$, efficient $(1 - 1/n, n^{-\epsilon})$ dispersers exist. Then at least one of the following statements hold:*

1. *$RP = P$.*
2. *For some $\epsilon' > 0$, every $t(n) > 2^n$ and infinitely many n, $DTIME(t) \subseteq DSPACE(t^{1-\epsilon'})$.*

As will become clear later, we are far from constructing such good dispersers as this theorem requires, and indeed we offer a way to go about it. However, [NW88] proved by a different mehod that a similar consequence can be proven without any assumptions if we relax (1).

Theorem 2.6 ([NW88]) *One of the following statements hold*

1. *$RP = \bigcap_{\epsilon} DTIME(2^{n^{\epsilon}})$.*
2. *For some $\epsilon' > 0$, $t(n) > 2^n$ and infinitely many n, $DTIME(t) \subseteq DSPACE(t^{1-\epsilon'})$.*

2.3 Weak Random Sources

An *ensemble* Π is a family of probability distibutions $\Pi = (\Pi_1, \Pi_2, \ldots)$ with Π_n a distribution of Σ_n. A *source* S is a family of ensembles Π.

A probabilistic algorithm works with source S if on input x it computes a length $l = l(x)$, receives (w.l.o.g) a *single*[2] random string $\sigma \in \Pi_l$ from some ensemble $\Pi \in S$, and from that point proceeds deterministically. The class $RP(S)$ is the class of languages that are recognised by probabilistic poly-time algorithms working with S. Thus, if *UNIFORM* is the source comprised of the single ensemble $U = (U_1, U_2, \ldots)$, with U_n being the uniform distribution on Σ^n, then $RP = RP(UNIFORM)$.

Much literature in recent years, e.g.[Blu84,SV84,VV85,CG85,CW89] is devoted to proving $RP = RP(S)$ for a variety of sources S, mostly sources with bounded entropy rate plus structural restrictions.

Say that an ensemble Π is α-*FAIR*[3] if for every n, every string $\sigma \in \Sigma^n$ has probability $\leq 2^{-\alpha n}$. S is α-*FAIR* if every $\Pi \in S$ is α–FAIR, and S is *FAIR* if it is α–FAIR for some $\alpha > 0$. FAIR sources include most of the sources studied in the literature, and many more.

Theorem 2.7 *If efficient* $(1 - 1/n, \alpha)$ *dispersers exist, then* $RP = RP(\alpha\text{-}FAIR)$. *If for every* $\alpha > 0$ *efficient* $(1 - 1/n, \alpha)$ *dispersers exist, then* $RP = RP(FAIR)$.

Cohen and Wigderson [CW89] improved the [AKS87] construction of dispersers, using explicit dense Ramanujan graphs ([Chu88,Hal86]). The following is the best known bound.

Theorem 2.8 ([CW89]) *There are efficeint* $(1 - 1/n, 1/2 + \epsilon)$ *dispersers f or every* $\epsilon > 0$.

The reason for the 1/2 barrier in the above theorem, and the ideas for breaking it are discussed in the next section.

3 Dispersers, Expanding Hypergraphs and Eigenvalues

Let us consider a mapping $\mu : \Sigma^m \to (\Sigma^n)^k$. We will fix n, though we continue to use asymptotic notation, and the reader should remember we are really working with an infinite family of objects indexed by n.

[2] In some of the literature on weak random sources multiple access to the source is allowed. However, single access serves to better clarify the probabilistic properties of the source.

[3] This seems the same as the PRB source of [CG85], only that they allow multiple access to the source (see previous footnote).

μ can be thought of (as in [CW89]) as a bipartite multigraph on $\Sigma^n \cup \Sigma^m$ with degree k for every vertex in Σ^m. Equivallently, it will be considered here as a k-uniform (ordered) hypergraph $G = G_\mu = (V, E)$, with $V = \Sigma^n$ and $E \subseteq V^k$, $|E| = 2^m$.

Let $T \subseteq V$. Define $E(T) = E \cap T^k$, i.e. all hyperedges contained in T. It is now natural to call G an (α, β) disperser if for every $T \subseteq V$ with $\log|T| \leq \alpha \log|V|$ we have $\log|E(T)| \leq \beta \log|E|$. In words, if sparse subsets of V contain a sparser set of edges, G is a good disperser. For graphs, (k=2), this is the defining property of expansion, which is why dispersers can be thought of as expanding hypergraphs. Our aim is thus the explicit construction of expanding hypergraphs, where explicit means that the 1-1 mapping $\mu : \Sigma^m \to E$ is efficiently computable (in P).

Let $N = |V| + 2^n$, $|E| = M = 2^m$. For the hypergraph G and any $\alpha = \alpha(n)$, let

$$\beta_G(\alpha) = \max\left\{ \frac{\log|E(T)|}{\log|N|} \; : \; \frac{\log|T|}{\log N} \leq \alpha \right\}.$$

By definition, G is a $(\alpha, \beta_G(\alpha))$ disperser for every α. The question is when do we have $\beta_G(\alpha) < \alpha$? We shall mainly deal with fixed $\alpha \in [0, 1]$, and ignore $o(1)$ terms (i.e. $\alpha \approx \alpha + o(1)$).

A natural idea is to compose hypergraphs. If $G = (V, E)$ is k-uniform and $H = (E, F)$ is l-uniform, then the kl-uniform hypergraph $G \circ H = (V, R)$ is defined by $(v_{11}, \ldots, v_{1k}, \ldots, v_{l1}, \ldots, v_{lk}) \in R$ iff $(e_1, e_2, \ldots, e_l) \in F$ and for every $1 \leq i \leq l$, $e_i = (v_{i1}, \ldots, v_{ik}) \in E$.

Composition is clearly useful: if G is an (α, β) disperser and H is a (β, γ) disperser, then $G \circ H$ is an (α, γ) disperser. Also, if $G^{(t)}$ is the t-fold composition of G with itself, then $G^{(t)}$ is an $(\alpha, \beta_G^{(t)}(\alpha))$ disperser for every α and t.

Remark: One must remember that we restrict $k(n)$ to be $n^{O(1)}$, which limits the number of compositions allowed. However, a fixed number of compositions is always allowed.

To understand the tools for constructing dispersers and their limitations, consider simple graphs, i.e. $k = 2$. Assume $G = (V, E)$ is d-regular, and λ_1, λ_2 the first and second (in absolute value) eigenvalues of its adjacency matrix. Then $\lambda_1 = d$ and λ_2 give a basic upper bound on $|E(T)|$ for every $T \subseteq V$:

Theorem 3.1 (Easy) $|E(T)| \leq \frac{d}{n}|T|^2 + \lambda_2|T|$.

It is convenient to consider all quantities in logarithms to the base N, so let $T = N^\alpha$, $d = N^\delta$, $\lambda_2 = N^\gamma$. The theorem above gives a function $\bar{\beta_G} \geq \beta_G$, with

$$\bar{\beta_G}(\alpha) = \begin{cases} \frac{2\alpha+\delta-1}{1+\delta} & \text{if } \alpha \geq 1+\gamma-\delta \\ \frac{\gamma+\alpha}{1+\delta} & \text{if } \alpha \leq 1+\gamma-\delta \end{cases}$$

Note that $\bar{\beta}(\alpha) \leq \alpha$ only if $\alpha \in [\gamma/\delta, 1+\gamma-\delta]$. The proof of thm [CW89] is essentially an application of this fact, together with the composition of (explicitely constructable) graphs with $\gamma/\delta = 1/2$ for dense enough sequences of value δ.

This method gets stuck at $1/2$, as no explicit construction with $\gamma/\delta < 1/2$ (i.e. $\lambda_2 - d^{1/2-\epsilon}$) is known (for interesting d). In fact, if $d \leq N/2$, this is impossible.

Theorem 3.2 (Easy) $\lambda_2 \geq \sqrt{d(n-d)/n}$.

To see how bad this eigenvalue estimate $\bar{\beta}$ is relative to β, even for random graphs, consider as an example a random graph G of degree $d = \sqrt{N}$, and sets $T, |T| = \sqrt{N}$. With high probability, $\lambda_2 = O(N^{1/4}) = O(d^{1/2})$, so $\bar{\beta_G}(1/2) = 1/2$. However, with high probability, $|E(T)| = O(\sqrt{N})$, so

$$\frac{\log |E(T)|}{\log |E|} = \frac{1}{3},$$

and $\beta_G(1/2) = 1/3$. This means that most graphs G with degree $\sqrt{N}$ are $(1/2, 1/3)$ dispersers, but the eigenvalue bound does not prove it.

This is bad news, since for explicitely constructed graphs, the eigenvalues of the associated matrices are almost the only means to bound $|E(T)|$.

To give a better bound for $|E(T)|$, we observe that for a k-uniform hypergraph $G = (V, E)$ on n vertices, with second eigenvalue with respect to d-regularity λ_2, the following generalization of proposition 3.1 holds.

Theorem 3.3 *For every* $T \subseteq V$, $|E(T)| \leq \frac{d}{n}|T|^k + \lambda_2 |T|^{k/2}$.

For every d-regular k-uniform hypergraph, $\lambda_1(G) = dn^{(k-2)/2}$, and for every k-uniform hypergraph on dn^{k-1} edges certainly $\lambda_1(G) \geq dn^{(k-2)/2}$. It is therefore surprising and promising that according to theorem 1.4 there exist k-uniform hypergraphs with $dn^{(k-2)/2}$ edges whose second eigenvalue is $\leq (c \log n)^{k/2}\sqrt{d}$.

Call a d-regular hypergraph G with $\lambda_2(G) = O(\sqrt{d})$ Ramanujan. It turns out that for our purposes, even explicitely constructible hypergraphs G with $\lambda_2(G) = 2^{O(k)} N^{o(1)} d^{O(1)}$ will suffice to construct better dispersers than the state of the art. We *cannot* build such expanding hypergraphs. In section 5 we describe methods for explicitly construcing hypergraphs with nontrivial bounds on λ_2, and hope these will be improved upon.

Let us consider again the eigenvalue estimate $\bar{\beta_G}$, for highly expanding hypergraphs. Recall that $T = N^\alpha$, $d = N^\delta$, $\lambda_2 = N^\gamma$. Then for $\alpha \leq \frac{2}{k}(1 + \gamma - \delta)$, $\bar{\beta_G}(\alpha) = \dfrac{\gamma + \frac{\alpha k}{2}}{\delta + k - 1}$. So $\bar{\beta_G}$ is decreasing in the interval $\left[\dfrac{2\gamma}{k + 2\delta - 2}, \dfrac{2}{k}(1 + \gamma - \delta)\right]$, with a unique fixed point in the left boundary. Composing G with itself can get us arbitrarily close to the left boundary, i.e. $\bar{\beta_G}(t)\left(\dfrac{2}{k}(1 + \gamma - \delta)\right) \longmapsto_{t \to \infty} \dfrac{2\gamma}{k + 2\delta - 2}$.

This motivates two definitions:

Definition 3.4 (k, δ, γ) *will denote that there is an explicitely constructed family of* k*-uniform hypergraphs* $G_i(V_i, E_i)$ *(of dense enough cardinalities) with* $d_i = |V_i|^\delta$ *and* $\lambda_2(G_i) = O((V_i)^\gamma)$.

For example, the graphs of projective geometries and the graphs of Chung respectively give:

Theorem 3.5 (Hall) *For every integer $r \geq 2$, $(2, 1 - 1/r, 1/2(1 - 1/r))$.*

Theorem 3.6 (Chung, Katz) *For every integer $r \geq 2$, $(2, 1/r, 1/2r)$.*

Definition 3.7 *$[\alpha \mapsto \beta]$ will denote that for every $\epsilon > 0$ there is an explicit dense enough family of $(\alpha + \epsilon, \beta + \epsilon)$-dispersers.*

A useful observation is:

Lemma 3.8 $[\alpha \mapsto \beta], [\beta \mapsto \gamma] \Rightarrow [\alpha \mapsto \gamma]$

The discussion above implies in this notation:

Theorem 3.9

$$(k, \delta, \gamma) \Rightarrow \left[\frac{2}{k}(1 + \gamma - \delta) \mapsto \frac{2\gamma}{k + 2\delta - 2}\right]$$

This and either of the theorems above ([Hall, Chung]) imply:

Theorem 3.10 $(1 - \tau \mapsto 1/2)$ *for every fixed $\tau > 0$.*

Using this theorem, Theorem [CW89] is implied from [AKS87].

For Ramanujan hypergraphs, $\gamma = \delta/2$. To realize their potential we prove:

Theorem 3.11 *[For every fixed k, $(k, 1/2, 1/4)$]* $\Rightarrow [1/2 \mapsto 0] \Rightarrow RP = RP(FAIR)$.

We picked $\delta = 1/2$ and $\gamma = 1/4$ as an example. The interested reader can compute the envelope of values for which the consequence $RP = RP(FAIR)$ holds.

Proof: The premise $(k, 1/2, 1/4)$ implies, by Theorem 3.9 that

$$[\frac{3}{2k} \mapsto \frac{1}{2(k-1)}]$$

for every k. Using Lemma 3.8 and composition, we get $[3/4 \mapsto 0]$. With [CW89] and Theorem 2.7 this implies $RP = RP(FAIR)$.

4 Basic Facts about Hypergraphs

In this section we will prove some basic facts about the eigenvalues of hypergraphs. We will call a multilinear form, τ, d-regular if it satisfies the conditions of regularity stated before for hypergraphs.

Theorem 4.1 *Let τ be a non-negative, d-regular k-linear form on n vertices. Then the first eigenvalue of H is $dn^{(k-2)/2}$, with $\vec{1} \otimes \cdots \otimes \vec{1}$ being an eigenvector.*

Proof We shall proceed by induction on k. For $k = 1$ this is clear, for then the associated multilinear form, τ, is just a one-dimensional vector proportional to $\vec{1}$. For general k, let the associated trilinear form, τ, take its maximum over the product of unit balls at $(u_1, \ldots, u_k)$. Since τ has non-negative coefficients, we can assume that each u_i has non-negative coefficients (or replace the u_i by such vectors while preserving their norm and without decreasing the absolute value of τ at $(u_1 \ldots, u_k)$). But viewing u_1 as fixed, the $k-1$ form $\tau(u_1, \cdot, \ldots, \cdot)$ is easily seen to be a non-negative, θ-regular form where θ is the sum of the components of u_1; therefore, by induction, then norm of this $k-1$-linear form is $\theta n^{(k-3)/2}$, with $\vec{1} \otimes \cdots \otimes \vec{1}$ being a coresponding eigenvector. But the ratio of θ to $\|u_1\|$ is maximized when u_1 is proportional to $\vec{1}$, and so the norm of τ is $dn^{(k-2)/2}$, with $\vec{1} \otimes \cdots \otimes \vec{1}$ being a corresponding eigenvector.

□

Theorem 4.2 *Let τ be a symmetric k-linear form. Then there is a vector $u \in L^2(V)$ such that $u \otimes \cdots \otimes u$ is a first eigenvector for τ.*

Lemma 4.3 *Let ν be a symmetric 2-linear form. If u and v are unit vectors such that $\|\nu\| = |\nu(u, v)|$, then also $\|\nu\| = |\nu(u, u)|$.*

Proof By linear algebra, ν is diagonalizable with an orthonormal basis. Let $\tilde{\nu}$ be the associated endomorphisn of $L^2(V)$. If $\tilde{\nu}$'s largest eigenvalue in absolute value is λ, then $\|\nu\| = |\lambda| = \|\tilde{\nu}\|$. Since

$$|\lambda| = |\nu(u, v)| = |(\tilde{\nu}(u), v)| \leq \|\tilde{\nu}(u)\|\|v\| = \|\tilde{\nu}(u)\|,$$

u must be an eigenvector with corresponding eigenvalue $= \pm\|\nu\|$, and so $\|\nu\| = |\nu(u, u)|$.

□

The theorem easily follows. If $u_1 \otimes \cdots \otimes u_k$ is a first eigenvector for τ, then applying the lemma to the bilinear form $\tau(\cdot, \cdot, u_3, \ldots, u_k)$ we see that we can replace u_2 by u_1. Similarly we can replace all the u_i by u_1, successively, proving the theorem.

□

We now give a lower bound for the second eigenvalue. Since the projection of a vector with a ones and $n - a$ zeros onto $\vec{1}^{\perp}$ has norm $\sqrt{a(n-a)/n}$, the following proposition follows immediately from the definition of second eigenvalue:

Proposition 4.4 *Let H be a k-uniform, d-regular hypergraph. For any subsets $U_1, \ldots U_k$ of V, the number of edges in $U_1 \times \cdots \times U_k$ is*

$$\frac{d}{n}|U_1| \cdots |U_k| + \theta\lambda_2 \sqrt{\frac{|U_1|(n - |U_1|)}{n}} \cdots \sqrt{\frac{|U_k|(n - |U_k|)}{n}}$$

for some θ with $|\theta| \leq 1$, and where λ_2 is the second eigenvalue of the hypergraph.

Now take any $U_1, \ldots, U_{k-1}$ consisting of one vertex. Then we can find a U_k of size $n - d$ for which E contains no edges in $U_1 \times \cdots \times U_k$. It follows that:

Proposition 4.5 *Let H be as in proposition 4.4. Then*

$$\lambda_2 \geq \sqrt{\frac{d(n-d)}{n}}.$$

In the literature there are many methods for obtaining upper and lower bounds on the second eigenvalue of graphs. We remark that when we consider the bilinear form $\tau(v_1, \ldots, v_{k-2}, \cdot, \cdot)$ with each v_i being one of the standard vectors, e_j, we get a d-regular graph, and can therefore apply lower bounds known for the second eigenvalue of graphs. The preceding proposition is an example. Applying another known lower bound for graphs (see [Alo86]) yields:

Proposition 4.6 *For some constant C we have that if $d \leq n^{1/C}$ then $\lambda_2 \geq 2\sqrt{d-1} - C/\log_d n$.*

As for obtaining upper bounds, the strongest methods don't seem to directly generalize. The so-called "trace method," which uses multiplication of graphs and taking the trace of a graph (its adjacency matrix), does not have an obvious generalization (that we see). However, for graphs which are Cayley graphs, such as those of [LPS86] and [Mar87], there is a natural generalization, and it turn out that to analyze their second eigenvalue is suffices to analyze the second eigenvalue of the corresponding graph (and to understand the representations of the underlying group), so the trace method can be applied there. For random graphs, we don't know how to apply results that use the trace method, such as those of [Wig55], [BS87], [Fri88]. However the method of Kahn and Szemerédi, in [FKS89], does generalize quite readily, and we can prove theorem 1.4, which will be done in section 6.

For applications we would like to come up with constructions to match this bound. Often we would like the graph to be constructible in poly log of the number of nodes, for example in using weak random sources. We do not have an explicit construction to match the bound in theorem 1.4, not even one which is constructable in polynomial time in the number of nodes. We will describe two types of constructions; one is very easy, and yields a weak eigenvalue bound. The second type is an analogue of Cayley graphs, which we call Cayley hypergraphs. This gives somewhat improved bounds, but we also give a lower bound to show that Cayley hypergraphs cannot match the bound of theorem 1.4. We find this interesting because most explicit constructions of graphs with small second eigenvalue that we know are Cayley graphs. Perhaps there is a better generalization of this concept that can yield hypergraphs with small second eigenvalue; one such construction might be buildings with finite groups or finite quotients of buildings of rank $t - 1$ Lie groups (for t-uniform hypregraphs) (see [BT71]). We finish this section by giving an easy construction for a graph with slightly small second eigenvalue.

Let $G = (V, E)$ be any graph which is d-regular, and let $\widetilde{G} = (V, V^{t-2} \times E)$.

Proposition 4.7 *If λ_2 is the second eigenvalue of G, then the second eigenvalue of $\widehat{G}$ is no greater than $\sqrt{n}|\lambda_2|$.*

Proof It is easy to see that the associated τ is d-regular. If $v, w \in \{\vec{1}\}^{\perp}$, then clearly $|\tau(e_i, v, w)| \leq \lambda_2 \|v\| \|w\|$ for any standard basis vector e_i. Thus

$$|\tau(u, v, w)| \leq \left(\sum_{i=1}^{n} |u_i| \right) \lambda_2 \|v\| \|w\| \leq \sqrt{n} \lambda_2 \|u\| \|v\| \|w\|.$$

5 Cayley Hypergraphs

Let G be a group, and H a subset of G. The *Cayley graph* on G generated by H is defined to be the graph with vertex set G and edge set

$$\{(x, y) \mid xy^{-1} \in H\}.$$

(We do not require that H generate G, nor that $H = H^{-1}$.) This gives a d-regular directed graph, and if $H = H^{-1}$ we can view the graph as undirected. We define the *Cayley sum graph* similarly, though taking the edge set to be

$$\{(x, y) \mid xy \in H\}.$$

Perhaps the easiest way to get a 3-regular hypergraph from this data is to keep G as the set of vertices and to take

$$\{(x, y, z) \mid xyz \in H\}$$

as the edge set; we will call this hypergraph the *(3-regular) Cayley hypergraph* on G and H.

The eigenvalues of Cayley graphs and, as we shall see, hypergraphs, can often be estimated when one understands the decomposition of $L^2(G)$ under the right regular representation. We recall the following facts about representations of finite groups (see [Rob83] for details). We will, for the moment, let $L^2(G)$ denote the space of complex-valued functions on G with the usual inner product:

$$(u, v) = \sum_{g \in G} u_g \overline{v_g}.$$

$L^2(G)$ can be decomposed into subspaces

$$L_2(G) = \bigoplus_{i=1}^{r} E_i$$

with the following conditions:

1. Each E_i is invariant under the natural action of G on $L_2(G)$, given by $g(u(x)) \equiv u(gx)$.

2. $\dim(E_i) = d_i^2$ for some d_i corresponding to the dimension of an irreducible unitary representation of G, $\rho_i : G \to \mathrm{Gl}(\mathbf{C}, d_i)$, in the sense that a complete orthogonal basis for E_i is given by the d_i^2 entries of ρ_i with respect to any basis of $\mathbf{C}^n$. Also, the norm of each coefficient of ρ_i, as an element of $L^2(G)$, has norm $\sqrt{n/d_i}$.

3. r is equal to the number of conjugacy classes in G.

It follows that the matrix A of any Cayley graph on G vanishes outside the $E_i \times E_i$ blocks, in the sense that $\tau(u, v) = 0$ if u and v are contained in different E_i's. More generally, the t-regular hypergraph generated by G and H vanishes outside the $E_i \times \cdots \times E_i$ blocks. Let us assume, for simplicity, that the Cayley graph is generated by an H which satisfies $H{=}H^{-1}$. It then follows that the eigenvalues of A are real, and there is an orthogonal set of real eigenvectors.

Theorem 5.1 *Let the eigenvalues of A restricted to E_i be $\lambda_1, \ldots \lambda_{d_i^2}$. Then the norm of the t-linear form associated to the t-regular Cayley hypergraph on G and H is*

$$\left(\frac{n}{d_i}\right)^{(t-2)/2} \max_j |\lambda_j|.$$

Corollary 5.2 *The second eigenvalue of a t-uniform Cayley hypergraph of degree d is at least $\sqrt{d(n-d)/n}\, n^{(t-2)/2}$.*

Proof For simplicity we will first prove this for 3-regular hypergraphs, and then indicate how to generalize the proof. Let A and τ, by abuse of notation, denote the adjacency matrix and bilinear form, respectively, associated to the Cayley graph restricted to the subspace E, corresponding to the d dimensional representation ρ. For any basis of $\mathbf{C}^d$, we can consider the coefficients of ρ, $\{\rho_{i,j}\}$, with respect to a given basis for E, $w_1, \ldots, w_d$. For any i, j, k, l, we have

$$\sum_{xy=h} \rho_{i,j}(x)\rho_{k,l}(y) = \sum_{g \in G} \rho_{i,j}(g)\rho_{k,l}(g^{-1}h)$$

which, using the fact that ρ is a unitary representation, is

$$\sum g \in G \sum_{m=1}^{d} \rho_{i,j}(g)\overline{\rho_{l,m}(g)}\rho_{m,k}(h) = \delta_{i,k}\left(\frac{n}{d}\right)\rho_{j,l}(h),$$

where $\delta_{i,k}$ is the Kronecker delta function. It follows that τ is given by

$$\tau\left(\sum_{i,j}\alpha_{i,j}\rho_{i,j},\sum_{k,l}\beta_{k,l}\rho_{k,l}\right)=\sum_{i,j,k,l}\alpha_{i,j}\beta_{k,l}\delta_{i,k}M_{j,l}=\sum_{i,j,l}\alpha_{i,j}\beta_{i,l}M_{j,l}$$

where $M_{j,l}$ is given by the matrix equation

$$M=\left(\frac{n}{d}\right)\sum_{h\in H}\rho(h).$$

Since $H=H^{-1}$, it follows that M is a real symmetric matrix, and therefore is diagonalizable by a set of real eigenvectors in $\mathbf{R}^d$ with real eigenvalues. Take these eigenvectors as the basis of $\mathbf{C}^d$, $w_1,\ldots,w_d$. Then M becomes a diagonal matrix. Since the norm of each $\rho_{i,j}$ is $\sqrt{n/d}$, it follows that for each i,j, the function $\rho_{i,j}$ is an eigenvector of the adjacency matrix, with eigenvalue 0 if $i\neq j$, and eigenvalue $\lambda_{i,i}=\frac{d}{n}M_{i,i}$ if $i=j$.

Now let ν be the trilinear form associated to the 3-regular hypergraph generated by G and H. A similar calculation shows that ν is given by

$$\nu\left(\sum_{i,j}\alpha_{i,j}\rho_{i,j},\sum_{k,l}\beta_{k,l}\rho_{k,l},\sum_{q,r}\gamma_{q,r}\rho_{q,r}\right)=\sum\alpha_{i,j}\beta_{k,l}\gamma_{q,r}\delta_{i,k}\delta_{j,q}\delta_{l,r}\frac{n}{d}M_{l,l}$$

$$=\sum\alpha_{i,j}\beta_{i,l}\gamma_{j,l}\frac{n}{d}M_{l,l}$$

It follows that $\nu(\rho_{i,i},\rho_{i,i},\rho_{i,i})$ is $\frac{n}{d}M_{i,i}$, and so

$$\|\nu\|\geq\sqrt{\frac{n}{d}}\lambda_{i,i}$$

for each i. We claim this hitting the ν with $\rho_{i,i}$ is the best that we can do, i.e.

Proposition 5.3 *We have*

$$\|\nu\|=\max_i\sqrt{\frac{n}{d}}|\lambda_{i,i}|.$$

Proof Let $\alpha_{i,j},\beta_{k,l},\gamma_{q,r}$ be given with

$$\sum_{i,j}\alpha_{i,j}^2=\sum_{i,j}\beta_{i,j}^2=\sum_{i,j}\gamma_{i,j}^2=1.$$

It suffices to obtain the estimate

$$\sum_{i,j,k}\alpha_{i,j}\beta_{i,k}\gamma_{j,k}\leq 1.$$

Letting

$$a_{j,k} = \sum_i \alpha_{i,j}\beta_{i,k},$$

we see that applying Cauchy-Schwartz inequality yields

$$\sum_{j,k} a_{j,k}^2 \leq \sum_{j,k} \left(\sum_i \alpha_{i,j}^2\right)\left(\sum_i \beta_{i,k}^2\right) = 1$$

and so

$$\sum_{j,k} a_{j,k}\gamma_{j,k} \leq \left(\sum_{j,k} a_{j,k}^2\right)^{1/2}\left(\sum_{j,k} \gamma_{j,k}^2\right)^{1/2} \leq 1.$$

□

The generalization of theorem 5.1 to t-uniform hypergraphs follows from similar considerations, and the generalized estimate:

Proposition 5.4 *Let $\alpha_{i,j}^1, \alpha_{i,j}^2, \cdots, \alpha_{i,j}^s$ be s $n\times n$ arrays of numbers with $\sum_{i,j}(\alpha_{i,j}^r)^2 = 1$ for each r. Then*

$$\sum \alpha_{i_1,i_2}^1 \alpha_{i_1,i_3}^2 \alpha_{i_2,i_4}^3 \alpha_{i_3,i_5}^4 \cdots \alpha_{i_{k-2},i_k}^{k-1} \alpha_{i_{k-1},i_k}^k \leq 1$$

Proof Let

$$a_{j,k}^r = \sum \alpha_{i_1,i_2}^1 \alpha_{i_1,i_3}^2 \alpha_{i_2,i_4}^3 \alpha_{i_3,i_5}^4 \cdots \alpha_{i_{r-2},j}^{r-1} \alpha_{i_{r-1},k}^r$$

for $r < k$. By induction on r, using Cauchy-Schwartz as before, one can prove that

$$\sum_{j,k}(a_{j,k}^r)^2 \leq 1$$

for all r, and therefore prove the proposition, by estimating

$$\sum_{j,k} a_{j,k}^{k-1}\alpha_{j,k}^k \leq \left(\sum_{j,k}(a_{j,k}^{k-1})^2\right)^{1/2}\left(\sum_{j,k}(\alpha_{j,k}^k)^2\right)^{1/2} \leq 1.$$

6 Random Hypergraphs

First we will prove theorem 1.4, and then theorem 1.5 for arbitrary t. In this section, C, C', C_1, etc. will each denote various absolute positive constants unless otherwise indicated. Throughtout this section the phrase "with high probability" is short for there exist constants C, C' such that the event occurs with probability at least $1 - Cn^{-C'\min(\sqrt{d}, n^{1/(t+3)})/t}$. When we prove theorem 1.4, we take $t = 3$ in this formula, but for for theorem 1.5 this is to hold for all t with C and C' are independent of t.

Proof of theorem 1.4 To avoid issues of the model we work with the probability space $\mathcal{G}_{n,d}$ of random 3-uniform hypergraphs on n vertices which has dn^2 directed edges chosen at random independently and uniformly from all possible n^3 "edge slots." A graph in $\mathcal{G}_{n,d}$ won't in general be d-regular, but we will estimate its second eigenvalue with respect to d-regularity. $\mathcal{G}_{n,d}$ graphs are approximately d-regular in the following two senses:

Lemma 6.1 *Let $d \geq \rho \log n$ with $\rho \geq C_0$. With probability $1 - n^{C_1 \rho}$, a $\tau \in \mathcal{G}_{n,d}$ will satisfy*

$$\sum_{i \in A, j \in B, 1 \leq k \leq n} \tau_{ijk} \leq C_2 d|A||B|$$

for every $A, B \subset V$, and the same with τ_{ijk} replaced by τ_{ikj} or τ_{kij}.

Proof The probability that τ does not satisfy the equation is bounded by

$$\sum_{a,b=1}^{n} \binom{n}{a}\binom{n}{b} B(dn^2, Cdab, ab/n^2),$$

where $B(r,s,t)$ denotes the probability of at least s successes in r Bernoulli trials of probability t. Since $B(r,s,t) \leq r\binom{r}{s}t^s \leq (ert/s)^s r$, we have

$$B(dn^2, Cdab, ab/n^2) \leq (e/C)^{Cdab} dn^2,$$

and the right-hand-side of the above is clearly bounded by $n^{-C_1\rho}$ for some positive C_1.

Lemma 6.2 *With probability $1 - Cn^{-C'n^{1/5}}$, a τ in $\mathcal{G}_{n,d}$ will have*

$$\left|\tau(\vec{1}/\sqrt{n}, u, v)\right| \leq C(\log n)\sqrt{d}\|u\|\|v\|$$

for all $u, v \in \{\vec{1}\}^\perp$, and similarly with the $\vec{1}/\sqrt{n}, u, v$ permuted in any order. The same holds with u replaced by $\vec{1}/\sqrt{n}$.

Proof The matrix, A, given by

$$A_{ij} = \sum_{k=1}^{n} \tau_{kij}$$

is a random graph on n vertices with average degree nd. It would be a standard fact that with very large probability

$$\left\|A - d\vec{1} \otimes \vec{1}\right\|_2 \leq C\sqrt{nd}$$

for a slightly different model of random A, e.g. if A's entries were chosen independently from some distribution. This is not the case here, but since we are willing to settle for a $\leq C\sqrt{d}\log n$ estimate, it is easiest to note that this follows from the generalization of theorem 6.3, theorem 6.7, for 2-uniform hypergraphs with d replaced by dn. (The proof of theorem 6.3 and its generalization does not use lemma 6.2.) This proves the first part of lemma 6.2. The second part of the lemma is just a direct calculation— it suffices to show that dn^2 Bernoulli trials with success probability $1/n$ have $dn \pm O(\sqrt{dn})$ successes with large probability. This indeed occurs with probability $\geq 1 - C^n$ with some $C < 1$.

To prove theorem 1.4, we estimate $\lambda_{2,d}(\tau)$ by writing

$$\tau(u_1, \ldots, u_3) = \tau(\widetilde{u}_1 + \alpha_1\vec{1}, \ldots, \widetilde{u}_3 + \alpha_3\vec{1})$$

with $\widetilde{u}_i \in \{\vec{1}\}^\perp$ and $\alpha_i \in \mathbf{R}$, and use multilinearity. By lemma 6.2, we see that theorem 1.4 follows from:

Theorem 6.3 *With probability* $1 - n^{-C'\min(\sqrt{d}, n^{1/6})}$, *a* τ *in* $\mathcal{G}_{n,d}$ *satisfies*

$$|\tau(u, v, w)| \leq C\sqrt{d}(\log n)^{3/2}\|u\|\|v\|\|w\| \tag{6.1}$$

for all $u, v, w \in \{\vec{1}\}^\perp$.

Proof For the proof we use the Kahn-Szemerédi approach of estimating the second eigenvalue of a random graph, see [FKS89]. Clearly it suffices to prove this for all u, v, w in some large finite set, $\mathcal{M} \subset \{\vec{1}\}^\perp$; in fact one can take

$$\mathcal{M} = \left(\frac{1}{4\sqrt{n}}\mathbf{Z}\right)^n \cap \{\vec{1}\}^\perp \cap B_1(0),$$

where $B_1(0)$ denotes the unit ball, for it is easy to see that any $u \in \{\vec{1}\}^\perp \cap B_1(0)$ can be written as

$$u = \sum_{i=-1}^{\infty} \left(\frac{1}{4}\right)^i u_i$$

with $u_i \in \mathcal{M}$, and thus

$$|\tau(u, v, w)| \leq \left|\tau\left(\sum\left(\frac{1}{4}\right)^i u_i, \sum\left(\frac{1}{4}\right)^i v_i, \sum\left(\frac{1}{4}\right)^i w_i\right)\right| \leq C\sqrt{d}(\log n)^{3/2}\left(\sum_{i=-1}\left(\frac{1}{4}\right)^i\right)^3$$

if $u_i, v_i, w_i \in \mathcal{M}$, assuming that equation 6.1 holds for all $u, v, w \in \mathcal{M}$. It is easy to check, see [FKS89], that $|\mathcal{M}| \leq C^n$.

Fix $u, v, w \in \mathcal{M}$, and write $u = \sum u_i e_i$, and similarly for v and w, where e_i are the standard basis. We break the sum

$$\tau(u, v, w) = \sum u_i v_j w_k \tau_{ijk}$$

into two sums, by setting

$$S = \{(i,j,k) \mid |u_i v_j w_k| \leq \sqrt{d}/n\}.$$

S represents the "small tuples" $u_i v_j w_k$. We now argue first that:

Lemma 6.4 *For any C there is a C' such that for a given u, v, w,*

$$\left| \sum_{(i,j,k) \in S} u_i v_j w_k \tau_{ijk} \right| \leq C' \sqrt{d}$$

for $\tau \in \mathcal{G}_{n,d}$ with probability $\geq 1 - C^{-n}$.

Since there are $|\mathcal{M}|^3 = C^{3n}$ such tuples, lemma 6.4 holds for all $(u, v, w) \in \mathcal{M}^3$ with high probability.

On the other hand we claim that the sum outside S will, for most hypergraphs, be a priori bounded by $C\sqrt{d}$, a condition that is only violated when certain "unusually high edge densities" occur in τ.

Lemma 6.5 *With high probability, a $\tau \in \mathcal{G}_{n,d}$ satisfies*

$$\left| \sum_{\overline{S}} u_i v_j w_k \tau_{ijk} \right| \leq C(\log n)^{3/2} \sqrt{d}$$

for all u, v, w.

It remains to prove the above two lemmas. First notice that

$$\left| \sum_{S} u_i v_j w_k \right| = \left| \sum_{\overline{S}} u_i v_j w_k \right| \leq \frac{n}{\sqrt{d}},$$

since

$$\left| \sum_{\overline{S}} u_i v_j w_k \right| \leq \frac{n}{\sqrt{d}} \sum_{\overline{S}} u_i^2 v_j^2 w_k^2 \leq \frac{n}{\sqrt{d}}.$$

Lemma 6.4 is now a standard calculation: we have

$$\mathrm{E}\left\{ e^{\sum_S \lambda u_i v_j w_k \tau_{ijk}} \right\} = \left(\sum_S \frac{1}{n^3} e^{\lambda u_i v_j w_k} \right)^{dn^2}$$

$$= \left(e^{\frac{1}{n^3} \sum_S \lambda u_i v_j w_k} e^{\frac{1}{n^3} \sum_S O(\lambda u_i v_j w_k)^2} \right)^{dn^2} = e^{\frac{d}{n} \lambda \sum_S u_i v_j w_k} e^{\frac{d}{n} \lambda^2 O(1)}$$

provided that $\lambda u_i v_j w_k = O(1)$ for all $(i,j,k) \in S$, with an absolute constant in the $O(\,)$ notation. Taking $\lambda = n/\sqrt{d}$ yields $e^{O(n)}$ in the above, and thus

$$\Pr\left\{\sum_S u_i v_j w_k \tau_{ijk} \geq C\sqrt{d}\right\} \leq e^{-Cn} e^{O(n)}.$$

Taking $\lambda = -n/\sqrt{d}$ yields a similar exponential lower bound:

$$\Pr\left\{\sum_S u_i v_j w_k \tau_{ijk} \leq -C\sqrt{d}\right\} \leq e^{-Cn} e^{O(n)}$$

for a positive constant C.

To prove lemma 6.5, we consider for subsets $A, B, C \subset V$,

$$e(A,B,C) \equiv \sum_{i\in A, j\in B, k\in C} \tau_{ijk}$$

and

$$\tilde{e}(A,B,C) \equiv \sum_{i\in A, j\in B, k\in C, (i,j,k)\in \overline{S}} \tau_{ijk}.$$

The standard counting argument and estimates yield:

Lemma 6.6 *With probability $1-\epsilon$ we have that a $\tau \in \mathcal{G}_{n,d}$ satisfies for all subsets $A, B, C \subset V$,*

$$e(A,B,C) \leq k(|A|,|B|,|C|)$$

where $k = k(a,b,c)$ is any function of a, b, c satisfying

$$\binom{n}{a}\binom{n}{b}\binom{n}{c}\left(\frac{e\frac{d}{n}abc}{k}\right)^k \leq \frac{\epsilon}{n^3}$$

for all a, b, c.

Proof A standard calculation, using the estimate $\binom{a}{b} \leq (ea/b)^b$.

The above lemma gives a condition that the hypergraph have no "irregular (overly large) edge densities." Now we set for $i > 0$

$$A_i = \{j \mid 2^{-i} < u_j \leq 2^{-i+1}\} \subset \{1, \ldots, n\}$$

and

$$A_{-i} = \{j \mid 2^{-i} < -u_j \leq 2^{-i+1}\} \subset \{1, \ldots, n\}$$

and similarly define B_i and C_i; note that for $m = \lfloor \log(4\sqrt{n}) \rfloor$, we have that the above sets are empty when $i > m$. We now estimate

$$\sum_{u_i, v_j, w_k > 0, \overline{S}} u_i v_j w_k \tau_{ijk} \leq \sum_{\alpha,\beta,\gamma=1}^{m} \sum_{\overline{S} \cap A_\alpha \times B_\beta \times C_\gamma} \tilde{e}(A_\alpha, B_\beta, C_\gamma) \overline{u}_\alpha \overline{v}_\beta \overline{w}_\gamma \qquad (6.2)$$

where $\overline{u}_\alpha = 2^{-\alpha+1}$ and similarly for $\overline{v}$ and $\overline{w}$. (Thus $\overline{u}_\alpha$ approximates every u_i with $i \in A_\alpha$; we use this notation to be suggestive.) The game becomes to choose an appropriate $k = k(a,b,c)$ in the above lemma to get good estimates in equation 6.2. Then we do the same for the analogue of equation 6.2 with the sum over all other sign combinations for $\alpha, \beta, gamma$, and we are done.

To estimate the right-hand-side of equation 6.2, note that we can assume that $|C_\gamma|$ is at least as large as $|A_\alpha|$ and $|B_\beta|$ (and then repeating the argument when $|A_\alpha|$ and then $|B_\beta|$ are the largest). We may also assume that

$$\overline{u}_\alpha \overline{v}_\beta \leq \sqrt{d}\, \overline{w}_\gamma, \tag{6.3}$$

for each term, for if not then take out such terms and note that since $\sum_\gamma \widetilde{e} \leq \sum_\gamma e \leq Cd|A_\alpha||B_\beta|$,

$$\sum \widetilde{e}(A_\alpha, B_\beta, C_\gamma)\overline{u}_\alpha \overline{v}_\beta \overline{w}_\gamma \leq \sum C\sqrt{d}|A_\alpha||B_\beta|\overline{u}_\alpha^2 \overline{v}_\beta^2 \leq C\sqrt{d}.$$

It is easy to see that for $d \leq n^{1/3}$ the choice of

$$k(a,b,c) = \begin{cases} C\sqrt{d} & \text{if } c \leq C'\sqrt{d} \\ C(\log n)^{3/2}\frac{d}{n}\, abc & \text{if } ab \geq \frac{n}{d}(\log n)^{-1/2} \\ (\log n)^{1/2} c & \text{if } ab \leq \frac{n}{d}(\log n)^{-1/2} \text{ and } c \geq C'\sqrt{d} \end{cases}$$

when $c = \max(a,b,c)$ satisfies lemma 6.6 with $\epsilon = n^{-C''\sqrt{d}}$, while either of the three bounds for k (and therefore $\widetilde{e}$) above give an estimate for the sum in equation 6.2: first of all, where $k = C\sqrt{d}$ we can estimate the sum in equation 6.2

$$\sum_{\alpha,\beta,\gamma=1}^{m} \widetilde{e}(A_\alpha, B_\beta, C_\gamma)\overline{u}_\alpha \overline{v}_\beta \overline{w}_\gamma \leq \sum_{\alpha,\beta,\gamma=1}^{m} C\sqrt{d}\overline{u}_\alpha \overline{v}_\beta \overline{w}_\gamma \leq C'\sqrt{d}$$

since each of $\overline{u}$, $\overline{v}$, and $\overline{w}$ are geometrically decreasing sequences (i.e. $1, 1/2, 1/4, \ldots$). Second, where k is $C(\log n)^{3/2}(d/n)abc$ we can estimate the sum by

$$\sum C(\log n)^{3/2}\frac{d}{n}|A_\alpha||B_\beta||C_\gamma|\overline{u}_\alpha \overline{v}_\beta \overline{w}_\gamma$$

$$\leq \sum C(\log n)^{3/2}\sqrt{d}|A_\alpha||B_\beta||C_\gamma|\overline{u}_\alpha^2 \overline{v}_\beta^2 \overline{w}_\gamma^2 \leq C'(\log n)^{3/2}\sqrt{d},$$

the first inequality using that fact that each of the tuples is in $\overline{S}$. Finally, where k is $C(\log n)^{1/2}c$ we can estimate the sum by

$$\leq \sum_{\alpha,\beta,\gamma} C(\log n)^{1/2}|C_\gamma|\overline{u}_\alpha \overline{v}_\beta \overline{w}_\gamma \leq \sum_\gamma C(\log n)^{1/2}|C_\gamma| \sum_{\alpha,\beta} \overline{u}_\alpha \overline{v}_\beta,$$

where the leftmost sum is over all α and β satisfying equation 6.3. Since the $\overline{u}$'s and $\overline{v}$'s are a geometric sequence, and since the α and β range over $m = O(\log n)$ values, we have that for any fixed γ,

$$\sum_{\overline{u}_\alpha \overline{v}_\beta \leq \sqrt{d}\overline{w}_\gamma} \overline{u}_\alpha \overline{v}_\beta \leq$$

$$\sqrt{d}\overline{w}_\gamma \left(m + \frac{1}{2}m + \frac{1}{4}m + \cdots \right) \leq \sqrt{d}\overline{w}_\gamma C' \log n.$$

Hence the sum in equation 6.2 can be bounded by

$$\sum_\gamma C(\log n)^{3/2}\sqrt{d}|C_\gamma|\overline{w}_\gamma^2 \leq C(\log n)^{3/2}\sqrt{d}.$$

To finish the proof of theorem 1.4, we must alter slightly our choice of k for $d \geq n^{1/3}$. For this case we can chose k as before except replacing each $\sqrt{d}$ by a $n^{1/6}$, giving us an ϵ of $n^{-C''n^{1/6}}$. This completes the proof of theorem 1.4.

□

Proof of theorem 1.5 We indicate what needs to be changed in the previous proof. To prove theorem 1.5 for t-uniform hypergraphs, we estimate $\lambda_{2,d}(\tau)$ by writing

$$\tau(u_1, \ldots, u_t) = \tau(\widetilde{u}_1 + \alpha_1\vec{1}, \ldots, \widetilde{u}_t + \alpha_t\vec{1})$$

with $\widetilde{u}_i \in \{\vec{1}\}^\perp$ and $\alpha_i \in \mathbf{R}$, and use multilinearity. As before, it suffices to prove a generalization of theorem 6.3 for a random t-linear form, where we prove the estimate

$$|\tau(u_1, \ldots, u_t)| \leq \sqrt{d}(C\log n)^{t/2}\|u_1\| \cdots \|u_t\|$$

with high probability. To do so, first, in lemma 6.1 we replace "$d \geq \rho \log n$" by "$d \geq \rho t \log n$," and "probability $1 - n^{\rho C_1}$" with "high probability;" the proof is the same. In lemma 6.5 we replace $3/2$ by $t/2$. The proof of the lemma is almost identical, except that we choose k slightly differently, namely

$$k(a_1, \ldots, a_t) = \begin{cases} C\sqrt{d} & \text{if } a_t \leq C'\sqrt{d} \\ C(\log n)^{t/2}\frac{d}{n}\, a_1 \cdots a_t & \text{if } a_1 \cdots a_{t-1} \geq \frac{n}{d}(\log n)^{(2-t)/2} \\ (\log n)^{(4-t)/2}a_t & \text{if } a_1 \cdots a_{t-1} \leq \frac{n}{d}(\log n)^{(2-t)/2} \text{ and } a_t \geq C'\sqrt{d} \end{cases}$$

for $d \leq n^{2/(t+3)}$ (assuming $a_t = \max(a_1, \ldots, a_t)$), and for larger d we replace the $\sqrt{d}$'s by $n^{1/(t+3)}$'s. The only difference in the estimates from this point on is that the old estimate

$$\sum_{\overline{u}_\alpha \overline{v}_\beta \leq \sqrt{d}\overline{w}_\gamma} \overline{u}_\alpha \overline{v}_\beta \leq \sqrt{d}\,\overline{w}_\gamma C \log n$$

must be replaced by the analogous sum being estimated by

$$\sqrt{d}\,\overline{w}_\gamma C(\log n)^{t-2}.$$

In lemma 6.4 we replace $C'\sqrt{d}$ by $C't\sqrt{d}$, and prove the inequality with probability $1 - C^{-nt}$. We need this because we are counting over $\mathcal{M}^t$, whose size is C^{nt}. Thus we can establish the generalization of theorem 6.3:

Theorem 6.7 *With high probability, a τ in $\mathcal{G}_{n,d,t}$ satisfies*

$$|\tau(u_1, \ldots, u_t)| \leq C\sqrt{d}(\log n)^{t/2}\|u_1\| \cdots \|u_t\|$$

for all $u_1, \ldots, u_t \in \{\vec{1}\}^{\perp}$.

Here $\mathcal{G}_{n,d,t}$ is the space of random t-uniform hypergraphs with n vertices and with dn^{t-1} edges chosen at random from all n^t possible "edge slots."

The analog of lemma 1.2 is that with high probability, a τ in $\mathcal{G}_{n,d,t}$ satisfies

$$|\tau(u_1, \ldots, u_t)| \leq C\sqrt{d}(\log n)^{(t-1)/2}$$

for all $u_1, \ldots, u_t$ with at least one of the u_i's $= \vec{1}/\sqrt{n}$, and with each other u_i's either $= \vec{1}/\sqrt{n}$ or in $\{\vec{1}\}^{\perp} \cap B_1(0)$. As before, this follows from theorem 6.7. Since each of the $\leq 2^t$ ways of fixing a subset of u's to $\vec{1}/\sqrt{n}$ satisfies the desired inequality with high probability, all $\leq 2^t$ inequalities are satisfied simultaneously with high probability (note the definition of high probability has a $1/t$ factor in the exponent). To finish, we use multilinearity, as described in the beginning of the proof of theorem 1.5.

□

We remark that our choice of k is more relaxed that the choice of k in [FKS89], and consequently the analysis is easier at the cost of possibly forsaking a $\log n$ factor. If we replaced, say, the first k in the proof of theorem 1.4 by

$$k(a,b,c) = \begin{cases} C\sqrt{d} & \text{if } c \leq C'\sqrt{d} \\ C\frac{d}{n}abc & \text{if } ab \geq \frac{n}{d}(\log n) \\ r(a,b,c) & \text{if } ab \leq \frac{n}{d}(\log n) \text{ and } c \geq C'\sqrt{d} \end{cases}$$

where $r = r(a,b,c)$ satisfies for some large enough C

$$r \log \frac{r}{abcd/n} = Cc \log \frac{n}{c},$$

one could perhaps perform a more careful analysis to recover a $\log n$ factor (see [FKS89]). Perhaps one could also replace the $(C \log n)^{t/2}$ entirely by a small constant, but we don't see how to do this at this point.

References

[AKS87] M. Aitai, J. Komlós, and E. Szemerédi. Deterministic simulation in LOGSPACE. *STOC*, 1987.

[Alo86] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[Blu84] M. Blum. Independent unbiased coin flips from a correlated sources: a finite state markov chain. *FOCS*, pages 425–433, 1984.

[BS87] Andrei Broder and Eli Shamir. On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science*, pages 286–294, 1987.

[BT71] F. Bruhat and J. Tits. *Publ. Math. IHES*, 1971.

[CG85] B. Chor and O. Goldriech. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *FOCS*, pages 429–442, 1985.

[Chu88] F. R. K. Chung. Diameters and eigenvalues. preprint, 1988.

[CW89] A. Cohen and A. Wigderson. Dispersers, deterministic amplification, and weak random sources. *FOCS*, 1989.

[FKS89] J. Friedman, J. Kahn, and E. Szemerédi. On the second eigenvalue of random graphs. *STOC*, 1989.

[Fri88] J. Friedman. On the second eigenvalue of random d-regular graphs. Technical report, Princeton University, August 1988. To appear in *Combinatorica*.

[Fri89] J. Friedman. Some graphs with small second eigenvalue. Technical report, Princeton University, October 1989.

[Hal86] M. Hall. *Combinatorial Theory*. John Wiley & Sons, 2nd edition, 1986.

[IZ89] R. Impagliazzo and D. Zuckerman. How to recycle random bits. *FOCS*, 1989.

[LPS86] A. Lubotzky, R. Phillips, and P. Sarnak. Explicit expanders and the ramanujan conjectures. In *18th Annual ACM Symposium on Theory of Computing*, pages 240–246, 1986.

[Mar87] G. Margulis. Manuscript in Russian on graphs with large girth, 1987.

[NW88] N. Nissan and A. Wigderson. Hardness vs. randomness. *FOCS*, pages 2–11, 1988.

[Rob83] Alain Robert. *Introduction to the Representation Theory of Compact and Locally Compact Groups*. Cambridge University Press, Cambridge, 1983.

[San86] M. Santha. On using detrnimistic functions to reduce randomness in probabilistic algorithms. 1986. manuscript.

[Sip86] M. Sipser. Expanders randomness or time versus space. *Structure and Complexity Theory*, 1986.

[SV84] M. Santha and U. V. Vazirani. Generating quasi-random sequences from slightly random sources. *FOCS*, pages 434–440, 1984.

[VV85] U. V. Vazirani and V. V. Vazirani. Random polynomial time is equal to slightly random polynomial time. *FOCS*, pages 417–428, 1985.

[Wig55] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Math.*, 63(3):548–564, 1955.