Lecture 3-6: Spectral HDX and the Local-to-Global Method

April 29, 2025

Lecturer: Max Hopkins *Scribe: Max Hopkins*

1 High Dimensional Expanders

The past several lectures have showcased some of the classical successes of expanders graphs in computation, ranging from analysis of stochastic processes to the construction of good error correcting codes. While the literature contains countless further such examples (in networks, derandomization, PCPs, extractors, regularity decompositions, etc...), it also contains a litany of settings where expanders don't seem to achieve the desired results. Oftentimes this is due to a simple but inherent mismatch: many problems we care about in computer science (high arity CSPs, optimization of many variables, sampling multi-feature distributions...) are *high dimensional tasks*, while expanders are inherently *1-dimensional objects*. This suggests a natural question:

Can we push the success of expander graphs to higher dimensions?

In this mini-course we will develop the core concepts and applications of exactly such a theory of *high dimensional expanders* (HDX) that has seen incredible success over the past decade in resolving the above gap. We will see in particular the breakthrough application of this theory to rapid mixing of classical high dimensional walks (matroid bases), and develop a subset of the core tools used in further applications to solving high arity CSPs, list-decoding, quasilinear PCPs, fault tolerant network design, and (surprisingly) even to the construction of better 1-D expanders!

1.1 Weighted Graphs

Before jumping into high dimensional expanders, we need to spend just a bit more time on the low dimensional setting. In particular, in the previous lecture we covered spectral expansion first for regular graphs, then saw how to extend the theory to irregular graphs by weighting the adjacency matrix and inner product. Our final step in the low-dimensional case is to move to general *weighted* graphs, which will be important for handling the high dimensional setting:

Definition 1.1 (Weighted Graphs)**.** A weighted graph on vertex set $[n]$ is a pair (E, π_2) where $E \subset \binom{[n]}{2}$ and π_2 is a distribution over E.

With general weights, we have to be a bit careful how we define the normalized adjacency matrix. The correct way to do this is to define A as the transition matrix of the underlying random walk on G, that is we should take $A(v, w) = \mathbf{Pr}[v \to w]$ in this walk.

With this in mind, how does the walk on G proceed in the weighted setting? There is really only one natural option: starting at a vertex v, we should move to a neighbor w proportional to the weight of $\pi_2(v, w)$, that is

$$\mathbf{Pr}[v \to w] = \frac{\pi_2(v, w)}{\sum\limits_{u \sim N(v)} \pi_2(v, u)}$$

where we have normalized by the *total* weight around the vertex. We thus define the normalized adjacency matrix as

$$A(v,w) = \begin{cases} \frac{\pi_2(v,w)}{\sum\limits_{u \sim N(v)} \pi_2(v,u)} & \text{if } \{v,w\} \in E \\ 0 & \text{else} \end{cases}$$

Note that this is exactly the weighting we used last lecture for irregular graphs when we take the weights to be uniform. In this case the normalizing factor in each row is just $\frac{1}{\deg(v)}$.

As in the irregular case, A is self-adjoint when considered with respect to the correct weighted inner product. Namely, define the *marginal vertex distribution* of G, denoted π_1, by

1. Sampling a random edge $\{v, w\} \sim \pi_2$
2. Subsampling either v or w uniformly at random

and the corresponding inner product

$$\langle f, g \rangle = \mathbb{E}_{v \sim \pi_1}[f(v)g(v)].$$

Claim 1.2. *A is self-adjoint with respect to this inner product, namely:*

$$\langle f, Ag \rangle = \mathbb{E}_{\{v,w\} \sim \pi_2}[f(v)g(w)] = \langle Af, g \rangle$$

Why? Recall the action of A on g is to take the average value across g's neighbors with respect to the 'local distribution' around v

$$Ag(v) = \mathbb{E}_{w \sim \pi_1^v}[g(w)]$$

where

$$\pi_1^v(w) := \frac{\pi_2(v,w)}{\sum\limits_{u \sim N(v)} \pi_2(v,u)} = \frac{\pi_2(v,w)}{2\pi_1(v)}$$

You should really think of π_1^v as the distribution of w *conditioned on* v. Then the inner product

$$\langle f, Ag \rangle = \mathbb{E}_{v \sim \pi_1}[f(v)\mathbb{E}_{w \sim \pi_1^v}[g(w)]] = \mathbb{E}_{v \sim \pi_1, w \sim \pi_1^v}[f(v)g(w)]$$

but drawing $v \sim \pi_1$ from the marginal on vertices, then w conditioned on v is exactly like drawing an edge $\{v, w\} \sim \pi_2$, so this is indeed just $\mathbb{E}_{\{v,w\} \sim \pi_2}[f(v)g(w)]$. We'll make this formal in the bipartite setting momentarily, and leave the reader to verify the general claim directly.

We can now define spectral expansion. Namely by the above A has a spectral decomposition with eigenvalues

$$1 = \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n \geqslant -1$$

and we can define spectral expansion exactly as before as λ_2 (one-sided setting) or $\max\{\lambda_2, |\lambda_n|\}$ (two-sided setting).

The Distributional View of Expansion A very useful way to think about graphs and expansion is to think about a graph $G = (V, E)$ as a 2-dimensional distribution or *random variable*

$$X = \{X_1, X_2\},$$

taking values on unordered pairs $\binom{V}{2}$ generated by sampling a random edge the graph. In some sense in discussing conditional distributions and expectations above, we have already been using this view implicitly, but it is worth taking a moment to make the connection very explicit.

Let's focus for the moment on the setting of *bipartite graphs* where this distributional view is a bit cleaner. Here we can really view X as a standard 2-D joint random variable:

$$X = (X_1, X_2),$$

where X_1 is supported on L, X_2 is supported on R, and (X_1, X_2) is jointly distributed on $L \times R$:

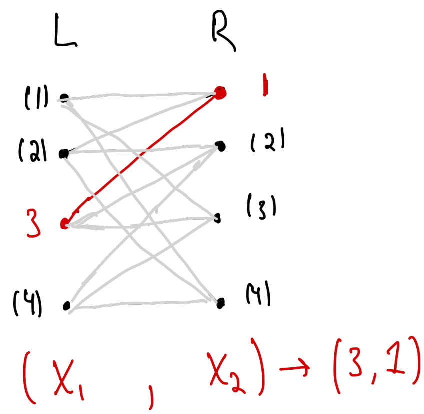

Figure 1: Bipartite graph $G = (L, R, E)$ vs. 2-D random variable (X_1, X_2)

Note in this view our claim about A being self-adjoint becomes a completely trivial application of the law of total probability. Why? For fixed $v \in L$, the local distribution π_1^v over R is *exactly* the distribution of X_2 conditioned on $X_1 = v$. Thus drawing $v \sim X_1$, then $w \sim X_2$ conditioned on v is exactly the same as drawing (v, w) from the joint distribution (X_1, X_2) (i.e. an edge) by the law of total probability. The same holds first sampling $w \in X_2$, then $v \in X_1$ conditioned on w.

How can we interpret expansion in this distributional view? In fact, we already saw in our previous lecture that spectral expansion is exactly a statement about the variables X_1 and X_2 appearing *approximately independent*. In other words, G is a λ-expander iff for all $f : L \to \mathbb{R}$ and $g : R \to \mathbb{R}$, the expectation of the product fg is close to the product of their expectations:

$$\mathbb{E}[fg] \in \mathbb{E}[f]\mathbb{E}[g] \pm \lambda\sqrt{\text{Var}(f)\text{Var}(g)}. \tag{1}$$

This is pretty amazing, since in reality for sparse graphs the variables are *extremely correlated* — for any fixed value of X_1, there may be as few as 3 possible values for X_2! Nevertheless, we saw there are ways to construct (degree-3) graphs for which these correlations 'cancel' on average, giving the above guarantee (indeed we saw 'almost all' 3-regular graphs satisfy this!).

1.2 Hypergraphs and Simplicial Complexes

To talk about high dimensional expansion, we'll first need to actually define what set of high dimensional objects we're interested in studying. In today and tomorrow's lectures, we'll focus on the simplest (but still widely powerful) setting of weighted *uniform hypergraphs*:

Definition 1.3 (Uniform Hypergraph)**.** A (weighted) uniform hypergraph on vertex set $[n]$ is a pair (H, π_d) where $H \subset \binom{[n]}{d}$ is a collection of (unordered) d-tuples, and π_d is a distribution over H.

Notice in the $d = 2$ case, where edges are pairs, H is simply a (weighted) graph. In the context of expansion, we will soon see it is useful to view H through the lens of *pure simplicial complexes*. Formally, this is just the downward closure of H in the following sense:

Definition 1.4 (Pure Simplicial Complex)**.** The weighted (pure) simplicial complex corresponding to a hypergraph (H, π_d) is the disjoint union

$$X_H = X(0) \cup X(1) \cup \ldots \cup X(d)$$

where $X(d) = H$ and $X(i) \subset \binom{[n]}{i}$ consists of all i-sets sitting in any d-set in $X(d)$ and is generated distributionally by the following downward closure process:

1. Sample a hyperedge $s = \{v_1, \ldots, v_d\} \in X(d)$ (distributed according to π_d).
2. Subsample $t = \{w_1, \ldots, w_i\} \subset s$ uniformly at random

Stated more explicitly, this results in a distribution π_i over $X(i)$ given by:

$$\pi_i(t) = \frac{1}{\binom{d}{i}} \mathbf{Pr}_{s \sim \pi_d}[t \subset s]$$

which we can view as the 'i-dimensional marginal' of X. We denote the resulting weighted simplicial complex by (X, π).

A simple and incredibly useful way to view X is through its inclusion diagram. In other words, draw X as a layered graph (one per $X(i)$), and connect each adjacent layer by set inclusion:

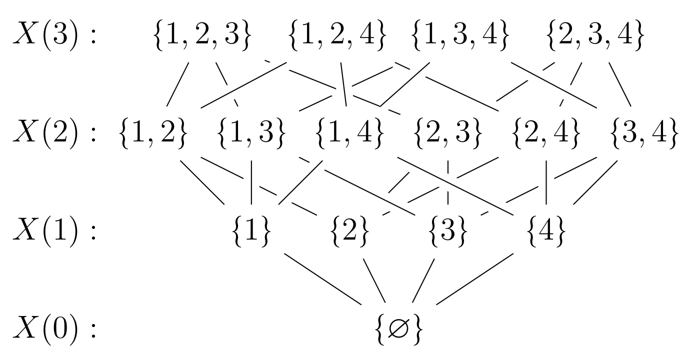

Figure 2: The complete 3-dimensional simplicial complex over 4 vertices.

The elements of X are typically referred to as *faces* rather than hyperedges (this stems from the geometric view of simplicial complexes). Departing from geometric conventions, we will say the *dimension* of a face $\sigma \in X$ is its size $|\sigma|$. This differs from geometric/topological notation, in which the sets of size i are thought of as i-simplices (e.g. sets of size 3 are triangles, sets of 4 are pyramids, etc.) and therefore have dimension $i-1$. In probabilistic and combinatorial analysis however the latter notation becomes extremely cumbersome, so with apologies to the math community we will just use dimension to refer to size.

Finally, now that we have our equivalent notion of pure simplicial complexes, we will typically just drop the initial hypergraph H and deal with X directly.

The Distributional View There is of course a natural analog of the random variable view in the high dimensional setting, where we think of a d-dimensional simplicial complex X as a *d-dimensional random variable*:

$$X = \{X_1, \dots, X_d\}$$

supported on $\binom{[n]}{d}$ generated by drawing a random hyperedge from H.

Again this is perhaps most convenient when X is *partite*, meaning there is a partition of the vertices $X(0)$ into d disjoint subsets

$$X(0) = P_1 \amalg P_2 \amalg \dots \amalg P_d$$

such that every top level face $\sigma \in X(d)$ has exactly one vertex from each part. In this case, thinking of the parts as corresponding to d 'coordinates', drawing a random d-face gives a joint d-dimensional random variable

$$X = (X_1, \dots, X_d)$$

where each X_i is supported on P_i, and X is therefore supported on the product $P_1 \times \dots \times P_d$.

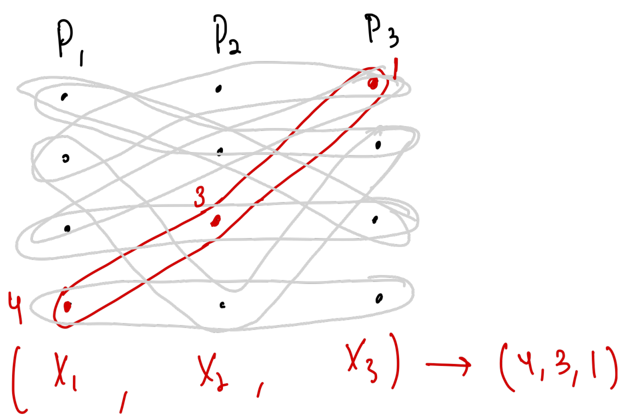

Figure 3: 3-Partite Complex vs. 3-dim random variable (X_1, X_2, X_3)

We note this connection goes both ways: any d-dimensional variable $(X_1, \dots, X_d)$ can also be written as a partite simplicial complex. Namely writing Supp$(X_i) = P_i$, we have a one-to-one correspondence with the d-partite complex X_P whose vertex parts are $P_i \times \{i\}$, with top level faces

$$X_P(d) = \{\{(x_1, 1), \dots, (x_d, d)\} : x_i \in P_i \wedge (x_1, \dots, x_d) \in \text{Supp}(\pi_P)\}$$

and corresponding weights $\pi_P(\{(x_1, 1), \dots, (x_d, d)\}) = \mathbf{Pr}[(X_1, \dots, X_d) = (x_1, \dots, x_d)]$.

This correspondence means we can write many classical objects of study in theoretical computer science like the hypercube (i.e. product of Bernoulli r.v.'s) as partite simplicial complexes:

$$X_{\text{hypercube}}(0) = \{0,1\} \times [d],$$
$$X_{\text{hypercube}}(d) = \left\{\{(x_1, 1), \dots, (x_d, d)\} : (x_1, \dots, x_d) \in \{0,1\}^d\right\}.$$

Indeed in the following lectures we will see as a general theme how many traditional results in TCS on the cube and beyond (Fourier analysis, concentration of measure, mixing of random walks) have a beautiful unifying theory through the lens of simplicial complexes and HDX.

1.3 Spectral HDX

When should we call a hypergraph expanding? Let's first focus on the partite case. Above, we reviewed how graph expansion in this setting captures whether the variables X_1 and X_2 are "approximately independent". This suggests a very natural way to extend expansion to partite hypergraphs: we should call X expanding if all d variables in $(X_1, \dots, X_d)$ 'look independent'.

Now there are several reasonable ways to go about formalizing this intuition. One natural way is to require all the *pairwise marginals* (X_i, X_j) in $(X_1, \dots, X_d)$ are approximately independent. Another would be to require some direct extension of Equation (1) by requiring that $\forall \{f_i : P_i \to \mathbb{R}\}_{i \in [d]}$:

$$\mathbb{E}\left[\prod_{i=1}^{d} f_i(x_i)\right] \approx \prod_{i=1}^{d} \mathbb{E}[f_i]. \tag{2}$$

Both of these are useful but miss a critical property of independence that only shows up in dimension 3 and beyond: the behavior of X under *conditioning*.

Let's take a moment to discuss. A key property of independent variables is that conditioned on the value of any one variable (say $X_1 = a$), the marginal distribution over the rest $(X_2 \dots X_d)$ *remains independent*. Incorporating this requirement leads to the following natural notion of (partite) high dimensional expansion called a λ-*product*:

Definition 1.5 (λ-Products [GLL21]). A d-partite complex X is a λ-product if for all $i, j \in [d]$:

1. (X_i, X_j) is a λ-spectral expander
2. This remains true under all (valid) conditionings[1] $X_S = y_S$ for $i, j \notin S$.

Remark 1.6. Why haven't we chosen to work with Equation (2) directly in the above? It turns out it is possible to derive a form of Equation (2) from the pairwise guarantee above (and vice-versa). In general the pairwise notion tends to be somewhat cleaner in practice and more closely related to the standard notion of 'local spectral expansion' we'll cover now.

The General Case: Let's now drop the partite requirement and return to the case of general pure simplicial complexes. Is there a natural adaptation of 'pairwise independence' when $X = \{X_1, \dots, X_d\}$ is coordinate-free?

[1] In other words, y_S should actually appear in the support of X_S.

In fact, there is a very natural way to look at pairwise correlations in a general complex X: instead of looking at a specific 2D marginal (X_i, X_j) (which no longer makes sense without coordinates), we should look at the distribution of a *random pair* chosen from $\{X_1, \ldots, X_d\}$. The astute reader might observe this is actually the second time we've seen this definition; it is *exactly* how we defined the (weighted) edges of X! In other words, combinatorially, the natural '2-D marginal' of $\{X_1, \ldots, X_d\}$ is nothing more than X's *1-skeleton* or *underlying graph* $G_X = (X(1), X(2), \pi_2)$:

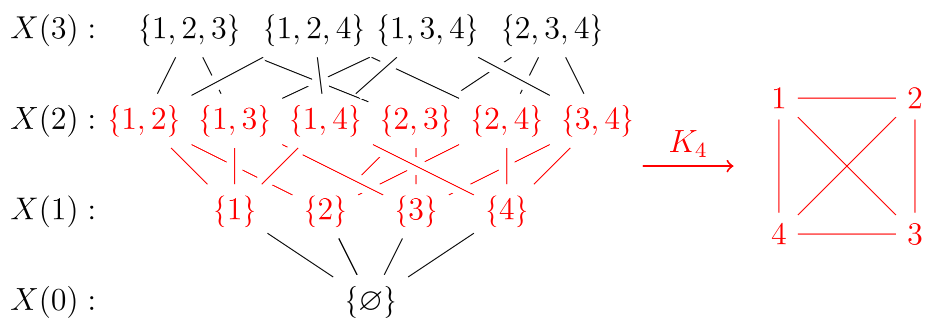

Figure 4: The 1-skeleton of X, here a copy of K_4.

Thus, the natural analog of our first HDX condition is to require the 1-skeleton of X expands. As before, we will also require that this holds *locally* within X even after conditioning on the value of a subset of the variables $\{X_1, \ldots, X_d\}$. These local conditionings are usually called the *links* of X, and can be thought of in two equivalent ways. Distributionally, the link of a face $\tau \in X$ is simply the $(d - |\tau|)$-dimensional distribution generated by conditioning $\{X_1, \ldots, X_d\}$ on containing τ, and marginalized over the remaining unset values. Combinatorially, the link is the *local neighborhood* of τ, the subcomplex of X generated by taking the collection of faces in X containing τ (conditioning), then removing τ (marginalizing over remaining variables):

Definition 1.7 (Link)**.** The link of $t \in X$ is the $(d - |t|)$-dimensional pure simplicial complex generated by taking all faces in X that include t and removing t:

$$X_t := \{s \backslash t : t \subset s \in X\}.$$

As discussed, the corresponding (top level) distribution over X_t is simply the induced distribution given by conditioning on containing t, i.e.

$$\pi^t_{d-|t|}(s) = \frac{\pi(s \cup t)}{\underset{s' \sim X(d)}{\mathbf{Pr}}[t \subset s']},$$

and the level-i distribution π^τ_i is generated as before by sampling $s \sim \pi^t_{d-|t|}$, then sub-sampling i elements from s uniformly at random. The link of an i-face $t \in X(i)$ is often called an 'i-*link*'.

We are now ready to define the modern notion of spectral high dimensional expansion in full generality: X a spectral HDX if the 1-skeleton of every link expands.

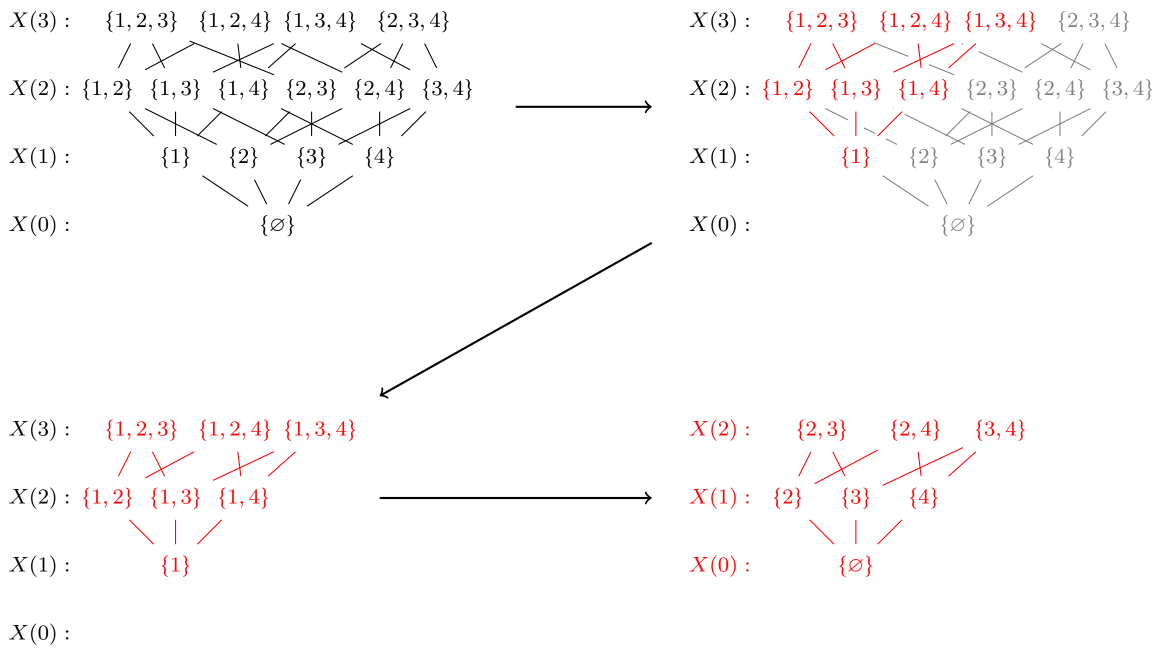

Figure 5: The link of vertex $\{1\}$ in X, in this case the graph K_3.

Definition 1.8 (Local-Spectral Expansion [KKL16, EK16, Opp18])**.** A d-dimensional pure simplicial complex X is called a (one-sided) λ-local-spectral expander if the 1-skeleton of every i-link for $0 \leqslant i \leqslant d-2$ is a (one-sided) λ-spectral expander.[2]

We emphasize local-spectral expansion is natural from both a combinatorial and probabilistic perspective. Combinatorially, it says the base graph of X and all its local neighborhoods expand. Probabilistically, it promises X is 'product-like' in the sense that all pairwise correlations in X are (spectrally) bounded, even after conditioning on a subset of the variables.

Remark 1.9 (A Historical Note). Local-spectral expansion originally arose in the context of algebra and topology, in particular the Ramanujan complexes of [CSŻ03, Li04, Sar04, LSV05] and the study of Gromov's topological overlap property [Gro10, KKL16, EK16].[3] While Definition 1.8 as stated is from [Opp18], essentially equivalent notions bounding combinatorial link expansion appeared earlier in [KKL16, EK16, KM16].[4]

Remark 1.10 (Local-Spectral vs γ-Product). When X is partite, local-spectral expansion is roughly the same definition as our γ-product notion from before. Looking at the 1-skeleton, the difference lies in whether one bounds the global spectrum of the entire multipartite graph (local-spectral), or looks individually at the bipartite components (λ-product). In the exercises, you will show these definitions are roughly the same up to factors in dimension.

1.4 Examples of High Dimensional Expanders

Let's now take a look at a few concrete examples of HDX. Just like in the graph setting, the first natural example is to look at the hypergraph with all possible edges, the *complete complex*:

[2]Note this includes the 1-skeleton of X itself, which is the link of the emptyset.
[3]See also related work on homological properties of simplicial complexes [Gar73, MW09].
[4]See e.g. [EK16, Theorem 1.10] and the strengthened notion of skeleton expansion in [KM16].

Example 1.11 (Complete Complex)**.** *The k-dimensional complete complex on n vertices $\Delta_n(d)$ is a 0-one-sided and $\frac{1}{n-d+1}$-two-sided HDX*

Proof. Every i-link of the complete complex is just the $(d-i)$-dimensional complete complex on $n-i$ vertices, so it is enough to argue the 1-skeleton of $\Delta_n(d)$ is a 0-one-sided and $\frac{1}{n-1}$-two-sided expander. This is immediate from the fact that the 1-skeleton is simply the complete graph K_n whose only non-trivial eigenvalues is $-\frac{1}{n-1}$. □

Indeed in the non-partite case, HDX can naturally be viewed as 'sparse models' for the complete complex in the same way expanders are sparse models for the complete graph. In general, it is a good motivating principle that any property which holds in the complete complex should also hold on strong enough HDX (so long as that property does not inherently require high degree).

Another basic example (indeed our initial motivating example) are *product spaces* like the hypercube. In the world of simplicial complexes these correspond to complete partite complexes.

Example 1.12 (Complete k-Partite Complex (Product Space))**.** *The complete partite complex (on any number of vertices and parts) is a one-sided 0-local-spectral expander.*

Proof. All links have complete multipartite 1-skeletons with uniform weights. One can check such graphs are one-sided 0-spectral expanders. □

The above are somewhat "uninteresting" examples of HDX (or more accurately, they are exactly the objects HDX are supposed to model). Many more interesting examples arise from looking at weakly dependent combinatorial structures on graphs (e.g. independent sets, matchings...). For instance, a famous example is the set of *spanning trees* of an arbitrary graph:

Example 1.13 (Spanning Trees)**.** *Fix a graph $G=(V,E)$, and consider the $(n-1)$-dimensional complex $X^G(n-1)$ whose vertices $X^G(1)$ are the edges of G, and whose $(n-1)$-faces are all sets of edges that form a spanning tree of G. X^G is a one-sided 0-local-spectral expander.*

Proof. We leave the proof as an exercise. □

In fact, it turns out that the above can be generalized to any simplicial complex corresponding to bases of a matroid (an important observation due to [ALOV19]). The proof of this is actually completely elementary given the definition of a matroid and the results of the next section and is also left as an exercise.

Sparse HDX: The above examples of HDX are very *high degree* in the sense that each vertex is contained in a huge (unbounded) number of hyperedges. Usually in application to coding and complexity we will pay a cost (e.g. in rate) based on the number of hyperedges of X. This motivates us to look for *bounded degree* HDX, where every vertex appears in only $O_d(1)$ faces. The first thing we might think to try to build such complexes is to look at the *random* case; this gave us nearly perfect expansion in the graph setting. Interestingly, while a random complex will satisfy the base notions of expansion we discussed (e.g. pairwise marginals or expander-mixing type lemma), they fail dramatically in the links:

Example 1.14 (Failure of Random Hypergraphs)**.** *Let $X(k, n, p)$ denote the k-uniform hypergraph in which every hyperedge is included independently with probability p. Fix any $d \in \mathbb{N}$. For all large enough n, $X(3, n, d/n)$ has a disconnected link with high probability. Similarly in k-dimensions, this means even a random complex of size roughly n^{k-1} is not a high dimensional expander!*

Why are random complexes so poorly behaved in this sense? Morally, the issue is that HDX should be globally sparse, but locally dense (s.t. expansion can hold in links). This tension cannot be realized by a naive random object, since the hyperedges will be well spread around the complex.

We emphasize this simple example illustrates an important philosophical difference between expanders and HDX. Expander graphs, while exceedingly useful, are *common* objects; we use them when we need an *explicit* way to mimic the properties of a sparse random network. High dimensional expanders are *rare* and *structured* objects; we (usually) use them when we need properties that go *beyond random*, such as in the study of local testability and PCPs.

Returning to the question of constructions, we also discussed how expanders can be built through either combinatorial or algebraic means. The latter turns out to generalize nicely to higher dimensions. Indeed this was done independently by several groups [CSŻ03, Li04, Sar04, LSV05] well before the above formalization of modern local-spectral expansion:

Theorem 1.15 (Bounded Degree HDX [LSV05])**.** *For every dimension d and $\lambda > 0$, there exists an infinite family $\{X_n\}$ of explicit λ-HDX with degree at most $\lambda^{-O(d^2)}$.*

We now know more elementary constructions of HDX based on simple matrix groups [KO18] (see [HS19] for a simplified exposition and analysis). We will not focus on constructions in these lectures, but we remark a number of fascinating questions remain like the existence of combinatorial constructions or understanding to what extent current constructions are optimal in terms of degree.

2 Trickling-Down and the Local to Global Method

Hopefully at this point you are convinced that spectral high dimensional expansion is a natural definition — let's show it's a useful one!

The core of essentially all variants and applications of high dimensional expansion is the *local-to-global method*, often attributed to early work of Garland [Gar73] in the 70s. At a high level, the paradigm follows a simple framework: say we are given a complex X and want to analyze some global quantity Q — perhaps the spectral gap of a high dimensional random walk, or local testability of a related code. Instead of trying to analyze Q directly, we'll break Q into an expectation over simpler local components $\{Q_v\}$ across the links $\{X_v\}$, analyze each local component directly using X's link expansion, and finally piece the results back together into a global bound on Q.

This is, admittedly, a bit abstract, but should become clear in the following lectures as we see several major concrete applications of the paradigm to bounding the 1-skeleton expansion of X and the spectral gap of global random walks using their localizations to links.

2.1 The Trickling Down Theorem

Our first application of the local-to-global method is an exceedingly useful result of Oppenheim [Opp18] called the 'Trickling-Down Theorem'. Recall that, as defined, local-spectral expansion is really both a local and global property. In particular it requires both expansion of the 1-skeleton of X itself (which measures *global* pairwise correlation in X), as well as the expansion of the links (which measure *local* correlations within X). Oppenheim showed that up to a small quantitative loss, it is actually possible to deduce a bound on former just from local information on the links!

Theorem 2.1 (Trickling Down Theorem (3D case))**.** *Let X be a 3-dimensional complex such that*

- *Every vertex link is a (one-sided) λ-expander*
- *The 1-skeleton of X is connected*

Then the 1-skeleton of X is a (one-sided) $\frac{\lambda}{1-\lambda}$-expander.

Applying this result inductively within X we get the full Trickling-Down Theorem: any connected complex with expanding co-dimension[5] 2 links is automatically a full local-spectral HDX!

Corollary 2.2 (Trickling Down Theorem)**.** *Let X be a d-dimensional complex such that*

- *Every $(d-2)$-link is a (one-sided) λ-expander*
- *The 1-skeleton of every link is connected*

Then X is a (one-sided) $\frac{\lambda}{1-(d-2)\lambda}$-local-spectral HDX.

Proof. We leave the proof as an exercise. □

Theorem 2.2 implies any connected complex whose co-dim 2 links have expansion better than $\frac{1}{d-1}$ are expanding in all links. This threshold is tight. By tensoring a poorly expanding graph G with the complete complex (and weighting faces appropriately), it is possible to construct complexes with co-dim 2 expansion *exactly* $\frac{1}{d-1}$ whose 1-skeletons have arbitrarily poor expansion [Gol21].

2.1.1 Garland's Method and the Proof of $d = 3$ Case

Let's now prove the $d = 3$ case of the Theorem. Write the normalized adjacency matrix of the 1-skeleton of X as $A = A_{\varnothing}$, and the 1-skeleton of the link of v as A_v. Recall $A \in \mathbb{R}^{X(1)\times X(1)}$ is the stochastic matrix corresponding to the underlying random walk on X's 1-skeleton with edge-weights π_2:

$$A(v,w) = \begin{cases} \frac{\pi_2(v,w)}{\sum\limits_{w' \in X_v(1)} \pi_2(v,w')} & \text{if } \{v,w\} \in X(2) \\ 0 & \text{Otherwise} \end{cases}$$

Like the graph case, the probability $\frac{\pi_2(v,w)}{\sum\limits_{w' \in X_v(1)} \pi_2(v,w')}$ is really nothing more than the weight of w 'viewed from' v, i.e. the link probability $\pi_1^v(w)$, so we can also write:

$$A(v,w) = \begin{cases} \pi_1^v(w) & \text{if } \{v,w\} \in X(2) \\ 0 & \text{Otherwise} \end{cases}$$

[5]The *co-dimension* of a link X_t is simply $d - |t|$. Thus co-dim 2 links are the 'top level' links that fix all but two of the variables.

We encourage the reader to actually prove this — it is a simple exercise and we will see a similar calculation later in the proof below.

Returning to the task at hand, our goal is now to bound $\lambda_2(A)$ which we recall from the prior lecture is the maximizer of A's Rayleigh quotient across $f \perp 1$:

$$\lambda_2(A) = \max_{f \perp 1, ||f||=1} \langle Af, f \rangle$$

where $\langle \cdot, \cdot \rangle$ is the weighted inner product:

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{v \sim \pi_1} [f(v)g(v)].$$

Let f denote the unit normalized eigenvector corresponding to $\lambda_2(A)$. Our goal is to show:

$$\langle f, Af \rangle \leqslant \frac{\lambda}{1 - \lambda}$$

Since we only have spectral information about the *links* of X, the crucial observation is that the left-hand *global* Rayleigh quotient can be broken into an expectation over *local* Rayleigh quotients of the links, each of which can then be analyzed by standard graph expansion. This is usually referred to as *Garland's method*:[6]

Claim 2.3 (Garland's Method)**.** *For any complex X and function $f : X(1) \to \mathbb{R}$:*

1. $\langle f, f \rangle = \mathbb{E}_{v \sim \pi_1} [\langle f_v, f_v \rangle_{X_v}]$
2. $\langle Af, f \rangle = \mathbb{E}_{v \sim \pi_1} [\langle A_v f_v, f_v \rangle_{X_v}]$

where $f_v : X_v \to \mathbb{R}$ is the restriction of f to X_v defined as $f_v(w) = f(w)$.

Let's now prove the Theorem assuming this claim. It suffices to show

$$\mathop{\mathbb{E}}_{v \in X(1)} [\langle A_v f_v, f_v \rangle_{X_v}] \leqslant \frac{\lambda}{1 - \lambda}$$

Looking at the LHS, we know $\lambda_2(A_v) \leqslant \lambda$ so for any $g : X_v(1) \to \mathbb{R}$ s.t. $\mathop{\mathbb{E}}_{X_v}[g] = 0$:

$$\langle A_v g, g \rangle \leqslant \lambda \langle g, g \rangle.$$

We'd like to apply this with $g = f_v$ above, but there's a problem: while $\mathbb{E}[f] = 0$ *globally*, we have no guarantee that the localizations $\mathop{\mathbb{E}}_{X_v}[f_v] = 0$ on the links.

We therefore take the standard strategy we saw e.g. in our previous analysis of random walks: we'll split f_v into parallel and perpendicular components $f_v^{||} = \mathbb{E}[f_v] \cdot \mathbf{1}_v$ and $f_v^{\perp} = f - f_v^{||}$, handle the latter using expansion, and separately argue the former has small contribution on average.

Toward this end, recall by orthogonality of $f_v^{||}$ and $f_v^{\perp}$ we can write:

$$\langle A_v f_v, f_v \rangle = \langle A_v (f_v^{||} + f_v^{\perp}), f_v^{||} + f_v^{\perp} \rangle = \langle A_v f_v^{||}, f_v^{||} \rangle_{X_v} + \langle A_v f_v^{\perp}, f_v^{\perp} \rangle_{X_v}$$

[6] This references an old work of Garland showing vanishing of cohomology of certain complexes using this method [Gar73].

Applying this inside our global expectation gives:

$$\begin{aligned}
\mathop{\mathbb{E}}_{v\in X(1)}\left[\langle A_v f_v, f_v\rangle_{X_v}\right] &= \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle A_v f_v^{\|}, f_v^{\|}\rangle_{X_v}\right] + \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle A_v f_v^{\perp}, f_v^{\perp}\rangle_{X_v}\right] \\
&\leqslant \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle A_v f_v^{\|}, f_v^{\|}\rangle_{X_v}\right] + \lambda \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v^{\perp}, f_v^{\perp}\rangle_{X_v}\right] \\
&= \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v^{\|}, f_v^{\|}\rangle_{X_v}\right] + \lambda \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v^{\perp}, f_v^{\perp}\rangle_{X_v}\right] \\
&= (1-\lambda)\mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v^{\|}, f_v^{\|}\rangle_{X_v}\right] + \lambda \mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v, f_v\rangle_{X_v}\right] \\
&= (1-\lambda)\mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v^{\|}, f_v^{\|}\rangle_{X_v}\right] + \lambda,
\end{aligned}$$

where the last step follows from Garland's method and the fact that $\langle f, f\rangle = 1$. It is left to analyze the parallel term. Recall that $f_v^{\|} = \mathop{\mathbb{E}}_{X_v}[f_v]\vec{1}_v$. Thus

$$\langle f_v^{\|}, f_v^{\|}\rangle_{X_v} = \mathop{\mathbb{E}}_{X_v}[f_v]^2\langle 1_v, 1_v\rangle = \mathop{\mathbb{E}}_{X_v}[f_v]^2$$

The key is now to realize that $\mathop{\mathbb{E}}_{X_v}[f_v]$ is just the average value of f over the neighborhood of v, which is just $Af(v)$!

$$\mathop{\mathbb{E}}_{X_v}[f_v] = Af(v).$$

This allows us to argue the parallel term is small on average, since

$$\mathop{\mathbb{E}}_{v\in X(1)}\left[\langle f_v^{\|}, f_v^{\|}\rangle_{X_v}\right] = \langle Af, Af\rangle = \lambda_2^2.$$

Altogether, we get the following bound on λ_2:

$$\lambda_2 = \langle Af, f\rangle \leqslant (1-\lambda)\lambda_2^2 + \lambda.$$

Solving this quadratic inequality yields either that $\lambda_2 \geqslant 1$ or $\lambda_2 \leqslant \frac{\lambda}{1-\lambda}$. Since we are promised that X is connected, we know that $\lambda_2 < 1$ which completes the proof.

Remark 2.4 (On Connectivity). While it may seem like requiring of connectivity on X is an artifact of the proof, the assumption is necessary. Namely, take X to be the disjoint union of two local spectral expanders. The graph underlying X is clearly disconnected, but every link is still expanding by definition.

Proof of Claim 2.5: Garland's Method. Recall our goal is to show

1. $\langle f, f\rangle = \mathbb{E}_{v\sim\pi_1}[\langle f_v, f_v\rangle_{X_v}]$
2. $\langle f, f\rangle = \mathbb{E}_{v\sim\pi_1}[\langle A_v f_v, f_v\rangle_{X_v}]$

In some way these statements (especially the first) are 'automatic' due to how we've set up the local and global distributions π and π^v. However, it takes time to get used to these sort of manipulations and they are core to the analysis of HDX so we take some time to discuss them carefully below.

Let's take a look at Item (1). Expanding out the inner product as expectations, our goal is to prove that for any f:

$$\mathbb{E}_{w\sim\pi_1}[f(w)^2] = \mathbb{E}_{v\sim\pi_1}\mathbb{E}_{w\sim\pi_1^v}[f(w)^2]$$

In other words, we need to show that the distribution over w generated by process 1) Sample $v \sim \pi_1$ then 2) Sample $w \sim \pi_1^v$, is exactly the marginal vertex distribution π_1.

Morally, this should be 'obvious' — after all we defined the link distribution π^v exactly to be the law of π conditioned on v. Thus if we look at the distribution of π_1^v over a random $v \sim \pi_1$, the law of total probability says we should just get back π. That said, we've played a bit fast and loose in defining π and π^v at different levels of the complex, so let's actually formalize this.

Fix some $w \in X(1)$. The probability of sampling w under the above process is exactly:

$$\begin{aligned}
\sum_{v\sim w} \pi_1(v)\pi_1^v(w) &= \sum_{v\sim w} \frac{1}{k}\mathbf{Pr}_{s\sim\pi_d}[v\in s]\cdot \frac{1}{k-1}\frac{\mathbf{Pr}_{s\sim\pi_d}[\{v,w\}\in s]}{\mathbf{Pr}_{s\sim\pi_d}[v\in s]} \\
&= \frac{1}{2}\sum_{v\sim w}\frac{\mathbf{Pr}_{s\sim\pi_d}[\{v,w\}\in s]}{\binom{k}{2}} \\
&= \frac{1}{2}\sum_{v\sim w}\pi_2(v,w) \\
&= \pi_1(w)
\end{aligned}$$

The last step follows because we can always generate π_1 by sampling from any π_i (including in particular π_2), then uniformly sampling a vertex.

We remark that a similar way to think about the above method is that sampling v, then w from the link of v is (up to ordering) the same procedure as just sampling the edge $\{v, w\} \sim \pi_2$. Thus we may think of the outer probability as sampling a random edge, and taking the expectation of f over a random endpoint which should be distributed as π_1 by definition. Formalizing this interpretation follows essentially exactly the same steps as above.

Let's now look at Item (2). This actually follows from almost exactly the same argument, except the inner expectation will be over an *edge* instead of a vertex. In particular, recall that by definition of the weighted adjacency matrix A, we may write

$$\langle Af, f\rangle = \mathbb{E}_{\{u,w\}\sim\pi_2}[f(u)f(w)]$$

Similarly, we have the righthand side

$$\mathbb{E}_{v\sim\pi_1}[\langle A_v f_v, f_v\rangle_{X_v}] = \mathbb{E}_{v\sim\pi_1}[\mathbb{E}_{\{u,w\}\sim\pi_2^v}[f(u)f(w)]]$$

Thus we need only check that the distribution of $\{u, w\}$ generated by sampling v, then $\{u, w\}$ conditioned on v is the same as sampling $\{u, w\}$ from the marginal π_2. Similarly, one can argue that drawing $\{v, u, w\}$ in this way is distributed as a random triangle from π_3. The calculations are essentially the same as for Item (1), so we leave this to the reader to verify.

3 Random walks on HDX

Previously we saw expander graphs are intimately connected to the mixing time of their underlying random walk. It is natural to wonder whether this extends in some way to high dimensions: does local-spectral expansion imply the mixing of *higher order* random walks on X?

First, what even is the natural generalization of the random walk on a graph G to a simplicial complex X? To answer this, let's look at the graph G as a 2-D simplicial complex. The (lazy) random walk on G has a natural interpretation on this diagram. We start at a vertex $v \in X(1)$, move up to an edge $e = \{v, w\} \in X(2)$ (with probability $\pi_1^v(w)$), then move back down to a uniformly random vertex $\{v\}$ or $\{w\}$ in e.

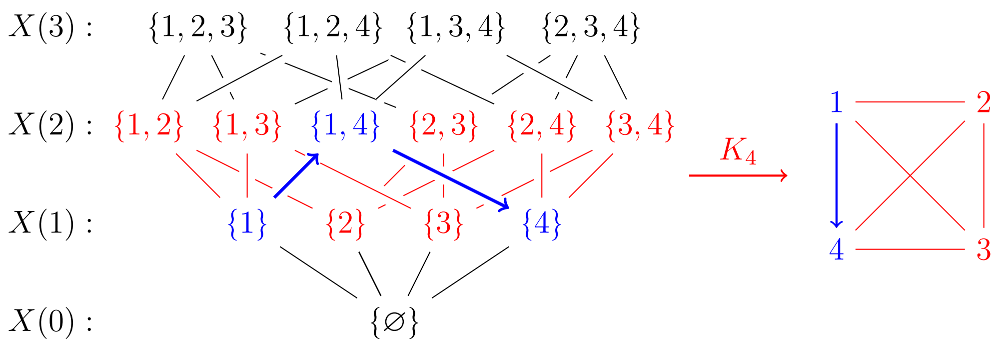

Figure 6: Random walk on 1-skeleton of X; simplicial vs standard view.

Now it is simple to observe this process has very natural high order analogs starting at any level i of the complex. We start at $s \in X(i)$, walk 'up' in the complex by adding a vertex v such that $s \cup v \in X(i+1)$ (distributed according to the link π_1^s), then move back down to $s' \in X(i)$ by removing a random vertex from $s \cup v$. Similarly, we can define the 'down-up' walk by first removing a random vertex to go to level $i-1$, then adding a vertex conditional on returning to a face in $X(i)$.

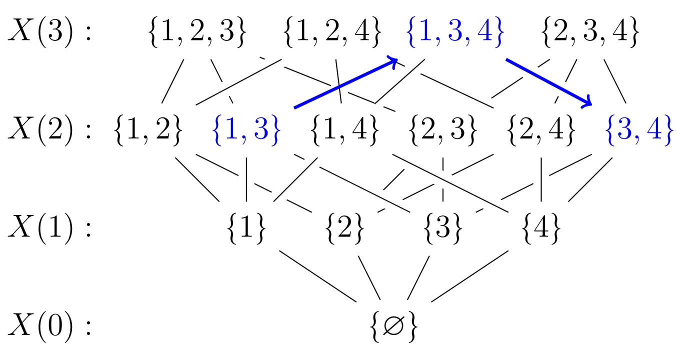

Figure 7: Edge-Triangle-Edge (up-down) walk on X.

High order random walks were introduced by Kaufman and Mass [KM16] in hopes of building improved PCPs from high dimensional expanders, a motivation that was realized only recently after a breakthrough line of works on the topic [DK17, GK22, DD23, BM24, DDL24, BLM24, BMV24]. Incredibly, PCPs are not the only breakthrough application of these walks. In 2018, Anari, Liu, Oveis-Gharan, and Vinzant [ALOV19, ALO20] realized these walks (in particular the down-up walk) actually generalize a natural and broadly studied class of Markov chains in TCS and statistical physics called the (single-site) *Glauber Dynamics*, leading to the resolution of a major conjecture regarding expansion of certain matroid-based chains [MV89] and an explosion of progress in MCMC analysis and approximate sampling. We'll touch on both of these breakthrough lines of work in the next two lectures, but for now we'll focus on the latter.

Let's look at the down-up walk. To see how it generalizes classical objects in TCS, let's return to our prototypical example of a 'perfect' (partite) HDX, the hypercube:

$$X_{\text{hypercube}}(0) = \{0,1\} \times [d],$$
$$X_{\text{hypercube}}(d) = \left\{\{(x_1, 1), \ldots, (x_d, d)\} : (x_1, \ldots, x_d) \in \{0,1\}^d\right\}.$$

What is the down-up walk on $X_{\text{hypercube}}$? It is exactly the procedure which, given a binary string $x \in \{0,1\}^d$, picks a random coordinate i and re-samples the value of x_i. This is, of course, just the (lazy) hypercube graph — a simple but fundamental object in TCS.

Let's take a look at a more interesting example: the spanning-tree complex. Here the down-up walk corresponds to the simple procedure of removing a random edge from the spanning tree, then adding back a random edge conditioned on creating a new spanning tree:

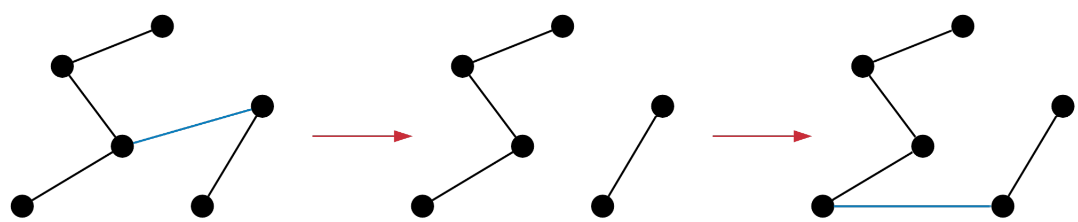

Figure 8: A single step of the down-up walk for spanning trees. A uniformly random edge is removed, then another added such that the result remains a spanning tree.

What can we say about the mixing time of these walks? While it is obvious the down-up walk on any product space mixes in polynomial time (indeed in roughly $\frac{1}{2}d\log(d)$ steps, we need only ensure each coordinate has been sampled at least once), it's substantially less clear this is true for spanning trees! Nevertheless, as we've discussed, the spanning-tree complex is a 0-local-spectral HDX, meaning its variables 'act' independent in a very strong sense. Can we therefore recover the mixing properties of products?

Let's start somewhere a bit more modest: spectral analysis. Basic Fourier analysis tells us that the second eigenvalue of the lower walk on any product (think the hypercube for simplicity) is

$$\lambda_2(P_d^\vee) = 1 - \frac{1}{d}. \tag{3}$$

Combined with the classic connection between expansion and mixing we saw in previous lectures, extending this from products to general 0-local-spectral HDX would at least imply *quadratic* mixing time e.g. for spanning trees since

$$T_{mix}(P_d^\vee, \varepsilon) \leqslant \frac{1}{1-\lambda_2} \log \frac{1}{\varepsilon \pi_{\min}} \approx d^2 \log(d)$$

In fact, to highlight how non-trivial such an extension would be, since general matroid bases are 0-local-spectral HDX extending Equation (3) to this case would resolve an old and storied conjecture of Mihail and Vazirani from the late 80s [MV89] on the edge expansion of the basis-exchange (down-up) walk on matroids [ALOV19] (combined with Cheeger and some basic manipulations).

In a breakthrough line of work, Kaufman and Mass [KM16], Dinur and Kaufman [DK17], Kaufman and Oppenheim [KO20], and finally Alev and Lau [AL20] showed the following elegant extension of Equation (3) not just to 0-local-spectral HDX like matroids, but to *any* HDX:

Theorem 3.1 ([AL20]). *Let X be a $(\gamma_0, \ldots, \gamma_{d-2})$-one-sided HDX. For any $1 \leqslant i \leqslant d$:*

$$\lambda_2(P_d^\vee) \leqslant 1 - \frac{1}{d}\prod_{j=0}^{d-2}(1-\gamma_j)$$

Where γ_j is the worst 1-skeleton expansion across j-links of X.

In particular, this indeed means any 0-local-spectral expander indeed has expansion $\lambda_2(P_d^\vee) \leqslant 1 - \frac{1}{d}$, matching the hypercube and resolving the Mihail-Vazirani conjecture [ALOV19].[7]

Remark 3.2 (Rapid Mixing and Log-Sobolev). Using more advanced tools (in particular stronger functional inequalities like modified log-sobelev), it turns out one can prove a faster mixing bound of $\widetilde{O}(d)$ for walks on matroids and beyond. This can actually be proved via an extremely similar local-to-global method as the proof of Theorem 3.1 we'll cover below, but requires replacing the local spectral conditions with local entropic conditions. We will not cover this in these lectures, but there is a beautiful theory called *entropic independence* [AJK$^+$21] capturing this phenomenon.

Remark 3.3 (Fourier Analysis). Classical Fourier analysis on products also goes far beyond the basic analysis of $\lambda_2(P_d^\vee)$, giving (for instance) a complete explicit description of *all* the eigenvectors and eigenvalues of this walk along with powerful higher moment tools like hypercontractivity. In the final lecture we will see that under sufficiently strong local-spectral expansion it is actually possible to extend this entire theory to HDX!

Random Walks and the 'Up' and 'Down' Operators: Let's now be a bit more careful with how we define our random walks. One very convenient formalization is through a composition of *averaging operators* that lift and lower functions between levels of X. Write $C_k = C_k(X, \mathbb{R})$ to denote the space of functions $\{f : X(k) \to \mathbb{R}\}$. Morally, the averaging operators simply act on functions by averaging up or down in the inclusion diagram. The *Up Operator* U_i lifts $f \in C_i$ to $U_i f \in C_{i+1}$ by averaging over lower neighbors.

$$U_i f(s) = \mathop{\mathbb{E}}_{t \subset s:\ |t|=i}[f(t)]$$

Similarly, the *Down Operator* D_i lowers $f \in C_i$ to $D_i f \in C_{i-1}$ by averaging over upper neighbors

$$D_i f(t) = \mathop{\mathbb{E}}_{s \supset t:\ |s|=i}[f(s)]$$

Visually, these have a nice representation on the inclusion diagram of X:

[7] We note Theorem 3.1 is overkill for Mihail-Vazirani and the 0-local-spectral case, which already followed from earlier work of Kaufman and Oppenheim [KO20]. The refined form above is critical, however, to many later breakthrough applications e.g. to independent sets, Ising models, etc.

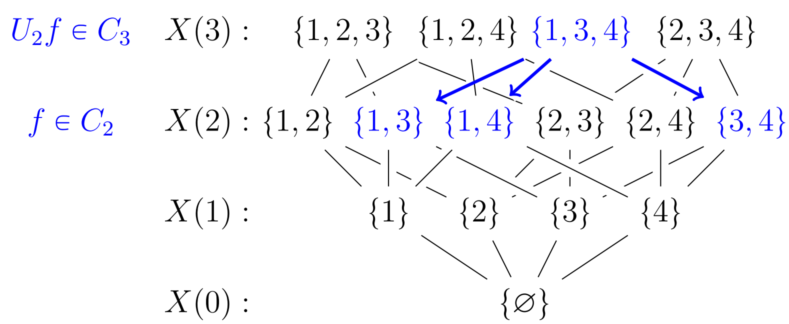

Figure 9: $U_2 f(\{1,3,4\})$ average f over lower neighbors $\{1,3\}$,$\{1,4\}$, and $\{2,4\}$

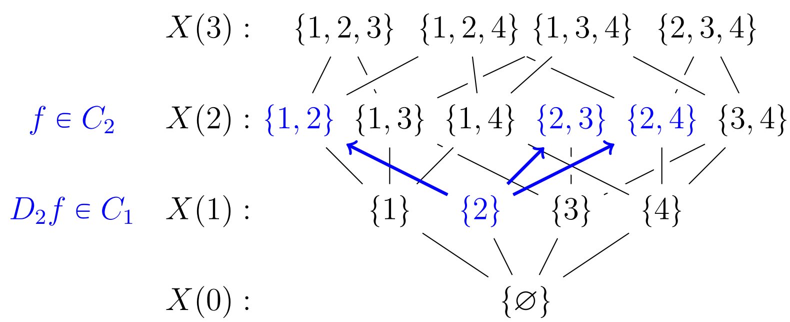

Figure 10: $D_2 f(\{2\})$ averages f over upper neighbors $\{1,2\}$, $\{2,3\}$, and $\{2,4\}$.

In particular, as is hopefully clear from the diagram, these operators are nothing more than the (normalized) *bipartite adjacency matrices* of the inclusion graph $(X(i), X(i-1))$.

Notice above we have not specified the distribution of our expectation very carefully. Let's now be formal about this, as we need to define them appropriately with respect to the complex weights $(\pi_0, \ldots, \pi_d)$. Because we are now moving between levels of the complex, the correct way to do this is to think of $\pi = (\pi_0, \ldots, \pi_d)$ as *jointly distributed* over chains of faces $\{\sigma_0 \subset \sigma_1 \subset \ldots \subset \sigma_d\}$ where we first draw $\sigma_d \sim \pi_d$, then iteratively draw each σ_i by removing a random vertex from σ_{i+1}. Note σ_i is then indeed distributed as π_i, since this process results in a uniformly random subset of σ_d of size i.

With this in mind, define the *up operator* $U_k : C_k \to C_{k+1}$ by averaging over neighbors below:

$$U_k f(s) := \mathop{\mathbb{E}}_{(\sigma_0,\ldots,\sigma_d)\sim\pi}[f(\sigma_k)|\sigma_{k+1} = s] = \frac{1}{k+1}\sum_{t\subset s} f(t)$$

and similarly the *down operator* by

$$D_{k+1} f(t) := \mathop{\mathbb{E}}_{(\sigma_0,\ldots,\sigma_d)\sim\pi}[f(\sigma_{k+1})|\sigma_k = t] = \sum_{s\supset t} \pi_1^t(s\backslash t) f(s)$$

Together, these operators make up the two steps of the up-down and down-up walks, denoted:

$$P_k^\vee = U_{k-1}D_k, \qquad P_k^\wedge = D_{k+1}U_k$$

We can also consider longer compositions of these operators that take multiple steps up and down the complex. Write $D_i^k := D_{i+1}\ldots D_k$ and $U_i^k : U_{k-1}\ldots U_i$ and define

$$P_{k\to i}^\vee := U_i^k D_i^k, \qquad P_{i\to k}^\wedge := D_i^k U_i^k$$

U_i^k and D_i^k are correspondingly the bipartite adjacency operators between levels $X(k)$ and $X(i)$, so it is easy to check this composition indeed gives the initially claimed walks, namely

Claim 3.4 (Action of the Upper and Lower Walks). *$P^{\vee}_{k\to i}$ acts on $s \in X(k)$ by*

1. *(Down-Step): Sampling $t \subset s$ of size i uniformly at random*
2. *(Up-Step): Sampling $s' \sim \pi^t_{k-i}$ from the link of t.*

Similarly $P^{\wedge}_{i\to k}$ acts on $t \in X(i)$ by

1. *(Up-Step): Sampling $s \sim \pi^t_{k-i}$ from the link of t*
2. *(Down-Step): Sampling $t' \subset s$ of size i uniformly at random*

Remark 3.5. Somewhat confusingly, as Markov operators U and D actually act in the opposite directions, i.e. U is the *down step* of the walk while D is the *up step* (this is because when thinking of performing the walk on a vector v, one *left*-multiplies to get the resulting position $w = vP$). As a result, you may see the definition of these operators flipped in the sampling literature.

The lower and upper walks are self-adjoint and therefore have a spectral decomposition. This follows from the fact that U_i and D_{i+1} are themselves adjoint since, much like our analysis of the standard weighted adjacency matrix, we can write:

$$\langle U_i f, g\rangle = \mathbb{E}_{(\sigma_0,\ldots,\sigma_d)\sim\pi}[f(\sigma_i)g(\sigma_{i+1})] = \langle f, D_{i+1} g\rangle.$$

We can now fully formally state the main result we'll prove this lecture, an extension of Alev and Lau's bound to longer walks:

Theorem 3.6 ([AR24]). *Let X be a $(\gamma_0, \ldots, \gamma_{d-2})$-one-sided HDX. For any $1 \leqslant i \leqslant d$:*

$$\lambda_2(P^{\vee}_{d\to i}) \leqslant 1 - \frac{d-i}{d}\prod_{j=0}^{i-1}(1-\gamma_j)$$

Here γ_j denotes the worst 1-skeleton expansion across j-links of X.

3.1 Variance Contraction and the Proof of Theorem 3.1

Recall our goal is to prove an upper bound on $\lambda_2(P^{\vee}_{k\to i})$, or by the variational characterization:

$$\lambda_2 = \max_{f\perp 1}\frac{\langle P^{\vee}_{k\to i} f, f\rangle}{\langle f, f\rangle}$$

A nice observation, originally due to [KM20] in this context, is that bounding this quantity is completely equivalent to understanding *variance contraction* of the down operator:

$$\lambda_2 = \max_{f\in C_k}\frac{\mathrm{Var}(D_i^k f)}{\mathrm{Var}(f)}$$

This is because (as we discussed) the down and up operators are adjoint, so letting $f^\perp = f - \mathbb{E}[f]\mathbf{1}$

$$\frac{\text{Var}(D_i^k f)}{\text{Var}(f)} = \frac{\langle D_i^k f^\perp, D_i^k f^\perp \rangle}{\langle f^\perp, f^\perp \rangle} = \frac{\langle P_{k\to i}^{\vee} f^\perp, f^\perp \rangle}{\langle f^\perp, f^\perp \rangle}$$

for any non-constant f. With this in mind, the proof will now follow from two main tools. The first is a variant of Garland's method in this setting, a simple local-to-global decomposition for Variance. To state it, we'll need a slightly different notion of localization for $f \in C_K(X)$ to links than before. In particular given $v \in X(1)$ define the *localization of f to X_v* by

$$f|_v(\sigma) = f(v \cup \sigma)$$

Note $f|_v \in C_{k-1}(X_v)$, which differs from our previous notion of a *restriction* $f_v \in C_k(X_v)$. We can now state our main (but totally elementary!) workhorse for this result: the Chain Rule for Variance.

Lemma 3.7 (Chain Rule for Variance)**.** *Let (X, π) be a pure d-dimensional simplicial complex. Then for any $k \leqslant d$ and $f \in C_k$*

$$\begin{aligned} Var_{\pi_k}(f) &= \mathop{\mathbb{E}}_{v\sim\pi_1}\left[Var_{\pi_{k-1}^v}(f|_v)\right] + \mathop{Var}_{v\sim\pi_1}(\mathbb{E}_{\pi_{k-1}^v}[f|_v]) \\ &= \mathop{\mathbb{E}}_{v\sim\pi_1}\left[Var_{\pi_{k-1}^v}(f|_v)\right] + \mathop{Var}_{\pi_1}(D_1^k f) \end{aligned}$$

Proof. We leave the proof as an exercise. □

Notice that this decomposition essentially breaks the variance of f into a local variance term inside the 1-links, and a global variance term sitting on the vertices $X(1)$. The second key lemma is to argue that this global term is actually controlled by the expansion of the 1-skeleton of X:

Lemma 3.8. *The variance contraction of D_1^d is exactly:*

$$\max_f \left\{ \frac{Var(D_1^d f)}{Var(f)} \right\} = \frac{1}{d} + \frac{d-1}{d}\lambda_2(A)$$

Now, hopefully, the overall proof strategy should be clear. We will bound Var$(D_i^d f)$ by a simple induction on i. In particular after decomposing by the chain rule, the lefthand term should become the local variance contraction within links and can be bounded inductively, while the latter term can be upper bounded by the expansion of X's 1-skeleton.

Let's formalize this. The base case of our induction is when $i = 1$. We need to show:

$$\frac{\text{Var}(D_1^d f)}{\text{Var}(f)} \overset{?}{\leqslant} 1 - \frac{d-1}{d}(1 - \lambda_2(A)) = \frac{1}{d} + \frac{d-1}{d}\lambda_2(A)$$

which is exactly Theorem 3.8.

Now assume the result holds for all complexes and $j \leqslant i$. Applying the Chain Rule:

$$\begin{aligned} \text{Var}_{\pi_i}(D_i^d f) &= \mathop{\mathbb{E}}_{v\sim\pi_1}\left[\text{Var}_{\pi_{d-1}^v}((D_i^d f)|_v)\right] + \mathop{\text{Var}}_{\pi_1}(D_1^i D_i^d f) \\ &= \mathop{\mathbb{E}}_{v\sim\pi_1}\left[\text{Var}_{\pi_{d-1}^v}(((D_v)_{i-1}^{d-1} f|_v)\right] + \mathop{\text{Var}}_{\pi_1}(D_1^d f) \end{aligned}$$

where we've used the fact that the down operator respects localization, i.e.

$$(D_i^d f)|_v(w) = \mathbb{E}_{\sigma\sim\pi_{d-i}^{w\cup v}}[f(\sigma\cup w\cup v)] = (D_v)_{i-1}^{d-1} f|_v(w)$$

where D_v is the local down operator on the link of v. For notational simplicity, denote

$$c = \frac{i}{d-1}\prod_{j=1}^{i-1}(1-\gamma_j)$$

Applying the inductive hypothesis:

$$\begin{aligned}
\mathrm{Var}(D_i^d f) &\leqslant (1-c)\mathop{\mathbb{E}}_{v\sim\pi_1}\left[\mathrm{Var}_{\pi_{d-1}^v}(f|_v)\right] + \mathop{\mathrm{Var}}_{v\sim\pi_1}(D_1^d f) && \text{(Ind Hyp)}\\
&= (1-c)(\mathrm{Var}_{\pi_d}(f) - \mathrm{Var}(D_1^d f)) + \mathop{\mathrm{Var}}_{v\sim\pi_1}(D_1^d f) && \text{(Chain Rule)}\\
&= (1-c)\mathrm{Var}_{\pi_d}(f) + c\mathop{\mathrm{Var}}_{v\sim\pi_1}(D_1^d f)\\
&= \left(1 - c\left(1 - \left(\frac{1}{d} + \frac{d-1}{d}\gamma_0\right)\right)\right)\mathrm{Var}_{\pi_d}(f) && \text{(Lemma 3.8)}
\end{aligned}$$

where we've now applied the chain rule for variance 'backwards' in the second step to lift our expected local variance back to a global bound. Finally we are left to analyze

$$\begin{aligned}
1 - \frac{i}{d-1}\prod_{j=1}^{i-1}(1-\gamma_j)\left(1 - \left(\frac{1}{d} + \frac{d-1}{d}\gamma_0\right)\right) &= 1 - \frac{i}{d}\prod_{j=1}^{i-1}(1-\gamma_j)\left(\frac{d}{d-1} - \left(\frac{1}{d-1} + \gamma_0\right)\right)\\
&= 1 - \frac{i}{d}\prod_{j=0}^{i-1}(1-\gamma_j)
\end{aligned}$$

as desired

Remark 3.9 (Beyond Variance). It turns out essentially the entire proof we have just given extends beyond variance to a more general class of measures called Φ-Entropies. Namely it shows how to infer global contraction of Φ-Entropy from local contraction of D_1^d on all the links. Taking, for instance, the classical Shannon entropy results in a lower bound on the MLSI constant of the down-up walk (which can lead to near-linear mixing time for the continuous time chain). In fact, a variant of this method was initially developed for entropy on the complete complex by Lee and Yau in the 90s [LY98].

Proof of Lemma 3.8 It is left to prove our key lemma relating variance contraction of D_1^k to expansion of the 1-skeleton (well, also to prove the Chain Rule, but you'll do this as an exercise).

Recall it is equivalent to bound $\lambda_2(P^{\vee}_{d\to 1})$. To prove this, we first claim we may instead analyze the upper walk.

Claim 3.10. *For any weighted complex* (X,π)*:* $\lambda_2(P^{\vee}_{d\to 1}) = \lambda_2(P^{\wedge}_{1\to d})$

This follows since $P^{\vee}_{d\to 1} = U_1^d D_1^d$ and $P^{\wedge}_{1\to d} = D_1^d U_1^d$. Indeed it is not hard to see these operators have the same non-zero spectrum.

Why have we moved to analyzing the upper walk? It turns out that any upper walk from the vertex level is, up to laziness, exactly the 1-skeleton of X. In particular:

Claim 3.11. $P^{\wedge}_{1\to d} = \frac{1}{d}I + \frac{d-1}{d}A_{\varnothing}$

This of course completes the result, since we then have $\lambda_2(P^{\wedge}_{1\to d}) = \lambda_2(P^{\wedge}_{1\to d}) = \frac{1}{d} + \frac{d-1}{d}\lambda_2(A_{\varnothing})$ as desired. To prove the claim, let's examine the probability of moving from $v \to w$ in $P^{\wedge}_{1\to d}$. Recall $P^{\wedge}_{1\to d}$ walks from $v \to w$ by:

1. Sampling $\sigma \sim \pi^v_{d-1}$
2. Sampling $w \in \sigma \cup v$ uniformly

Thus

$$P^{\wedge}_{1\to d}(v,w) = \frac{1}{d}\mathbf{Pr}_{\sigma\sim\pi^v_{d-1}}[w \in \sigma \cup v]$$

There are now two cases. First, if $w = v$, the righthand probability is always 1. Second, if $w \neq v$, then the righthand probability is just $(d-1)\pi^v_1(w)$, but this is just $(d-1)A(v,w)$! Thus taken together we have

$$P^{\wedge}_{1\to d}(v,w) = \frac{1}{d}I(v,w) + \frac{d-1}{d}A(v,w)$$

for all $v, w \in X(1)$, or in other words that $P^{\wedge}_{1\to d} = \frac{1}{d} + \frac{d-1}{d}A$ as desired.

4 Beyond Spectral Analysis

Our last lecture will be devoted to an overview of two core toolsets going beyond the basic spectral analysis covered in the previous section: concentration of measure and Fourier analysis. The former has become a central tool in the local-to-global method on HDX and is critical in the breakthrough line of work constructing new agreement tests and PCPs from high dimensional expanders. The latter has not yet found such success, but is a fundamental extension of a powerful tool in the TCS toolkit, and with close ties to related problems in hardness of approximation there is reasonable hope the ideas therein may yet be of use.

4.1 Sampling and Concentration of Measure

We'll first take a look at *concentration of measure*, perhaps the most ubiquitously used property of product spaces in the TCS literature and a driving force behind the power of HDX in application.

So, what exactly does it mean for a generic d-uniform hypergraph (simplicial complex) X to be concentrated? Perhaps the simplest interpretation is to look at how well X *samples* a subset $A \subset [n]$ of its vertices. In other words, over worst-case choice of A, what is the probability that the density of A inside a random hyperedge is off by more than ε from the mean?

$$\underset{(x_1,\ldots,x_d)\in X(d)}{\mathbf{Pr}}\left[\left|\frac{1}{d}\sum_{i=1}^{d} 1_A(x_i) - \mu_A\right| \geqslant \varepsilon\right] \overset{?}{\leqslant} \beta(\varepsilon,d).$$

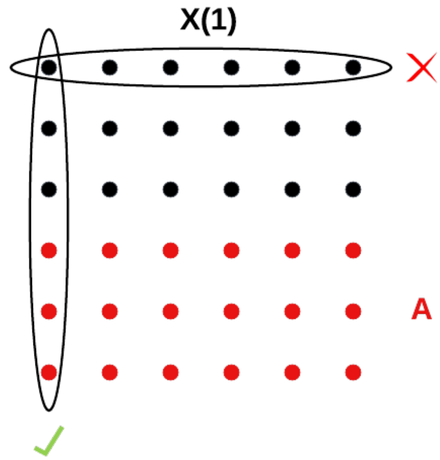

Figure 11: Hypergraph X and groundset A. The vertical hyperedge samples A in the right proportion, while the horizontal edge misses A completely.

When $X = \binom{[n]}{d}$, the reader might recognize this problem as the classical *Chernoff-Hoeffding* bound (without replacement), easily the most widely-used tail bound in TCS:

$$\mathbf{Pr}_{(x_1,\ldots,x_d)\in\binom{[n]}{d}}\left[\left|\frac{1}{d}\sum_{i=1}^{d} 1_A(x_i) - \mu_A\right| \geqslant \varepsilon\right] \lesssim e^{-\varepsilon^2 d}.$$

Thus, if HDX 'model' the complete complex, we might reasonably hope they give a 'derandomization' of Chernoff, a sparse family $X \subset \binom{[n]}{d}$ where the above concentration nevertheless holds.

In fact it's worth pausing to ask a more general question: what types of hypergraph (distribution) families are already known to have Chernoff-type tails? Unsurprisingly, this question is extremely well studied [SS89, SS90, SSS95, DR96, PS97, Pem01, BLM03, BLM13, IK10, Pau14, KS18, KKS21, AJK$^+$22]. Moreover, while most general conditions implying concentration require X to be dense, we have long known a simple and near-optimally sparse construction: simply take a random walk on an expander!

Theorem 4.1 (Expander-Chernoff [AKS87, Gil98])**.** *Let G be a λ-expander and $A \subset [n]$:*

$$\mathbf{Pr}_{(x_1,\ldots,x_d)\sim Walk^d(G)}\left[\left|\frac{1}{d}\sum_{i=1}^{d} 1_A(x_i) - \mu_A\right| \geqslant \varepsilon\right] \lesssim e^{-(1-\lambda)\varepsilon^2 d}.$$

If you haven't seen this result before, take a moment to internalize it. Say G is constant degree (even degree 3). After sampling x_1 (which is uniform on $[n]$), each following x_i is one of only three choices. In other words, these variables are *incredibly* correlated. Nevertheless, they near-optimally 'fool' functions of the vertices, acting as if each x_i is completely independent. This can be used, for instance, to greatly amplify the success probability of randomized algorithms with almost no additional randomness, and is a core tool is pseudorandomness, derandomization, and cryptography [Wig09, Gol11].

So if we've had essentially optimal constructions since the 90s, why has concentration become such a powerful tool in the context of HDX? The reason is that modern applications in complexity often require concentration against *broader classes of functions*. Above, we restricted our attention to sampling functions sitting on the *vertices* of X. In some sense, these correspond to *degree-1* or *linear* functions. Applications in pseudorandomness and PCPs often require strong concentration bounds against *degree-i* functions. In our setting, this corresponds to looking at functions that sit on 'i-sets' of X.[8] In particular, given a bounded function $f : X(i) \to [0,1]$, we'd like to bound the probability the conditional density of f in a d-face is far from its expectation:

$$\mathbf{Pr}_{s\in\binom{[n]}{d}}\left[\left|\mathop{\mathbb{E}}_{t\subset s}[f(t)] - \mathbb{E}[f]\right| \geqslant \varepsilon\right] \overset{?}{\leqslant} \beta(\varepsilon, i, d).$$

Concentration for degree-i functions was first considered explicitly by Impagliazzo, Kabanets, and Wigderson (IKW) [IKW12] for the complete complex $X = \binom{[n]}{d}$ as a core component in the construction of low soundness PCPs. For n sufficiently large, they proved the following generalization of Chernoff-Hoeffding to this setting

$$\mathbf{Pr}_{s\in\binom{[n]}{d}}\left[\left|\mathop{\mathbb{E}}_{t\subset s}[f(t)] - \mathbb{E}[f]\right| \geqslant \varepsilon\right] \lesssim e^{-\varepsilon^2\frac{d}{i}}. \tag{4}$$

Roughly speaking, the dependence $\frac{d}{i}$ corresponds to viewing a d-set in X as $\frac{d}{i}$ independent i-sets, and applying Chernoff to this system. We will not cover this in this lecture, but it turns out one can also show this bound is tight by reduction to tightness of Chernoff itself [DH24].

Unlike the degree-1 case, constructing derandomized systems with degree-i concentration remained largely[9] elusive for many years. Walks on expanders, for instance, fail drastically — even at $i = 2$ one can easily construct functions for any high girth graph such that walks fail to sample with better than constant probability.

This brings us back to high dimensional expanders: as sparse models of products and the complete complex $\binom{[n]}{d}$, HDX are a natural candidate for this problem. In fact, we've already proved that HDX are at least weakly concentrated in this sense! Why? The reader may have already noticed at this point that degree-i concentration can be rephrased in terms of a key local-to-global property we studied in the first lecture: *sampler graphs*. Recall a bipartite graph is a sampler if for any function f on the lefthand side, a random vertex on the right 'sees' the correct proportion of f with very high probability. Formally:

Definition 4.2 (Sampler graphs)**.** A bipartite graph $G = (L, R, E)$ is called an (α, β)-sampler graph if for any function $f : R \to [0,1]$:

$$\mathbf{Pr}_{v\in L}\left[\left|\mathop{\mathbb{E}}_{w\sim N(v)}[f(w)] - \mathbb{E}[f]\right| > \alpha\right] < \beta$$

In other words, our question about concentration of degree-i functions is *exactly* the question of whether the inclusion graph $(X(d), X(i))$ is a sampler. In a prior lecture we saw the following connection between expanders and samplers.

[8]Note our choice of the word 'degree' here is purposeful; we will see in the next subsection that there is a formal sense in which these functions are exactly the ith Fourier level of the space of functions on $X(d)$.

[9][IKW09] did give a partial derandomization by studying the *Grassmann* (the complex of subspaces). Without going into detail, this complex gives a polynomial size construction with a weak Chebyshev-type tail — a far cry from a *bounded degree* construction with a *Chernoff-type* tail.

Lemma 4.3. *Every λ-expander is an (α, β)-sampler graph for $\beta \leqslant O(\lambda^2/\alpha^2)$.*

But as we discussed in the previous lecture, the normalized adjacency matrix of the inclusion graph $(X(d), X(i))$ is just the down operator D_i^d. Moreover we proved

$$\lambda_2^2(D_i^d) = \lambda_2(U_i^d D_i^d) = \lambda_2(P^{\vee}_{d\to i}) \approx \frac{i}{d},$$

so we immediately get the following 'Chebyshev-type' tail for HDX:

Corollary 4.4 (HDX-Chebyshev)**.** *Let X be a one-sided $O(\frac{1}{d})$-local-spectral expander. Then for all $i \leqslant d$ and $f : X(i) \to [0,1]$*

$$\mathbf{Pr}_{s\sim X(d)}\left[\left|\mathop{\mathbb{E}}_{t\subset s}[f(t)] - \mathbb{E}[f]\right| \geqslant \alpha\right] \leqslant O\left(\frac{i}{d\alpha^2}\right).$$

This is already quite a useful result (indeed it is the main component, e.g., of the original work connecting HDX and agreement testing [DK17] and gives the first construction of bounded-degree hypergraphs with some sort of degree-i concentration), but it's a far cry from the *subgaussian* concentration exhibited by the complete complex. Recently, Dikstein and Hopkins [DH24] made progress in this direction, showing *exponential* concentration for weak HDX and subgaussian concentration matching the complete complex under stronger assumptions.

Theorem 4.5 (HDX-Chernoff [DH24])**.** *Let X be a d-dimensional one-sided $\frac{c}{d-1}$-local-spectral HDX for any constant $c < 1$. Then for all $i \leqslant d$ and functions $f : X(i) \to [0,1]$:*

$$\mathbf{Pr}_{s\sim X(d)}\left[\left|\mathop{\mathbb{E}}_{t\subset s}[f(t)] - \mathbb{E}[f]\right| \geqslant \varepsilon\right] \leqslant e^{-\Omega\left(\sqrt{\alpha^2\frac{d}{i}}\right)}.$$

If X is a two-sided (or one-sided partite) $2^{-O(d)}$-HDX, this can be improved to match Δ_n:

$$\mathbf{Pr}_{s\sim X(d)}\left[\left|\mathop{\mathbb{E}}_{t\subset s}[f(t)] - \mathbb{E}[f]\right| \geqslant \varepsilon\right] \leqslant e^{\left(-\Omega\left(\alpha^2\frac{d}{i}\right)\right)}$$

Unfortunately the proof of this Theorem is outside the scope of these lectures (though it is not too involved!). We refer the interested reader to [DH24] and a corresponding lecture on the topic: `https://www.youtube.com/watch?v=3KsyVea6_Fo`. It remains an interesting question whether the former of these bounds is tight — do HDX with expansion bounded away form $\frac{1}{d-1}$ satisfy a full Chernoff bound, even for $i = 1$?

The 'Tricking-Down' Threshold: Like the Trickling-Down Theorem, Theorem 4.5 also exhibits a *phase transition* in its behavior at expansion $\frac{1}{d-1}$. In particular, as long as the co-dimension 2 links have expansion slightly bounded away from $\frac{1}{d}$, say even $\frac{1}{d-1} - \frac{1}{\text{poly(d)}}$, it is possible to achieve concentration $\exp(-\text{poly}(d))$. On the other hand, at expansion exactly $\frac{1}{d-1}$ one can construct complexes with arbitrarily poor concentration. This is a fairly striking behavior, and in part explains the great gap in difficulty in constructing HDX with expansion better than $\frac{1}{d-1}$, where only algebraic constructions are known.

4.1.1 Agreement Testing

Before we move on to our final topic, let's try to get a feel for why degree-i concentration and sampling on HDX are useful beyond what one can do with expander graphs. As mentioned above, the prototypical application of this result is to *agreement testing*.

Given a d-dimensional simplicial complex X, an agreement test over X takes as input a family of *local* functions on each d-face $\mathcal{F} := \{f_\sigma : \sigma \to \{0,1\}\}_{\sigma \in X(d)}$, and (via query access to $\mathcal{F}$) aims to check whether $\{f_\sigma\}$ are really restrictions of a single *global* function on the vertices, that is if there exists some $g : X(1) \to \{0,1\}$ such that for all $\sigma \in X(k)$ and $x \in \sigma$:

$$f_\sigma(x) = g(x).$$

Agreement tests form the combinatorial core of essentially all PCP constructions. Roughly speaking, one should think of starting with an original 'weakly hard' problem on the vertices and trying to amplify it to a harder 'd-wise' variant over the faces of $X(d)$. The agreement test ensures that the solutions to the d-wise version actually come from a real solution to the original problem. Since we've talked a bit about local testability, we note this can also be thought of as a local tester for the 'direct product code'

$$C := \left\{ f \in (\mathbb{F}_2^d)^{|X(d)|} : \exists g \in \mathbb{F}_2^{|X(1)|} \text{ s.t. } f(s) = g|_s \right\},$$

in other words, the codewords of C come from concatenating the symbols of any binary string on the vertices $X(1)$ across the d-faces $X(d)$.

Formally, an agreement test is a procedure that reads $\mathcal{F}$ in a few random places and accepts or rejects based on the local agreement behavior. We'd like our test to be *complete*, meaning if $\mathcal{F}$ truly is global the test will always pass, and *sound* meaning that if the test passes with high probability, then $\mathcal{F}$ is truly close to a global function.

We will focus on the 2-query setting, where there is really only one natural test: query two overlapping s, s' and check whether they agree on the intersection:

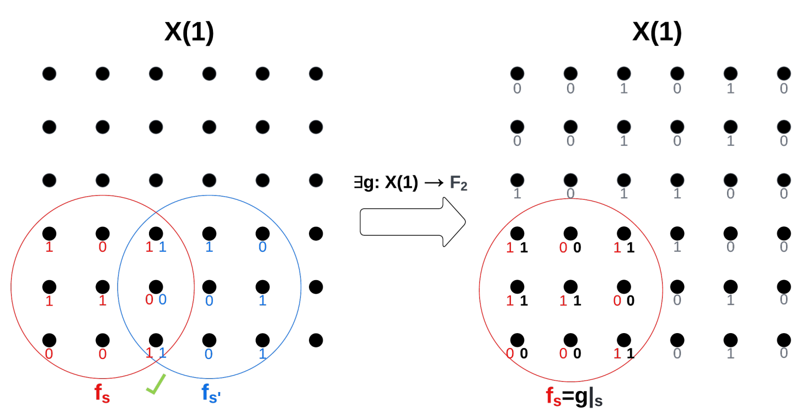

Figure 12: If f_s and f_s' often agree on their intersection, we'd like to infer f agrees with a global function g on the vertices.

In our case we will generate our intersecting sets through the $d \to \sqrt{d} \to d$ down-up walk, now indexed by triples (s,t,s') generated as:

1. Sample $s \sim \pi_d$
2. Sub-sample $t \subset s$ of size $\sqrt{d}$
3. Sample $s' \supset t$ from the link

Instead of checking exact agreement, we'll actually work in a slightly different regime which tests *approximate* closeness to a global function, namely we'll test

$$Agr_\eta(\mathcal{F}) := \mathbf{Pr}_{(s,t,s')}[f_s(t) \overset{\eta}{\approx} f_{s'}(t)]$$

where the inner notation denotes f_s and $f_{s'}$ agree outside an η fraction of t.[10]

Note this test is clearly complete. If $\mathcal{F}$ is truly global (or even close in the sense that all $f_s \approx g|_s$), the test will pass w.h.p. We'll show the test is also sound:

Theorem 4.6 (Approximate Agreement Testing (99%-Regime))**.** *Let X be a $\frac{1}{2d}$-one-sided local-spectral expander for any $c < 1$. For any $\varepsilon > \exp(-O_\eta(d^{1/4}))$:*

$$Agr_\eta(\mathcal{F}) \geqslant 1 - \varepsilon \implies \mathop{\mathbf{Pr}}_{s \in X(d)}\left[f_s \overset{O(\varepsilon+\eta)}{\approx} g_{maj}(s)\right] \leqslant 1 - O(\varepsilon)$$

where $g_{maj}(v) := maj_{s \ni v}\{f_s(v)\}$ is the majority vote over d-faces containing v.

We remark that when $\eta = 0$ (the exact agreement setting), it is actually possible to remove the constraints on ε and achieve *exact* agreement in the soundness [DK17]. This also relies strongly on sampling, but the proof is more involved than we have time to cover. The approximate setting we cover is also important in application, e.g. versions of this are used critically in hardness amplification (both for average-case hardness amplification [IJKW08] and in recent HDX-based PCP constructions [BMV24]).

Let's move now to the proof, focusing mainly on the role of sampling and concentration. To start, we'll take the following 'average-case' agreement guarantee for granted:

Claim 4.7. *$\mathcal{F}$ 'typically' agrees with g_{maj} over a random (vertex, d-set) pair:*

$$\mathbf{Pr}_{v \in s}[f_s(v) \neq g_{maj}(v)] \leqslant 2(\varepsilon + \eta).$$

This follows from expansion of the lower walk conditioned on containing a given vertex v, and is a good exercise in working with expander graphs.

With this in hand, consider the following sets of 'good' and 'bad' d-faces:

$$A := \{s \in X(d) : \mathbf{Pr}_{v \in s}[f_s(v) \neq g_{\text{maj}}(v)] \leqslant 4(\varepsilon + \eta)\}$$
$$B := \{s \in X(d) : \mathbf{Pr}_{v \in s}[f_s(v) \neq g_{\text{maj}}(v)] \geqslant 8(\varepsilon + \eta)\}.$$

[10] Note formally the intersection could be more than just t (though it is quite unlikely on a strong HDX). Restricting to t just makes the proof a bit simpler.

Note that by Claim 4.7 and Markov's inequality, the fraction of good faces is at least $\mathbf{Pr}[A] > 1/2$. Our goal is to show this implies the fraction of bad faces is quite small: $\mathbf{Pr}[B] \leqslant O(\varepsilon)$.

The idea behind the proof is simple, though a bit technical to execute. We know our test passes except with probability ε. If B is too large, we will use sampling to argue that many pairs (s, s') drawn from our test pass between A and B (say a 5ε fraction). However, A and B have very different agreement with g_{maj}, which means s and s' will typically disagree on a $\gg \eta$ fraction on their intersection. But then the test will reject with probability $\gg \varepsilon$, violating our initial assumption on $Agr_\eta(\mathcal{F})$.

Let's do this a bit more formally. Assume towards contradiction that B is large, say size $10\varepsilon \gg e^{-O_\eta(\sqrt{d})}$, and note that conditioned on our test pair (s, s') passing between sets A and B, our test should almost always reject. Why? Remember that (marginally) t is a uniformly random subset of s (likewise a uniformly random subset of s'). By standard Chernoff, a random $t \subset s \in A$ will see no more than a $5(\varepsilon + \eta)$ fraction of vertices that disagree with $g_{\text{maj}}(v)$, while a random $t \subset s' \in B$ will see at least $7(\varepsilon+\eta)$ such disagreements with extremely high probability. As long as these very high probability events occur, our test rejects since f_s and $f_{s'}$ cannot then be equivalent on t.[11]

With this in mind, we now argue the probability of passing between A and B is large

$$\mathbf{Pr}_{(s,t,s')}\left[s \in A, s' \in B\right] \geqslant 2\varepsilon \tag{5}$$

and therefore that the test rejects with probability greater than ε, a contradiction.

We claim Equation (5) follows from the sampling properties of $(X(d), X(\sqrt{d}))$, though as we will see we'll need to make a slight adjustment.

To start, notice that when sampling our test triple (s, t, s') from the down-up walk, while we previously thought of the procedure as sampling s, then $t \subset s$, and finally $s' \supset t$, we can instead *first* sample t, then $s, s' \supset t$ independently.[12] This means we can re-write the probability of passing between A and B as:

$$\mathbb{E}_{t\sim\pi_{\sqrt{d}}}\left[\mathbf{Pr}[s \in A | s \supset t] \cdot \mathbf{Pr}[s' \in B | s' \supset t]\right]$$

Now hopefully the relation to sampling becomes clear! We'd be done if we can argue that for, say, 99% of $t \sim \pi_{\sqrt{d}}$, the local density of A and B around t is at least $3/4$ what it should be, as then:

$$\mathbb{E}_{t\sim\pi_{\sqrt{d}}}\left[\mathbf{Pr}[s \in A | s \supset t] \cdot \mathbf{Pr}[s' \in B | s' \supset t]\right] \geqslant 0.99 \cdot \frac{3}{4}\mathbf{Pr}[A] \cdot \frac{3}{4}\mathbf{Pr}[B] \geqslant 2\varepsilon$$

This is almost exactly what sampling says, except it's in the wrong direction. The guarantee we currently have is that says for a set $S \subset X(\sqrt{d})$:

$$\mathbf{Pr}_{s\sim X(d)}[|\mathbf{Pr}[A|s] - \mathbf{Pr}[A]| > 0.01] \leqslant \exp(-\Omega(d^{1/4}))$$

[11] Technically we are cheating here, and one needs to ensure the failure probability of the Chernoff bound is much smaller than the probability of passing between A and B. This will indeed be the case in the rest of the argument, so this is not a real issue.

[12] Why? Note that once we fix t, it does not matter in the down-up walk which s we came from, the choice of s' is completely independent.

What we want is in some sense the 'flipped' guarantee, that for any subset $S \subset X(d)$:

$$\mathbf{Pr}_{t \sim X(\sqrt{d})}[|\mathbf{Pr}[A|s] - \mathbf{Pr}[A]| > \exp(-\Omega(d^{1/4}))] \leqslant 0.01$$

This is just the claim that $(X(\sqrt{d}), X(d))$ is an $(\exp(-\Omega(d^{1/4})), 0.01)$-sampler. Thankfully, it turns out that these 'flipped' views are essentially equivalent:

Claim 4.8. *If $G = (L, R, E)$ is an (α, β)-sampler, then $G^{rev} = (R, L, E)$ is an $((O(\beta), O(\alpha))$-sampler*

Proof. Exercise! □

As a corollary, we immediately get that $(X(\sqrt{d}), X(d))$ is a $(0.01, \varepsilon)$-sampler graph which completes the proof.

4.2 A Foray into Fourier

In Section 3 we showed how to leverage the approximate product structure of high dimensional expanders to prove the spectral gap of the down-up walk nearly matches its classical value on the cube. This raises a natural question: can we prove that HDX nearly match *all* the eigenvalues of product spaces? What can we say about the structure of their corresponding eigenspaces?

Let's take a step back. On the hypercube and product spaces, these questions are classically answered by the beautiful theory of *Fourier analysis*. In the case of the cube, for instance, the eigenvectors of the down-up walk (indeed any Cayley graph on $\{\pm 1\}^d$) are just the Fourier characters $\chi_S = \prod_{i \in S} x_i$

$$P_d^{\vee} \chi_S = \frac{d - |S|}{d} \chi_S,$$

so a function f's decomposition onto the eigenspaces of $P_d^{\vee}$ is exactly its Fourier decomposition:

$$f = \sum_{S \subset [d]} \widehat{f}(S) \chi_S$$

Moving to product spaces, there is a classical analog of the Fourier decomposition which breaks f into components coming from its projection onto each of the coordinate subsets $S \subset [d]$:

$$f = \sum_{S \subseteq [n]} f^{=S},$$

where on the cube $f^{=S}$ is $\widehat{f}(S)\chi_S$. In this setting, however, we will take a *combinatorial* approach to defining each contribution $f^{=S}$. Towards this end, it will be useful to define a set of *partite averaging operators*. Namely given any subset $S \subseteq [n]$, define E_S to be the operator that averages a function f over its values on S:

$$E_S[f](x) := \mathbb{E}[f(z) \mid z_S = x_S].$$

Equivalently, E_S can be thought of as re-randomizing f over all coordinates outside of S, or as a restricted version of the down-up walk (where the down-process is forced to walk to coordinate set S). In fact, this can be made formal by the following observation:

$$P^{\vee}_{d\to i} = \frac{1}{\binom{d}{i}} \sum_{|S|=i} E_S.$$

Back to the task at hand, what is the contribution to f coming from S? One natural idea is just to use $E_S f$. This is almost right, but the expression inherently counts contributions coming from all subsets of S as well. Using inclusion-exclusion to subtract out these contributions suggests the following formula:

$$f^{=S} = \sum_{T\subseteq S} (-1)^{|S\setminus T|} E_T f$$

We now have the following generalization of the classical Fourier Decomposition:

Theorem 4.9 (Efron-Stein Decomposition)**.** *For any d-dimensional product space X and function $f : X \to \mathbb{R}$, the Efron-Stein decomposition satisfies:*

1. ***Decomposition**: $f = \sum_{S\subset[d]} f^{=S}$*
2. ***Eigenbasis**: $P^{\vee}_{d} f^{=S} = \frac{d-|S|}{|S|} f^{=S}$*
3. ***Orthogonality**: $\forall S \neq S' : \langle f^{=S}, f^{=S'} \rangle = 0$*

The proof of these results relies on the following simple but key observation:

Claim 4.10. *For any d-dimensional product space X and any subsets $S, T \subseteq [n]$:*

$$E_S E_T = E_{S\cap T}$$

Proof. Recall E_T re-randomizes a function over $[n]\setminus T$. Applying E_S and E_T, we are re-randomizing over all coordinates except $S \cap T$. Since X is a product these operations are independent, so this is exactly $E_{S\cap T}$. □

In fact, this is the *only* assumption we will make on X (besides being a partite complex), which hints at what's to come!

Proof of Theorem 4.9.

Decomposition: We expand $\sum f^{=S}$

$$\begin{aligned}
\sum_{S\subset[d]} f^{=S} &= \sum_{S\subset[d]} \sum_{T\subset S} (-1)^{|S\setminus T|} E_T f \\
&= \sum_{T\subset[d]} \left(\sum_{j=0}^{d-|T|} (-1)^j \binom{[d]-|T|}{j} \right) E_T f \\
&= E_{[d]} f \\
&= f
\end{aligned}$$

since the sum of alternating binomial coefficients is 0 by the Binomial theorem (expand $(1-1)^d$).

Eiegenbasis: We can write $P_d^\vee$ in terms of the averaging operators as:

$$P_d^\vee = \frac{1}{d}\sum_{i\in[d]} E_{[d]\setminus i}$$

Then we can write:

$$P_d^\vee f^{=S} = \frac{1}{d}\sum_{i\in[d]} E_{[d]\setminus i} f^{=S}$$

Let's examine each $E_{[d]\setminus i} f^{=S}$ separately. There are two cases of interest based on whether $i \in S$. In particular, we claim

$$E_{[d]\setminus i} f^{=S} = \begin{cases} f^{=S} & \text{if } i \notin S \\ 0 & \text{if } i \in S \end{cases}$$

Once we have shown this we are done, since there are exactly $d - |S|$ coordinates s.t. $i \notin S$:

$$\begin{aligned} P_d^\vee f^{=S} &= \frac{1}{d}\sum_{i\in[d]} E_{[d]\setminus i} f^{=S} \\ &= \frac{1}{d}\sum_{i\notin S} f^{=S} \\ &= \frac{d-|S|}{d} f^{=S} \end{aligned}$$

Let's now analyze $E_{[d]\setminus i} f^{=S}$. We claim the first case, when $i \notin S$, is obvious. Namely $E_{[d]\setminus i}$ acts by re-randomizing the ith coordinate, but $f^{=S}$ doesn't depend on the ith coordinate so we immediately see $E_{[d]\setminus i} f^{=S} = f^{=S}$.

Now let's look at the case when $i \in S$. The intuition is that once we randomize over i, the function is now only depends on the set $S\setminus i$, so it should have no projection onto $f^{=S}$. Formally, expanding out $f^{=S}$

$$\begin{aligned} E_{[d]\setminus i} f^{=S} &= E_{[d]\setminus i}\left(\sum_{T'\subseteq S} (-1)^{|S\setminus T'|} E_{T'}[f]\right) \\ &= \left(\sum_{T'\subseteq S} (-1)^{|S\setminus T'|} E_{[d]\setminus i} E_{T'}[f]\right) && \text{(Linearity of } E_T\text{)} \\ &= \left(\sum_{T'\subseteq S} (-1)^{|S\setminus T'|} E_{\{T'\setminus\{i\}\}}[f]\right) && \text{(Claim 4.10)} \\ &= \left(\sum_{T\subseteq S\setminus i} ((-1)^{|S\setminus T|} + (-1)^{|S\setminus\{T\cup i\}|}) E_{\{T\}}[f]\right) \\ &= 0 \end{aligned}$$

Orthogonality: We now wish to show that for any $S \neq S'$:

$$\langle f^{=S}, f^{=S'} \rangle = 0$$

This follows from the self-adjointness of the operator E_T. Assume without loss of generality that there exists some $i \in S \backslash S'$, then we can write:

$$\begin{aligned}\langle f^{=S}, f^{=S'} \rangle &= \sum_{T \subseteq S \backslash \{i\}} (-1)^{|S \backslash T|} \langle (E_T - E_{T \cup \{i\}}) f, f^{=S'} \rangle \\ &= \sum_{T \subseteq S \backslash \{i\}} (-1)^{|S \backslash T|} \langle f, (E_T - E_{T \cup \{i\}}) f^{=S'} \rangle.\end{aligned}$$

But since $i \notin S'$ by definition

$$\begin{aligned}(E_T - E_{T \cup \{i\}}) f^{=S'} &= \sum_{T' \subseteq S'} (-1)^{|S \backslash T'|} (E_T - E_{T \cup \{i\}}) E_{T'} f \\ &= \sum_{T' \subseteq S'} (-1)^{|S \backslash T'|} (E_{T \cap T'} - E_{(T \cup \{i\}) \cap T'}) f && \text{(Claim 4.10)} \\ &= 0 && (i \notin T')\end{aligned}$$

□

We'd now like to extend this decomposition to (sufficiently strong) partite HDX. Because we only used the fact that $E_S E_T = E_{S \cap T}$, it is clearly sufficient (combined with the triangle inequality) to prove the following approximate variant for HDX:

Claim 4.11 (Lemma 3.3 [GLL21]). *For any d-dimensional partite γ-HDX X and any subsets $S, T \subseteq [n]$:*

$$\|E_S E_T - E_{S \cap T}\|_2 \leqslant O(|S||T|\gamma)$$

where $||A||_2$ is the operator norm $\max_f \frac{||Af||}{||f||}$

As an immediate corollary, we get a version of Theorem 4.9 for HDX.

Corollary 4.12 (Efron-Stein on HDX [GLL21, Hop24]). *For any d-dimensional partite γ-HDX X, Efron-Stein is an approximately orthogonal approximate eigenbasis:*

1. ***Decomposition**: $f = \sum_{S \subset [d]} f^{=S}$*
2. ***Approximate Eigenbasis**: $||P_d^{\vee} f^{=S} - \frac{d-|S|}{d} f^{=S}||_2 \leqslant 2^{O(d)} \gamma ||f||$*
3. ***Approximate Orthogonality**: $\forall S \neq S' : \langle f^{=S}, f^{=S'} \rangle \leqslant 2^{O(d)} \gamma ||f||_2$*

Proof. Repeat exactly the same arguments as in Theorem 4.9, but replace every use of Claim 4.10 with Claim 4.11. □

Note that since there are exp(d) many terms, Theorem 4.12 is only useful when $\gamma \leqslant 2^{-\Omega(d)}$, which is generally very far from what we required in previous sections. It would be quite nice to give a decomposition that works when $\gamma \approx 1/\text{poly}(d)$, or even near the Trickling-Down Threshold $1/d$.

Remark 4.13 (Beyond Spectral Analysis). Much of the power of Fourier analysis comes from developing a fine-grained understanding the structure of the Fourier components. For instance, the powerful theory of *hypercontractivity* shows that the low-degree Fourier components of any f are *well spread* in some formal sense. It turns out one can even generalize this broader theory to high dimensional expanders, but this is outside the scope of these lectures. See `https://www.wisdom.weizmann.ac.il/~dinuri/courses/22-HDX/L13.pdf` for some lecture notes on this topic.

References

[AJK+21] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence in high-dimensional expanders: Modified log-sobolev inequalities for fractionally log-concave polynomials and the ising model. *arXiv preprint arXiv:2106.04105*, 2021.

[AJK+22] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence: optimal mixing of down-up random walks. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1418–1430, 2022.

[AKS87] Miklós Ajtai, János Komlós, and Endre Szemerédi. Deterministic simulation in logspace. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 132–140, 1987.

[AL20] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. *arXiv preprint arXiv:2001.02827*, 2020.

[ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *arXiv preprint arXiv:2001.00303*, 2020.

[ALOV19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials ii: high-dimensional walks and an FPRAS for counting bases of a matroid. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 1–12, 2019.

[AR24] Vedat Levi Alev and Shravas Rao. Expanderizing higher order random walks. *arXiv preprint arXiv:2405.08927*, 2024.

[BLM03] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. Concentration inequalities using the entropy method. *The Annals of Probability*, 31(3):1583–1614, 2003.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 02 2013.

[BLM24] Mitali Bafna, Noam Lifshitz, and Dor Minzer. Constant degree direct product testers with small soundness. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 862–869. IEEE, 2024.

[BM24] Mitali Bafna and Dor Minzer. Characterizing direct product testing via coboundary expansion. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1978–1989, 2024.

[BMV24] Mitali Bafna, Dor Minzer, and Nikhil Vyas. Quasi-linear size pcps with small soundness from hdx. *arXiv preprint arXiv:2407.12762*, 2024.

[CSŻ03] Donald I Cartwright, Patrick Solé, and Andrzej Żuk. Ramanujan geometries of type an. *Discrete mathematics*, 269(1-3):35–43, 2003.

[DD23] Yotam Dikstein and Irit Dinur. Agreement theorems for high dimensional expanders in the small soundness regime: the role of covers. *arXiv preprint arXiv:2308.09582*, 2023.

[DDL24] Yotam Dikstein, Irit Dinur, and Alexander Lubotzky. Low acceptance agreement tests via bounded-degree symplectic hdxs. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 826–861. IEEE, 2024.

[DH24] Yotam Dikstein and Max Hopkins. Chernoff bounds and reverse hypercontractivity on hdx. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 870–919. IEEE, 2024.

[DK17] Irit Dinur and Tali Kaufman. High dimensional expanders imply agreement expanders. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 974–985. IEEE, 2017.

[DR96] Devdatt P Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *BRICS Report Series*, 3(25), 1996.

[EK16] Shai Evra and Tali Kaufman. Bounded degree cosystolic expanders of every dimension. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 36–48, 2016.

[Gar73] Howard Garland. p-adic curvature and the cohomology of discrete subgroups of p-adic groups. *Annals of Mathematics*, pages 375–423, 1973.

[Gil98] David Gillman. A chernoff bound for random walks on expander graphs. *SIAM Journal on Computing*, 27(4):1203–1220, 1998.

[GK22] Roy Gotlib and Tali Kaufman. List agreement expansion from coboundary expansion. *arXiv preprint arXiv:2210.15714*, 2022.

[GLL21] Tom Gur, Noam Lifshitz, and Siqi Liu. Hypercontractivity on high dimensional expanders. *arXiv preprint arXiv:2111.09375*, 2021.

[Gol11] Oded Goldreich. A sample of samplers: A computational perspective on sampling. In *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation: In Collaboration with Lidor Avigad, Mihir Bellare, Zvika Brakerski, Shafi Goldwasser, Shai Halevi, Tali Kaufman, Leonid Levin, Noam*

Nisan, Dana Ron, Madhu Sudan, Luca Trevisan, Salil Vadhan, Avi Wigderson, David Zuckerman, pages 302–332. Springer, 2011.

[Gol21] Louis Golowich. Improved product-based high-dimensional expanders. *arXiv preprint arXiv:2105.09358*, 2021.

[Gro10] M. Gromov. Singularities, expanders and topology of maps. part 2: from combinatorics to topology via algebraic isoperimetry. *Geom. Funct. Anal.*, 20:416526, 2010.

[Hop24] Max Hopkins. Hypercontractivity on hdx ii: Symmetrization and q-norms. *arXiv preprint arXiv:2408.16687*, 2024.

[HS19] Prahladh Harsha and Ramprasad Saptharishi. A note on the elementary construction of high-dimensional expanders of kaufman and oppenheim. *arXiv preprint arXiv:1912.11225*, 2019.

[IJKW08] Russell Impagliazzo, Ragesh Jaiswal, Valentine Kabanets, and Avi Wigderson. Uniform direct product theorems: simplified, optimized, and derandomized. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, pages 579–588, 2008.

[IK10] Russell Impagliazzo and Valentine Kabanets. Constructive proofs of concentration bounds. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 617–631. Springer, 2010.

[IKW09] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query pcps. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 131–140, 2009.

[IKW12] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query pcps. *SIAM Journal on Computing*, 41(6):1722–1768, 2012.

[KKL16] Tali Kaufman, David Kazhdan, and Alexander Lubotzky. Isoperimetric inequalities for ramanujan complexes and topological expanders. *Geometric and Functional Analysis*, 26(1):250–287, 2016.

[KKS21] Tali Kaufman, Rasmus Kyng, and Federico Soldá. Scalar and matrix chernoff bounds from ℓ_∞-independence. *arXiv preprint arXiv:2111.02177*, 2021.

[KM16] Tali Kaufman and David Mass. High dimensional combinatorial random walks and colorful expansion. *arXiv preprint arXiv:1604.02947*, 2016.

[KM20] Tali Kaufman and David Mass. Local-to-global agreement expansion via the variance method. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

[KO18] Tali Kaufman and Izhar Oppenheim. Construction of new local spectral high dimensional expanders. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 773–786, 2018.

[KO20] Tali Kaufman and Izhar Oppenheim. High order random walks: Beyond spectral gap. *Combinatorica*, pages 1–37, 2020.

[KS18] Rasmus Kyng and Zhao Song. A matrix chernoff bound for strongly rayleigh distributions and spectral sparsifiers from a few random spanning trees. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 373–384. IEEE, 2018.

[Li04] Wen-Ching Winnie Li. Ramanujan hypergraphs. *Geometric & Functional Analysis GAFA*, 14:380–399, 2004.

[LSV05] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. Explicit constructions of ramanujan complexes of type Ad. *European Journal of Combinatorics*, 26(6):965–993, 2005.

[LY98] Tzong-Yow Lee and Horng-Tzer Yau. Logarithmic sobolev inequality for some models of random walks. *The Annals of Probability*, 26(4):1855–1873, 1998.

[MV89] Milena Mihail and Umesh Vazirani. On the expansion of 0/1 polytopes. *Journal of Combinatorial Theory. B*, 1989.

[MW09] Roy Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Struct. Algorithms*, 34(3):408–417, 2009.

[Opp18] Izhar Oppenheim. Local spectral expansion approach to high dimensional expanders part i: Descent of spectral gaps. *Discrete & Computational Geometry*, 59(2):293–330, 2018.

[Pau14] Daniel Paulin. The convex distance inequality for dependent random variables, with applications to the stochastic travelling salesman and other problems. *Electronic Journal of Probability*, 19:1–34, 2014.

[Pem01] Sriram V Pemmaraju. Equitable coloring extends chernoff-hoeffding bounds. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 285–296. Springer, 2001.

[PS97] Alessandro Panconesi and Aravind Srinivasan. Randomized distributed edge coloring via an extension of the chernoff–hoeffding bounds. *SIAM Journal on Computing*, 26(2):350–368, 1997.

[Sar04] Alireza Sarveniazi. *Ramanujan Hypergraph Based on Bruhat–Tits Building*. PhD thesis, PhD thesis, University of Göttingen, 2004. 87, 88, 2004.

[SS89] Jeanette P Schmidt and Alan Siegel. On aspects of university and performance for closed hashing. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 355–366, 1989.

[SS90] Jeanette P Schmidt and Alan Siegel. The analysis of closed hashing under limited randomness. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, pages 224–234, 1990.

[SSS95] Jeanette P Schmidt, Alan Siegel, and Aravind Srinivasan. Chernoff–hoeffding bounds for applications with limited independence. *SIAM Journal on Discrete Mathematics*, 8(2):223–250, 1995.

[Wig09] Avi Wigderson. Randomness extractors–applications and constructions. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (2009)*, pages 471–473. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2009.

A Section 1-2 Exercises

Exercise A.1 (Links and Basic Conditioning)**.** *Let (X, π) be a weighted pure simplicial complex. Consider the following procedure sampling a vertex v and i-face σ:*

1. *Sample $v \sim \pi_1$*
2. *Sample $\sigma \sim \pi_i^v$ from the link of v*

Prove that

1. *σ is marginally distributed as π_i.*
2. *$\{v\} \cup \sigma$ is distributed as π_{i+1}*

Exercise A.2. *Using the above:*

1. *Prove the weighted adjacency matrix is self-adjoint*
2. *Finish the proof of Garland's method.*

Exercise A.3 (Trickle Down)**.** *Prove the full Trickling Down Theorem (Corollary 2.2).*

Exercise A.4 (The Flags Complex)**.** *Consider the graph $G = (V, E)$ whose vertices V are given by all subspaces of $\mathbb{F}_q^{d+1}$ and whose (undirected) edges are given by subspace inclusion. Prove that the d-dimensional* ***clique complex*** *of G, that is the simplicial complex whose top level faces are all d-cliques in G, is an $\Omega(1/\sqrt{q})$-one-sided HDX so long as $q \geqslant \Omega(d^2)$*

Exercise A.5 (Spanning Tree Complex)**.** *Prove $X_G(n-1)$ is a 0-local-spectral HDX.*

Exercise A.6 (γ-products vs local-spectral HDX)**.** *Prove that γ-products and partite local-spectral expanders are nearly equivalent:*

1. *Any d-dimensional γ-product is a $\frac{\gamma}{1-(d-2)\gamma}$-one-sided local spectral HDX*
2. *Any d-partite γ-local-spectral HDX is a $\frac{\gamma}{1-(d-2)\gamma}$-product*

Exercise A.7 (High Dimensional Expander-Mixing Lemma)**.** *Let d be a power of 2. Prove that given a d-partite γ-product X and functions $f_i : X[i] \to \mathbb{R}$,*

$$\left| \mathbb{E}_x\left[\prod_{i\in[d]} f_i\right] - \prod_{i\in[d]} \mathbb{E}[f_i]\right| \leqslant \textit{poly}(d)\gamma \prod_{i\in[d]} ||f_i||_{2^d},$$

where $X[i]$ are the vertices of color/coordinate i. Hint: You may find Exercise C.4 helpful.

B Section 3-4 Exercises

Exercise B.1. *Prove Claim 3.4: $P^{\vee}_{k\to i}$ acts on $\sigma \in X(k)$ by*

1. *(Down-Step): Sampling $\tau \subset \sigma$ of size i uniformly at random*
2. *(Up-Step): Sampling $\sigma' \sim \pi^{\tau}_{k-i}$*

Similarly $P^{\wedge}_{i\to k}$ acts on $\tau \in X(i)$ by

1. *(Up-Step): Sampling $\sigma \sim \pi^{\tau}_{k-i}$*
2. *(Down-Step): Sampling $\tau' \subset \sigma$ of size i uniformly at random*

Exercise B.2. *Prove the Chain Rule for Variance:*

$$Var_{\pi_k}(f) = \mathop{\mathbb{E}}_{v\sim\pi_1}\left[Var_{\pi^v_{k-1}}(f|_v)\right] + \mathop{Var}_{v\sim\pi_1}(\mathbb{E}_{\pi^v_{k-i}}[f|_v])$$

Exercise B.3 (Matroids)**.** *A matroid $(U, \mathcal{I})$ consists of a ground set U and a family of subsets of $\mathcal{I} \subset P(U)$ called 'independent sets' such that*

1. ***Downward Closure:*** *If $\sigma \in \mathcal{I}$, then all $\tau \subset \sigma$ are in $\mathcal{I}$*
2. ***Exchange Property:*** *given independent sets $|I_1| > |I_2|$, then $\exists v \in I_1$ s.t. $\{v\} \cup I_2 \in \mathcal{I}$*

The rank of a set $I \in \mathcal{I}$ is its size. Prove the simplicial complex generated by the uniform distribution on rank-k elements of any matroid $(U, \mathcal{I})$ is a 0-local-spectral HDX.

Exercise B.4 (Basis Exchange Walk)**.** *The rank-k basis exchange walk on matroid $(U, \mathcal{I})$ is the procedure which starting at $\sigma \in \mathcal{I}(k)$, removes a uniformly random vertex $v \in \sigma$, then adds a uniformly random vertex w such that $\sigma/v \cup \{w\} \in \mathcal{I}$. Deduce from the above that the basis exchange walk satisfies*

$$\lambda_2 \leqslant 1 - \frac{1}{k}.$$

Prove this is tight — there exists a matroid such that the basis exchange walk satisfies $\lambda_2 \geqslant 1 - \frac{1}{k}$

C Section 5

Exercise C.1 (Expander-Samplers)**.** *Prove every λ-expander is an (α, β)-sampler graph for*

$$\beta \leqslant O(\lambda^2/\alpha^2).$$

Exercise C.2 (Flipping Sampling)**.** *Prove if $G = (L, R, E)$ is an (α, β)-sampler, then $G^{rev} = (R, L, E)$ is an $(O(\beta), O(\alpha))$-sampler*

Exercise C.3 (Failure of Expander Walks)**.** *Prove there exists an expander graph G whose d-dimensional walk complex $X_{G_{walk}}$ fails d vs 2 sampling. I.e. construct $A \subset [n]$ such that*

$$\mathbf{Pr}_{s=(x_1,\ldots,x_d)\sim Walk_G(d)\}}\left[\left|\mathop{\mathbb{E}}_{\{x_i,x_j\}\subset s}[f(x_i,x_j)] - \mathbb{E}[f]\right| > 0.01\right] > 0.01.$$

In fact, prove such a construction is possible for any high girth graph G.

Exercise C.4 (Partite Swap Walks)**.** *Let X be a partite γ-product. Define the* ***partite swap walk*** *$E_{S,T}$ by*

$$E_{S,T}f(y_T) = \mathbb{E}_{x\sim\pi_d}[f(x_S) \mid x_T = y_T]$$

Prove $E_{S,T}$ expands when $S \cap T = \emptyset$:

$$\lambda_2(E_{S,T}) \leqslant |S||T|\gamma$$

What's the best bound you can prove?

Exercise C.5 (Key Efron-Stein Lemma)**.** *Deduce from the previous exercise by reducing the general case to the above in links of color $S \cap T$ that*

$$||E_S E_T - E_{S\cap T}|| \leqslant |S||T|\gamma$$