Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 75 (2018)

ECCC

Boolean function analysis on high-dimensional expanders*

Yotam Dikstein† Irit Dinur† Yuval Filmus‡ Prahladh Harsha§

January 29, 2019

Abstract

We initiate the study of Boolean function analysis on high-dimensional expanders. We give a random-walk based definition of high dimensional expansion, which coincides with the earlier definition in terms of two-sided link expanders. Using this definition, we describe an analogue of the Fourier expansion and the Fourier levels of the Boolean hypercube for simplicial complexes. Our analogue is a decomposition into approximate eigenspaces of random walks associated with the simplicial complexes. We then use this decomposition to extend the Friedgut-Kalai-Naor theorem to high-dimensional expanders.

Our results demonstrate that a high-dimensional expander can sometimes serve as a sparse model for the Boolean slice or hypercube, and quite possibly additional results from Boolean function analysis can be carried over to this sparse model. Therefore, this model can be viewed as a derandomization of the Boolean slice, containing only $|X(k-1)| = O(n)$ points in contrast to $\binom{n}{k}$ points in the (k)-slice (which consists of all n-bit strings with exactly k ones).

Our random-walk definition and the decomposition has the additional advantage that they extend to the more general setting of posets, which include both high-dimensional expanders and the Grassmann poset, which appears in recent works on the unique games conjecture.

1 Introduction

Boolean function analysis is an essential tool in theory of computation. Traditionally, it studies functions on the Boolean cube $\{-1,1\}^n$. Recently, the scope of Boolean function analysis has been extended further, encompassing groups [EFF15b, EFF15a, Pla15, EFF17], association schemes [OW13, Fil16a, Fil16b, FM16, FKMW16, DKK+18a, KMS18], error-correcting codes [BGH+15], and quantum Boolean functions [MO10]. Boolean function analysis on extended domains has led to progress in learning theory [OW13] and on the unique games conjecture [KMS17, DKK+18a, DKK+18b, BKS19, KMS18].

Another essential tool in theory of computation is expander graphs. Recently, high-dimensional expanders (HDXs), originally constructed by Lubotzky, Samuels and Vishne [LSV05a, LSV05b], have been used in computer science, with applications to property testing [DK17], lattices [KM18] and list decoding [DHK+19]. Just as expander graphs are sparse models of the complete graph, so are high-dimensional expanders sparse models of the *complete hypergraph*, and hence can be potentially used both for derandomization and to improve constructions of objects such as PCPs.

The goal of this work is to connect these two threads of research, by introducing Boolean function analysis on high-dimensional expanders.

We study Boolean functions on simplicial complexes. A pure d-dimensional simplicial complex X is a set system consisting of an arbitrary collection of sets of size $d+1$ together with all their subsets. The sets in a simplicial complex are called *faces*, and it is standard to denote by $X(i)$ the faces of X whose cardinality is $i+1$. Our simplicial complexes are *weighted* by a probability distribution Π_d on the top-level faces, which induces in a natural way probability distributions Π_i on $X(i)$ for all i: we choose

*A preliminary version of this paper appeared in *Proc. 20th RANDOM*, 2018 [DDFH18].

†Weizmann Institute of Science, ISRAEL. email: {yotam.dikstein,irit.dinur}@weizmann.ac.il. The research of the first and second authors was supported in part by Irit Dinur's ERC-CoG grant 772839.

‡Technion Israel Institute of Technology, ISRAEL. email: yuvalfi@cs.technion.ac.il. Taub Fellow — supported by the Taub Foundations. The research was funded by ISF grant 1337/16.

§Tata Institute of Fundamental Research, INDIA. email: prahladh@tifr.res.in. Research supported in part by UGC–ISF grant. Part of the work was done when the author was visiting the Weizmann Institute of Science.

ISSN 1433-8092

$s \sim \Pi_d$, and then choose an i-face $t \subset s$ uniformly at random. Our main object of study is the space of functions $f\colon X(d) \to \mathbb{R}$, and in particular, Boolean functions $f\colon X(d) \to \{0,1\}$.

1.1 Random-walk based definition of high-dimensional expanders

While much of our work applies to arbitrary complexes, our goal is to study complexes which are *high-dimensional expanders*. There are several different non-equivalent ways to define high-dimensional expanders, generalizing different properties of expander graphs. One of the main definitions, *two-sided link expansion* due to Dinur and Kaufman [DK17], extends the spectral definition of expander graphs by requiring two-sided spectral expansion in every link[1]. Dinur and Kaufman [DK17] shows how to construct complexes satisfying this definition from the Ramanujan complexes of Lubotzky, Samuels and Vishne [LSV05a, LSV05b].

We propose a new definition based on high-dimensional random walks on $X(i)$. Denote the real-valued function space on $X(i)$ by $C^i := \{f\colon X(i) \to \mathbb{R}\}$. There are two natural operators $U_i\colon C^i \to C^{i+1}$ and $D_{i+1}\colon C^{i+1} \to C^i$, which are defined by averaging:

$$U_i f(s) := \mathop{\mathbb{E}}_{t\sim\Pi_i}[f(t)|t \subset s] \quad \left(= \frac{1}{i+2}\sum_{t\subset s} f(t)\right),$$
$$D_{i+1} f(t) := \mathop{\mathbb{E}}_{s\sim\Pi_{i+1}}[f(s)|s \supset t].$$

The compositions $D_{i+1}U_i$ and $U_{i-1}D_i$ are Markov operators of two natural random walks on $X(i)$, the *upper random walk* and the *lower random walk*.

The first walk we consider is the *upper random walk* $D_{i+1}U_i$. Given a face $t_1 \in X(i)$, we choose its neighbour t_2 as follows: we pick a random $s \sim \Pi_{i+1}$ conditioned on $s \supset t_1$ and then choose uniformly at random $t_2 \subset s$. Note that there is a probability of $\frac{1}{i+2}$ that $t_1 = t_2$. We define the *non-lazy upper random walk* by choosing $t_2 \subset s$ conditioned on $t_1 \neq t_2$. We denote the Markov operator of the non-lazy upper walk by M_i^+.

Similarly, the *lower random walk* $U_{i-1}D_i$ is another random walk on $X(i)$. Here, given a face $t_1 \in X(i)$, we choose a neighbour t_2 as follows: we first choose a $r \in X(i-1)$ uniformly at random and then choose a $t_2 \sim \Pi_i$ conditioned on $t_2 \supset r$.

For instance, if X is a graph (a 1-dimensional simplicial complex), then the non-lazy upper random walk is the usual adjacency walk we define on a weighted graph (i.e. traversing from vertex to vertex by an edge). The (lazy) upper random walk has probability $\frac{1}{2}$ of staying in place, and probability $\frac{1}{2}$ of going to different adjacent vertex. The lower random walk on $V = X(0)$ doesn't depend on the current vertex: it simply chooses a vertex at random according to the distribution Π_0 on $X(0)$.

There are several works on these random walks on high-dimensional expanders, which naturally lead to analyzing both real-valued and Boolean-valued functions on $X(i)$, for example [KM18, DK17, KO18]. The most related work is by Kaufman and Oppenheim [KO18], who gave a correspondence between a function $f\colon X(i) \to \mathbb{R}$ and a sequence of functions $\{h_j\colon X(j) \to \mathbb{R}\}_{j=-1}^{i}$. This correspondence has the property that

$$\|f\|^2 \approx \sum_{j=-1}^{i} \|h_j\|^2,$$

and that

$$\langle M_i^+ f, f\rangle \approx \sum_{j=-1}^{i} \left(1 - \frac{j+1}{i+2}\right) \|h_j\|^2.$$

The error in the approximation depends on the one-sided expansion of the complex.

We are now ready to give our definition of a high dimensional expander in terms of these walks.

[1] A related and slightly weaker notion of one-sided spectral expansion appeared in earlier works of Kaufman, Kazhdan and Lubotzky [KKL14] and Evra and Kaufman [EK16].

Definition 1.1 (High-Dimensional Expander)**.** *Let $\gamma < 1$, and let X be a d-dimensional simplicial complex. We say X is a* γ-high dimensional expander *(or* γ-HDX*) if for all $0 \le i \le d-1$, the non-lazy upper random walk is γ-similar to the lower random walk in operator norm in the following sense:*

$$\left\| M_i^+ - U_{i-1} D_i \right\| \le \gamma.$$

In the graph case, this coincides with the definition of a γ-two-sided spectral expander: recall that the lower walk on $X(0)$ is by choosing two vertices $v_1, v_2 \in X(0)$ independently. Thus $\left\| M_i^+ - U_{i-1} D_i \right\|$ is the second eigenvalue of the adjacency random walk in absolute value. For $i \ge 1$, we cannot expect the upper random walk to be similar to choosing two independent faces in $X(i)$, since the faces always share a common intersection of i elements. Instead, our definition asserts that traversing through a common $(i+1)$-face is similar to traversing through a common $(i-1)$-face.

We show that this new definition coincides with the aforementioned definition of two-sided link expanders, thus giving these high-dimensional expanders a new characterization. Through this characterization, we decompose real-valued functions $f \colon X(i) \to \mathbb{R}$ in an approximately orthogonal decomposition that respects the upper walk and lower walk operators.

1.2 Decomposition of functions on $X(i)$

We being by recalling the *classical* decomposition of functions over the Boolean hypercube. Every function on the Boolean cube $\{0,1\}^n$ has a unique representation as a multilinear polynomial. In the case of the Boolean hypercube, it is convenient to view the domain as $\{1,-1\}^n$, in which case the above representation gives the *Fourier expansion* of the function. The multilinear monomials can be partitioned into "levels" according to their degree, and this corresponds to an orthogonal decomposition of a function into a sum of its homogeneous parts, $f = \sum_{i=0}^{\deg f} f^{=i}$, a decomposition which is a basic concept in Boolean function analysis.

These concepts have known counterparts for the *complete complex*, which consists of all subsets of $[n]$ of size at most $d+1$, where $d+1 \le n/2$. The facets (top-level faces) of this complex comprise the *slice* (as it is known to computer scientists) or the *Johnson scheme* (as it is known to coding theorists), whose spectral theory has been elucidated by Dunkl [Dun76]. For $|t| \le d+1$, let $y_t(s) = 1$ if $t \subseteq s$ and $y_t(s) = 0$ otherwise (these are the analogs of monomials). Every function on the complete complex has a unique representation as a linear combination of monomials $\sum_t \tilde{f}(t) y_t$ (of various degrees) where the coefficients $\tilde{f}(t)$ satisfy the following *harmonicity* condition: for all $i \le d$ and all $t \in X(i)$,

$$\sum_{a \in [n] \setminus t} \tilde{f}(t \cup \{a\}) = 0.$$

(If we identify y_t with the product $\prod_{i \in t} x_i$ of "variables" x_i, then harmonicity of a multilinear polynomial P translates to the condition $\sum_{i=1}^n \frac{\partial P}{\partial x_i} = 0$.) As in the case of the Boolean cube, this unique representation allows us to orthogonally decompose a function into its homogeneous parts (corresponding to the contribution of monomials y_t with fixed $|t|$), which plays the same essential part in the complete complex as its counterpart does in the Boolean cube. Moreover, this unique representation allows extending a function from the "slice" to the Boolean cube (which can be viewed as a superset of the "slice"), thus implying further results such as an invariance principle [FKMW16, FM16].

We generalize these concepts for complexes satisfying a technical condition we call *properness*, which is satisfied by both the complete complex and high-dimensional expanders. We show that the results on unique decomposition for the complete complex hold for arbitrary proper complexes, with a generalized definition of harmonicity which incorporates the distributions Π_i. In contrast to the case of the complete complex (and the Boolean cube), in the case of high-dimensional expanders the homogeneous parts are only approximately orthogonal.

The homogeneous components in our decomposition are "approximate eigenfunctions" of the Markov operators defined above, and this allows us to derive an approximate identity relating the total influence (defined through the random walks) to the norms of the components in our decomposition, in complete analogy to the same identity in the Boolean cube (expressing the total influence in terms of the Fourier expansion). All of this is summarized in Theorem 4.6.

1.3 Decomposition of posets

The decomposition we suggest in this paper holds for the more general setting of graded partially ordered sets (posets): A finite *graded poset* $(X, \leq, \rho)$ is a poset $(X, \leq)$ equipped with a *rank* function $\rho \colon X \to \{-1\} \cup \mathbb{N}$ that respects the order, i.e. if $x \leq y$ then $\rho(x) \leq \rho(y)$. Additionally, if y is minimal with respect to elements that are greater than x (i.e. y *covers* x), then $\rho(y) = \rho(x) + 1$. Denoting $X(i) = \rho^{-1}(i)$, we can partition the poset as follows:

$$X = X(-1) \cup X(0) \cup \cdots \cup X(d).$$

We consider graded posets with a unique minimum element $\emptyset \in X(-1)$.

Every simplicial complex is a graded poset. Another notable example is the *Grassmann poset* $\mathrm{Gr}_q(n, d)$ which consists of all subspaces of F_q^n of dimension at most $d+1$. The order is the containment relation, and the rank is the dimension minus one, $\rho(W) = \dim(W) - 1$. The Grassmann poset was recently studied in the context of proving the 2-to-1 games conjecture [KMS17, DKK+18a, DKK+18b, KMS18], where a decomposition of functions of the Grassmann poset was useful. Such a decomposition is a special case of the general decomposition theorem in this paper.

Towards our goal of decomposing functions on graded posets, we generalize the notion of random walks on $X(i)$ as follows: A *measured poset* is a graded poset with a sequence of measures $\vec{\Pi} = (\Pi_{-1}, \ldots, \Pi_d)$ on the different levels $X(i)$, that allow us to define operators U_i, D_{i+1} similar to the simplicial complex case (for a formal definition see Section 8). The *upper random walk* defined by the composition $D_{i+1}U_i$ is the walk where we choose two consecutive $t_1, t_2 \in X(i)$ by choosing $s \in X(i+1)$ and then $t_1, t_2 \leq s$ independently. The *lower random walk* $U_{i-1}D_i$ is the walk where we choose two consecutive $t_1, t_2 \in X(i)$ by choosing $r \in X(i-1)$ and then $t_1, t_2 \geq r$ independently.

Stanley studied a special case of a measured poset that is called a *sequentially differential poset* [Sta88]. This is a poset where

$$D_{i+1}U_i - r_i I - \delta_i U_{i-1} D_i = 0, \tag{1}$$

for all $0 \leq i \leq d$ and some constants $r_i, \delta_i \in \mathbb{R}$. There are many interesting examples of sequentially differential posets, such as the Grassmann poset and the complete complex. Definition 1.1 of a high-dimensional expander resembles an approximate version of this equation: in a simplicial complex, one may check that the non-lazy version is $\frac{i+1}{i+2} M_+^i = D_{i+1}U_i - \frac{1}{i+2} I$. Thus

$$\left\| M_i^+ - U_{i-1}D_i \right\| \leq \gamma$$

is equivalent to

$$\left\| D_{i+1}U_i - \frac{1}{i+2} - \frac{i+1}{i+2} U_{i-1}D_i \right\| \leq \frac{i+1}{i+2}\gamma,$$

which suggests a relaxation of (1) to an *expanding poset* (eposet).

Definition 1.2 (Expanding Poset (eposet)). *Let $\vec{r}, \vec{\delta} \in \mathbb{R}^k_{\geq 0}$, and let $\gamma < 1$. We say X is an $(\vec{r}, \vec{\delta}, \gamma)$-*expanding poset *(or $(\vec{r}, \vec{\delta}, \gamma)$-*eposet*) if for all $0 \leq i \leq k-1$:*

$$\| D_{i+1}U_i - r_i I - \delta_i U_{i-1} D_i \| \leq \gamma. \tag{2}$$

As we can see, γ-HDX is also an $(\vec{r}, \vec{\delta}, \gamma)$-eposet, for $r_i = \frac{1}{i+2}, \delta_i = \frac{i+1}{i+2}$. In Lemma 8.18 we prove that the converse is also true: every simplicial complex that is an $(\vec{r}, \vec{\delta}, \gamma)$-eposet is an $O(\gamma)$-HDX, under the assumption that the probability $\Pr_{t_1, t_2 \sim U_{i-1}D_i}[t_1 = t_2]$ is small.

It turns out that eposets are the correct setup to generalize our decomposition of simplicial complexes: in all eposets we can uniquely decompose functions $f \colon X(i) \to \mathbb{R}$ to

$$f = \sum_{j=-1}^{i} f^{=j},$$

where the functions $f^{=j}$ are "approximate eigenvectors" of $D_{i+1}U_i$. Furthermore, this decomposition is "approximately orthogonal". Fixing i, the error in both approximations is $O(\gamma)$.

1.4 An FKN theorem

Returning to simplicial complexes, as a demonstration of the power of this setup, we generalize the fundamental result of Friedgut, Kalai, and Naor [FKN02] on Boolean functions almost of degree 1. We view this as a first step toward developing a full-fledged theory of Boolean functions on high-dimensional expanders.

An easy exercise shows that a Boolean degree 1 function on the Boolean cube is a *dictator*, that is, depends on at most one coordinate; we call this the *degree one theorem* (the easy case of the FKN Theorem with zero-error) . The FKN theorem, which is the robust version of this degree one theorem, states that a Boolean function on the Boolean cube which is *close* to a degree 1 function is in fact close to a dictator, where closeness is measured in L_2.

The degree one theorem holds for the complete complex as well. Recently, the third author [Fil16a] extended the FKN theorem to the complete complex. Surprisingly, the class of approximating functions has to be extended beyond just dictators.

We prove an a degree one theorem for arbitrary proper complexes, and an FKN theorem for high-dimensional expanders. In contrast to the complete complex, Boolean degree 1 functions on arbitrary complexes correspond to *independent sets* rather than just single points, and this makes the proof of the degree one theorem non-trivial.

Our proof of the FKN theorem for high-dimensional expanders is very different from existing proofs. It follows the same general plan as recent work on the biased Kindler–Safra theorem [DFH19]. The idea is to view a high-dimensional expander as a convex combination of small sub-complexes, each of which is isomorphic to the complete k-dimensional complex on $O(k)$ vertices. We can then apply the known FKN theorem separately on each of these, and deduce that our function is approximately well-structured on each sub-complex. Finally, we apply the agreement theorem of Dinur and Kaufman [DK17] to show that the same holds on a global level.

1.5 Our results

Our first result is a decomposition for functions on any high-dimensional expander:

Theorem 1.3 (Decomposition theorem for functions on HDX)**.** *Let X be a proper d-dimensional simplicial complex.*[2] *Every function $f\colon X(\ell) \to \mathbb{R}$, for $\ell \le d$, can be written uniquely as $f = f_{-1} + \cdots + f_\ell$ such that:*

- *f_i is a linear combination of the functions $y_s(t) = 1_{[t \supseteq s]}$ for $s \in X(i)$, i.e. $y_s(t) = 1$ when $t \supseteq s$.*
- *Interpreted as a function on $X(i)$, f_i lies in the kernel of the Markov operator of the lower random walk UD.*

If X is furthermore a γ-high dimensional expander, then the above decomposition is an almost orthogonal decomposition in the following sense:

- *For $i \neq j$, $|\langle f_i, f_j\rangle| \approx 0$.*
- *$\|f\|^2 \approx \|f_{-1}\|^2 + \cdots + \|f_\ell\|^2$.*
- *If $\ell < k$ then $DUf_i \approx (1 - \frac{i+1}{\ell+2})f_i$, and in particular $\langle DUf, f\rangle \approx \sum_{i=-1}^{\ell}(1 - \frac{i+1}{\ell+2})\|f_i\|^2$.*

(For an exact statement in terms of the dependence of error on γ, see Theorem 4.6).

In Section 8 we give a more general version of this theorem that applies to arbitrary expanding posets.

Our proof of Theorem 1.3 uses the random-walk based definition of high dimensional expander, which appears in Section 4. In Section 5 we show that this definition is equivalent to the earlier notion of a two-sided link expander due to Dinur and Kaufman [DK17], up to a constant factor:

Theorem 1.4 (Equivalence between high-dimensional expander definitions)**.** *Let X be a d-dimensional simplicial complex.*

[2] A simplicial complex is *proper* if the Markov operators of the upper random walks DU have full rank. All high-dimensional expanders satisfy this property.

1. *If X is a γ-two-sided link expander according to the definition in [DK17] then X is a γ-HDX according to the definition we give.*

2. *If X is a γ-HDX then X is a $3d\gamma$-two-sided link expander according to the definition in [DK17] .*

Equipped with the decomposition theorem, we prove the following degree one theorem and its robust version, the FKN theorem on high-dimensional expanders.

Definition 1.5 (1-skeleton)**.** *The* 1-skeleton *of a simplicial complex X is the graph whose vertices are $X(0)$, the 0-faces of the complex, and whose edges are $X(1)$, the 1-faces of the complex.*

Theorem 1.6 (Degree one theorem on simplicial complexes)**.** *Let X be a d-dimensional simplicial complex whose 1-skeleton is connected. If $f\colon X(d) \to \{0,1\}$ has degree 1, then f is the indicator of either intersecting or not intersecting an independent set of X.*

Theorem 1.7 (FKN theorem on HDX (informal))**.** *Let X be a d-dimensional γ-HDX . If $F\colon X(d) \to \{0,1\}$ is ε-close (in L_2^2) to a degree 1 function then there exists a degree 1 function g on $X(d)$ such that $\Pr[F \neq g] = O_{\gamma,d}(\varepsilon)$.*

Paper organization

We describe our general setup in Section 2. We describe the property of properness and its implications — a unique representation theorem and decomposition of functions into homogeneous parts — in Section 3. We introduce our definition of high-dimensional expanders in Section 4. In Section 5 we show equivalence between our definition and the earlier one of two-sided link expanders. We prove our degree one theorem in Section 6, and our FKN theorem in Section 7.

In Section 8 we define expanding posets, and through them prove that the decomposition in Theorem 3.2 is almost orthogonal. We also show that expanding posets that are simplicial complexes, are in fact high-dimensional expanders. Theorem 4.6 summarizes these results for simplicial complexes.

Theorem 1.3 is a combination of Theorem 3.2 (first two items) and Theorem 4.6 (other three items). Theorem 1.4 is a restatement of Theorem 5.5. Theorem 1.6 is a restatement of Theorem 6.2. Theorem 1.7 is a restatement of Theorem 7.3.

2 Basic setup

A d-dimensional simplicial complex X is a non-empty collection of sets of size at most $d+1$ which is closed under taking subsets. We call a set of size $i+1$ an *i-dimensional face* (or i-face for short), and denote the collection of all i-faces by $X(i)$. A d-dimensional simplicial complex X is *pure* if every i-face is a subset of some d-face. We will only be interested in pure simplicial complexes.

Let X be a pure d-dimensional simplicial complex. Given a probability distribution Π_d on its top-dimensional faces $X(d)$, for each $i < d$ we define a distribution Π_i on the i-faces using the following experiment: choose a top-dimensional face according to Π_d, and remove $d-i$ points at random. We can couple all of these distributions to a random vector $\vec{\Pi} = (\Pi_d, \ldots, \Pi_{-1})$ of which the individual distributions are marginals.

Let $C^i := \{f\colon X(i) \to \mathbb{R}\}$ be the space of functions on $X(i)$. It is convenient to define $X(-1) := \{\emptyset\}$, and we also let $C^{-1} := \mathbb{R}$. We turn C^i to an inner product space by defining $\langle f, g\rangle := \mathbb{E}_{\Pi_i}[fg]$ and the associated norm $\|f\|^2 := \mathbb{E}_{\Pi_i}[f^2]$.

For $-1 \le i < d$, we define the **Up** operator $U_i\colon C^i \to C^{i+1}$ as follows:[3]

$$U_i g(s) := \frac{1}{i+2}\sum_{x\in s} g(s \setminus \{x\}) = \mathop{\mathbb{E}}_{t\subset s}[g(t)] \,,$$

where t is obtained from s by removing a random element. Note that if $s \sim \Pi_{i+1}$ then $t \sim \Pi_i$.

[3] The Up and Down operators differ from the boundary and coboundary operators of algebraic topology, which operate on linear combinations of *oriented* faces.

Similarly, we define the **Down** operator $D_{i+1}\colon C^{i+1} \to C^i$ for $-1 \le i < d$ as follows:

$$D_{i+1}f(t) := \frac{1}{(i+2)\cdot \Pi_i(t)} \sum_{x\notin t\colon\, t\cup\{x\}\in X(i+1)} \Pi_{i+1}(t\cup\{x\})\cdot f(t\cup\{x\}) = \mathop{\mathbb{E}}_{s\supset t}[f(s)] \ ,$$

where s is obtained from t by conditioning the vector $\vec{\Pi}$ on $\Pi_i = t$ and taking the $(i+1)$th component.

The operators U_i, D_{i+1} are adjoint to each other. Indeed, if $f \in C^{i+1}$ and $g \in C^i$ then

$$\langle g, D_{i+1} f\rangle = \mathop{\mathbb{E}}_{(t,s)\sim(\Pi_i,\Pi_{i+1})}[g(t)f(s)] = \langle U_i g, f\rangle \ .$$

When the domain is understood, we will use U, D instead of U_i, D_{i+1}. This will be especially useful when considering powers of U, D. For example, if $f\colon X(i) \to \mathbb{R}$ then

$$U^t f \equiv U_{i+t-1}\dots U_{i+1}U_i f.$$

Given a face $s \in X$, the function y_s is the indicator function of containing s. Our definition of the Up operator guarantees the correctness of the following lemma.

Lemma 2.1. *Let $s \in X(i)$. We can think of y_s as a function in C^j for all $j \ge i$. Using this convention, $U_j y_s = (1 - \frac{i+1}{j+2})y_s$.*

Proof. Direct calculation shows that

$$(U_j y_s)(t) = \frac{1}{j+2}\sum_{x\in t} y_s(t\setminus\{x\}) = \frac{|t|-|s|}{j+2} y_s(t) \ ,$$

and so $U_j y_s = (1-\frac{i+1}{j+2})y_s$. □

For $0 \le i \le k$, the space of harmonic functions on $X(i)$ is defined as

$$H^i := \ker D_i = \{f \in C^i : D_i f = 0\} \ .$$

We also define $H^{-1} := C^{-1} = \mathbb{R}$. We are interested in decomposing C^k, so let us define for each $-1 \le i \le k$,

$$V^i := U^{k-i}H^i = \{U^{k-i}f \ : \ f \in H^i\} \ .$$

We can describe V^i, a sub-class of functions of C^k, in more concrete terms.

Lemma 2.2. *Every function $h \in V^i$ has a representation of the form*

$$h = \sum_{s\in X(i)} \tilde{h}(s) y_s \ ,$$

where the coefficients $\tilde{h}(s)$ satisfy the following harmonicity *condition: for all $t \in X(i-1)$,*

$$\sum_{s\supset t} \Pi_i(s)\tilde{h}(s) = 0 \ .$$

Furthermore, if U^{k-i} is injective on C^i then the representation is unique.

Proof. Suppose that $h \in V^i$. Then $h = U^{k-i}f$ for some $f \in H^i$, which by definition of H^i and the Down operator is equivalent to the condition

$$\sum_{s\supset t} \Pi_i(s) f(s) = 0$$

for all $t \in X(i-1)$. In other words, the $f(s)$'s satisfy the harmonicity condition. It is easy to check that $f = \sum_{s\in X(i)} f(s)y_s$, and so Lemma 2.1 shows that $h = \sum_{s\in X(i)} \tilde{h}(s)y_s$, where

$$\tilde{h}(s) = \left(1 - \frac{i+1}{k+1}\right)\cdots\left(1-\frac{i+1}{i+2}\right) f(s).$$

Thus, $\tilde{h}(s)$ is a scaling of $f(s)$ by a non-zero constant, it follows that the coefficients $\tilde{h}(s)$ also satisfy the harmonicity condition.

Now suppose that U^{k-i} is injective on C^i, which implies that $\dim H^i = \dim V^i$. The foregoing shows that the dimension of the space of coefficients $\tilde{h}(s)$ satisfying the harmonicity conditions is $\dim H^i$. Since $\dim H^i = \dim V^i$, this shows that the representation is unique. □

3 Decomposition of the space C^k and a convenient basis

Our decomposition theorem relies on a crucial property of simplicial complexes, *properness*.

Definition 3.1. *A k-dimensional simplicial complex is* proper *if $D_{i+1}U_i > 0$ (i.e. $D_{i+1}U_i$ is positive definite) for all $i \le k-1$. Equivalently, if it is pure and $\ker U_i$ is trivial for $-1 \le i \le k-1$.*

We remark that since DU is PSD, $\ker U = 0$ is equivalent to $DU > 0$. This is because for any $x \in \ker DU$, we would have $0 = \langle x, DUx\rangle = \|Ux\|^2$, implying that $x = 0$.

The complete k-dimensional complex on n points is proper iff $k+1 \le \frac{n+1}{2}$. A pure one-dimensional simplicial complex (i.e., a graph) is proper iff it is not bipartite. Unfortunately, we are not aware of a similar characterization for higher dimensions. However, in Section 5 we show that high-dimensional expanders are proper.

We can now state our decomposition theorem.

Theorem 3.2. *If X is a proper k-dimensional simplicial complex then we have the following decomposition of C^k:*

$$C^k = V^k \oplus V^{k-1} \oplus \cdots \oplus V^{-1} .$$

In other words, for every function $f \in C^k$ there is a unique choice of $h_i \in H^i$ such that the functions $f_i = U^{k-i}h_i$ satisfy $f = f_{-1} + f_0 + \cdots + f_k$.

Proof. We first prove by induction on ℓ that every function $f \in C^\ell$ has a representation $f = \sum_{i=-1}^{\ell} U^{\ell-i}h_i$, where $h_i \in H^i$. This trivially holds when $\ell = -1$. Suppose now that the claim holds for some $\ell < k$, and let $f \in C^{\ell+1}$. Since $D^{\ell+1}\colon C^{\ell+1} \to C^\ell$ is a linear operator, we have $C^{\ell+1} = \ker D_{\ell+1} \oplus \operatorname{im} D^*_{\ell+1} = \ker D_{\ell+1} \oplus \operatorname{im} U_\ell$, and therefore we can write $f = h_{\ell+1} + Ug$, where $h_{\ell+1} \in H^{\ell+1}$ and $g \in C^\ell$. Applying induction, we get that $g = \sum_{i=-1}^{\ell} U^{\ell-i}h_i$, where $h_i \in H^i$. Substituting this in $f = h_{\ell+1} + Ug$ completes the proof.

It remains to show that the representation is unique. Since $\ker U_{i-1} = \ker D^*_i$ is trivial, $\dim H^i = \dim C^i - \dim C^{i-1}$ for $i \ge 0$. This shows that $\sum_{i=-1}^{k} \dim H^i = \dim C^k$. Therefore the operator $\varphi\colon H^{-1} \times \cdots \times H^k \to C^k$ given by $\varphi(h_{-1}, \ldots, h_k) = \sum_{i=-1}^{k} U^{k-i}h_i$ is not only surjective but also injective. In other words, the representation of f is unique. $\square$

Corollary 3.3. *If X is a proper k-dimensional simplicial complex then every function $f \in C^k$ has a unique representation of the form*

$$f = \sum_{s \in X} \tilde{f}(s) y_s ,$$

where the coefficients $\tilde{f}(s)$ satisfy the following harmonicity *conditions: for all $0 \le i \le k$ and all $t \in X(i-1)$:*

$$\sum_{\substack{s \in X(i) \\ s \supset t}} \Pi_i(s)\tilde{f}(s) = 0 .$$

Proof. Follows directly from Lemma 2.2. $\square$

We can now define the degree of a function.

Definition 3.4. *The* degree *of a function f is the maximal cardinality of a face s such that $\tilde{f}(s) \neq 0$ in the unique decomposition given by Corollary 3.3.*

Thus a function has degree d if its decomposition only involves faces whose dimension is less than d. The following lemma shows that the functions y_s, for all $(d-1)$-dimensional faces s, form a basis for the space of all functions of degree at most d.

Lemma 3.5. *If X is a proper k-dimensional simplicial complex then the space of functions on $X(k)$ of degree at most $d+1$ has the functions $\{y_s : s \in X(d)\}$ as a basis.*

Proof. The space of functions on $X(k)$ of degree at most $d+1$ is spanned, by definition, by the functions y_t for $t \in X(-1) \cup X(0) \cup \cdots \cup X(d)$. This space has dimension $\sum_{i=-1}^{d} \dim H^i$. Since X is proper, $\dim H^i = \dim C^i - \dim C^{i-1}$ for $i > 0$, and so $\sum_{i=1}^{d} \dim H^i = \dim C^d = |X(d)|$.

Given the above, in order to complete the proof, it suffices to show that for every $i \le d$ and $t \in X(i)$, the function y_t can be written as a linear combination of y_s for $s \in X(d)$. This will show that $\{y_s : s \in X(d)\}$ spans the space of functions of degree at most $d+1$. Since this set contains $|X(d)|$ functions, it forms a basis.

Recall that $y_t(r) = 1_{r \supseteq t}$, where $r \in X(k)$. If r contains t then it contains exactly $\binom{k+1-|t|}{d+1-|t|}$ many d-faces containing r, and so

$$y_t = \frac{1}{\binom{k+1-|t|}{d+1-|t|}} \sum_{\substack{s \supseteq t \\ s \in X(d)}} y_s.$$

This completes the proof. □

We call f_i the "level i" part of f, and denote the weight of f above level i by

$$wt_{>i}(f) := \sum_{j>i} \|f_j\|_2^2.$$

We also define $f_{\le i} = f_{-1} + \cdots + f_i$ and $f_{>i} = f - f_{\le i}$.

4 How to define high-dimensional expansion?

In this section we define a class of simplicial complexes which we call γ-high-dimensional expanders (or γ-HDXs). We later show that these simplicial complexes coincide with the high-dimensional expanders defined by Dinur and Kaufman [DK17] via spectral expansion of the links. In addition, we will show the decomposition in Section 3 is almost orthogonal for γ-HDXs. We will define γ-HDXs through relations between random walks in different dimensions. It is easy to already state the definition using the U, D operators: a k-dimensional simplicial complex is said to be a γ-HDX if for all levels $0 \le j \le k-1$,

$$\left\| \frac{j+2}{j+1}\left(DU - \frac{1}{j+2}I\right) - UD \right\| \le \gamma. \tag{3}$$

We turn to explain the meaning of (3) being small by discussing these random walks.[4]

The operators U and D induce random walks on the jth level $X(j)$ of the simplicial complex. Recall that our simplicial complexes come with distributions Π_j on the j-faces.

Definition 4.1 (The upper random walk DU)**.** *Given $t \in X(j)$, we choose the next set $t' \in X(j)$ as follows:*

- *Choose $s \sim \Pi_{j+1}$ conditioned on $t \subset s$.*
- *Choose uniformly at random $t' \in X(j)$ such that $t' \subset s$.*

Definition 4.2 (The lower random walk UD)**.** *Given $t \in X(j)$, we choose the next set $t' \in X(j)$ as follows:*

- *Choose $t \sim \Pi_j$.*
- *Choose uniformly at random $r \in X(j-1)$ such that $r \subset t$.*
- *Choose $t' \sim \Pi_j$ conditioned on $r \subset t'$.*

[4] UD and DU are called high-dimensional Laplacians in some works, such as [KO18].

It is easy to see that the stationary distribution for both these processes is Π_j. However, these random walks are not necessarily the same. For example, if $j = 0$, we consider the graph $(X(0), X(1))$. The upper walk is the $\frac{1}{2}$-lazy version of the usual adjacency random walk in a graph. The lower random walk is simply choosing two vertices independently, according the distribution Π_0. In both walks, the first step and the third step are independent given the second step. In fact, we can view the upper walk (resp. lower walk) as choosing a set $s \in X(j+1)$ (resp. $r \in X(j-1)$), and then choosing independently two sets $t, t' \in X(j)$ given that they are contained in s (resp. given that they contain r).

One property of a random walk is its laziness:

Definition 4.3 (Laziness). *Let M be a random walk. The* laziness *of M is*

$$\ell z(M) = \Pr_{(x,y)\sim M}[x = y].$$

We say that an operator is non-lazy *if $\ell z(M) = 0$.*

It is easy to see that both walks have some laziness. In the upper walk, the laziness is $\frac{1}{j+2}$. We can decompose DU as

$$DU = \frac{1}{j+2} I + \frac{j+1}{j+2} M_j^+, \tag{4}$$

where M_j^+ is the non-lazy version of DU, i.e. the operator representing the walk when conditioning on $t' \neq t$. The laziness of the lower version depends on the simplicial complex itself, thus it doesn't admit a simple decomposition in the general case.

(4) can be written as

$$M_j^+ = \frac{j+2}{j+1}\left(DU - \frac{1}{j+2} I\right).$$

A γ-HDX is a simplicial complex in which the non-lazy upper walk is similar to the lower walk. Thus an equivalent way to state (3) is as follows.

Definition 4.4 (High-dimensional expander). *Let X be a simplicial complex, and let $\gamma < 1$. We say that X is a* γ-HDX *if for all $0 \le j \le k-1$,*

$$\|M_j^+ - UD\| \le \gamma. \tag{5}$$

This definition nicely generalizes spectral expansion in graphs, since if X is a graph, $\|M_j^+ - UD\|$ is the second largest eigenvalue (in absolute value) of the normalized adjacency random walk. In Section 5 we show that this definition is equivalent to the definition of high-dimensional two-sided local spectral expanders that was extensively studied in [DK17, Opp18] and other papers.

If $\gamma < \frac{1}{k+1}$ then any γ-HDX is proper, as shown by the following lemma.

Lemma 4.5. *Let X be a k-dimensional γ-HDX, for $\gamma < \frac{1}{k+1}$. Then X is proper.*

Proof. To prove this, we directly calculate $\langle U_j f, U_j f\rangle$ and show that it is positive when $f \neq 0$:

$$\begin{aligned}\langle Uf, Uf\rangle = \langle DUf, f\rangle &= \frac{1}{j+2}\langle f, f\rangle + \frac{j+1}{j+2}\langle M_j^+ f, f\rangle \\ &= \frac{1}{j+2}\langle f, f\rangle + \frac{j+1}{j+2}\langle (M_j^+ - UD + UD)f, f\rangle.\end{aligned} \tag{6}$$

From Cauchy–Schwartz,

$$\left|\langle (M_j^+ - UD)f, f\rangle\right| \le \left\|(M_j^+ - UD)f\right\| \|f\|,$$

and since X is a γ-HDX,

$$\left\|(M_j^+ - UD)f\right\| \le \gamma \|f\|.$$

Plugging this in (6), we get

$$\frac{1}{j+2}\langle f, f\rangle + \frac{j+1}{j+2}\langle (M_j^+ - UD + UD)f, f\rangle \ge \left(\frac{1}{j+2} - \frac{j+1}{j+2}\gamma\right)\langle f, f\rangle + \langle UDf, f\rangle.$$

The last part of the sum is non-negative: $\langle UDf, f\rangle = \langle Df, Df\rangle \geq 0$. Therefore, if $\gamma < \frac{1}{k+1} \leq \frac{1}{j+1}$ then

$$\left(\frac{1}{j+2} - \frac{j+1}{j+2}\gamma\right)\langle f, f\rangle + \langle DUf, f\rangle \geq \left(\frac{1}{j+2} - \frac{j+1}{j+2}\gamma\right)\langle f, f\rangle > 0.$$

Hence $\langle U_j f, U_j f\rangle > 0$. □

4.1 Almost orthogonality of the decomposition in HDXs

In Section 8 we prove that the decomposition in Theorem 3.2 is "almost orthogonal". We summarize our results below:

Theorem 4.6. *Let X be a k-dimensional γ-HDX, where γ is small enough as a function of k . For every function f on C^ℓ for $\ell \leq k$, the decomposition $f = f_{-1} + \cdots + f_\ell$ of Theorem 3.2 satisfies the following properties:*

- *For $i \neq j$, $|\langle f_i, f_j\rangle| = O(\gamma)\|f_i\|\|f_j\|$.*
- *$\|f\|^2 = (1 \pm O(\gamma))(\|f_{-1}\|^2 + \cdots + \|f_\ell\|^2)$, and for all i, $\|f\|^2 = (1 \pm O(\gamma))(\|f_{\leq i}\|^2 + \|f_{>i}\|^2)$.*
- *f_i are approximate eigenvectors with eigenvalues $\lambda_i = 1 - \frac{i+1}{\ell+2}$ in the sense that $\|DUf_i - (1 - \frac{i+1}{\ell+2})f_i\| = O(\gamma)\,\|f_i\|$.*
- *If $\ell < k$ then $\langle DUf, f\rangle = (1 \pm O(\gamma))\sum_{i=-1}^{\ell} \lambda_i \|f_i\|^2$.*
- *If $\gamma < \frac{1}{k+1}$ then X is proper.*

The hidden constants in the O notations depend only on k (and not on the size of X).

This result is analogous to [KO18, Theorem 6.2], in which a similar decomposition is obtained. However, whereas our decomposition is to functions $f_{-1}, \ldots, f_\ell$ in C^ℓ, the decomposition of Kaufman and Oppenheim [KO18] is to functions $h_{-1}, \ldots, h_\ell$, which live in different spaces.

5 High-dimensional expanders are two-sided link expanders

In Section 4 we defined γ-HDXs, see Definition 4.4. Earlier works, such as [EK16, DK17, KO18] for example, gave a different definition of high-dimensional expanders — *two-sided link expanders* — based on the local link structure. We recall this other definition and prove that the two are equivalent.

Definition 5.1 (Link)**.** *Let X be a d-dimensional complex with an associated probability distribution Π_d on $X(d)$, which induces probability distributions on $X(-1), \ldots, X(d-1)$. For every i-dimensional face $s \in X(i)$ for $i < d-1$, the* link *of s, denoted X_s, is the simplicial complex:*

$$X_s = \{r \setminus s : r \in X, r \supset s\}.$$

We associate X_s with the weights $\vec{\Pi}^s$ such that

$$\Pi^s_j(t) := \Pr_{r \sim \Pi_{i+j+1}}[r = s \cup t | r \supset s] = \frac{\Pi(s \cup t)}{\Pi(s)\binom{|s \cup t|}{|s|}}.$$

Definition 5.2 (Underlying graph)**.** *Let $i < d-1$. Given $s \in X(i)$, the* underlying graph *G_s is the weighted graph consisting of the first two levels of the link of s. In other words, $G_s = (V, E)$, where*

- *$V = X_s(0) = \{x \notin s : s \cup \{x\} \in X(i+1)\}$.*
- *$E = X_s(1) = \{\{x, y\} : s \cup \{x, y\} \in X(i+2)\}$.*

The weights on the edges are given by

$$w_s(\{x, y\}) = \Pr_{r \sim \Pi_{i+2}}[r = s \cup \{x, y\} | r \supset s] = \frac{\Pi(s \cup \{x, y\})}{\Pi(s)\binom{|s|+2}{|s|}}.$$

We can also consider directed edges, by choosing a random orientation:

$$w_s(x,y) = \frac{1}{2} w_s(\{x,y\}).$$

We define the weight of a vertex x to

$$w_s(x) := \Pi_0^s(x) = \Pr_{r \sim \Pi_{i+1}}[r = s \cup \{x\} | r \supset s].$$

We define an inner product for functions on vertices along the lines of Section 2:

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

We denote by A_s the adjacency operator of the *non-lazy upper-walk* on $X_s(0)$, given by

$$A_s f(x) = \mathop{\mathbb{E}}_{y \sim w_s}[f(y) | \{x,y\} \in E].$$

The corresponding quadratic form is

$$\langle f, A_s g \rangle = \mathop{\mathbb{E}}_{(x,y) \sim w_s}[f(x)g(y)].$$

By definition, A_s fixes constant functions, and is a Markov operator. It is self-adjoint with respect to the inner product above. Thus A_s has eigenvalues $\lambda_1 = 1 \geq \lambda_2 \geq \ldots \geq \lambda_m$, where m is the number of vertices. We define $\lambda(A_s) = \max(|\lambda_2|, |\lambda_m|)$. Orthogonality of eigenspaces guarantees that

$$|\langle f, A_s g \rangle - \mathbb{E}[f]\,\mathbb{E}[g]| \leq \lambda(A_s) \|f\| \|g\|. \tag{7}$$

Definition 5.3 (Two-sided link expander)**.** *Let X be a simplicial complex, and let $\gamma < 1$ be some constant. We say that X is a* γ-two-sided link expander *(called γ-HD expander in [DK17]) if every link X_s of X satisfies $\lambda(A_s) \leq \gamma$.*

Dinur and Kaufman [DK17] proved that such expanders do exist, based on a result of [LSV05a].

Theorem 5.4 ([DK17, Lemma 1.5])**.** *For every $\lambda > 0$ and every $d \in \mathbb{N}$ there exists an explicit infinite family of bounded degree d-dimensional complexes which are λ-two-sided link expanders.*

We now prove that *two-sided link expanders* per Definition 5.3 and *high-dimensional expanders* per Definition 4.4 are equivalent.

Theorem 5.5 (Equivalence theorem)**.** *Let X be a d-dimensional simplicial complex.*

1. *If X is a γ-two-sided link expander, then X is a γ-HDX.*
2. *If X is a γ-HDX then X is a $3d\gamma$-two-sided link expander.*

Proof. **Item 1.** Assume that X is a γ-two-sided link expander. We need to show that

$$\left\| M_i^+ - UD \right\| \leq \gamma,$$

for all $i < d$, where M_i^+ is the non-lazy upper walk. Let f be a function on $X(i)$, where $i < d$. We have

$$\langle M_i^+ f, f \rangle = \mathop{\mathbb{E}}_{t \sim \Pi_{i+1}} \mathop{\mathbb{E}}_{x \neq y \in t}[f(t \setminus \{x\}) f(t \setminus \{y\}].$$

Let $s = t \setminus \{x,y\}$. Since $t \sim \Pi_{i+1}$ and $x \neq y \in t$ are chosen at random, we have $s \sim \Pi_{i-1}$. Given such an s, the probability to get specific (t,x,y) is exactly $w_s(x,y)$ (the factor $1/2$ accounts for the relative order of x,y), and so

$$\langle M_i^+ f, f \rangle = \mathop{\mathbb{E}}_{s \sim \Pi_{i-1}} \mathop{\mathbb{E}}_{(x,y) \sim w_s}[f(s \cup \{x\}) f(s \cup \{y\})]. \tag{8}$$

In other words, we have shown that

$$\langle M_i^+ f, f \rangle = \mathop{\mathbb{E}}_{s \sim \Pi_{i-1}}[\langle A_s f_s, f_s \rangle], \tag{9}$$

where $f_s\colon X_s(0) \to \mathbb{R}$ is defined by

$$f_s(x) = f(s \cup \{x\}).$$

We now note that

$$\mathop{\mathbb{E}}_{x\sim w_s}[f(s\cup\{x\})] = (Df)(s).$$

Therefore we have, by (7), that

$$\left|\langle M_i^+ f, f\rangle - \langle UDf, f\rangle\right| = \Big|\mathop{\mathbb{E}}_{s\sim\Pi_{i-1}}\mathop{\mathbb{E}}_{(x,y)\sim w_s}[f(s\cup\{x\})f(s\cup\{y\})] - (Df)(s)^2\Big| \leq$$

$$\mathop{\mathbb{E}}_{s\sim\Pi_{i-1}}\left[\lambda(A_s)\mathop{\mathbb{E}}_{x\sim w_s}[f(s\cup\{x\})^2]\right].$$

If X is a γ-two-sided link expander then $\lambda(A_s) \leq \gamma$ for all s, and so

$$\left|\langle (M_i^+ - UD)f, f\rangle\right| \leq \gamma\|f\|^2.$$

Item 2. Assume now that X is a γ-HDX. Our goal is to show that for all $i < d-1$ and $r \in X(i)$,

$$\lambda(A_r) \leq 3(i+2)\gamma.$$

Using the convention that $X(-1)$ consists of the empty set, for $i = -1$ we have $A_\emptyset = M_0^+$, and so $U_{-1}D_0$ is zero on the space perpendicular to the constant function. Thus

$$\left\|M_0^+ - UD\right\| = \lambda(A_\emptyset),$$

and from our assumption $\lambda(A_\emptyset) \leq \gamma$.

Now assume $1 \leq i \leq d-1$, and fix some $r \in X(i-1)$. Let $f\colon X_r(0) \to \mathbb{R}$ be some eigenfunction of A_r, which is perpendicular to the constant function. In order to prove the theorem, we must show that

$$\left|\frac{\langle A_r f, f\rangle}{\langle f, f\rangle}\right| \leq 3(i+1)\gamma.$$

Define a function $\tilde{f} \in C^i$ by

$$\tilde{f}(s) = \begin{cases} f(s\setminus r) & \text{if } r \subset s,\\ 0 & \text{otherwise.}\end{cases}$$

Without loss of generality, we may assume that $\|\tilde{f}\| = 1$.

In order to obtain a bound on $\lambda(A_r)$, we bound $\langle \tilde{f}, \tilde{f}\rangle$, $\langle M_i^+\tilde{f}, \tilde{f}\rangle$, and $\langle UD\tilde{f}, \tilde{f}\rangle$ in terms of f and A_r.

Observe that the norms of f and $\tilde{f}$ are proportional:

$$\langle f, f\rangle = \frac{\langle \tilde{f}, \tilde{f}\rangle}{\Pi_{i-1}(r)(i+1)} = \frac{1}{\Pi_{i-1}(r)(i+1)}. \tag{10}$$

Furthermore, from what we showed in (9) we obtain that

$$\langle M_i^+\tilde{f}, \tilde{f}\rangle = \mathop{\mathbb{E}}_{r'\in X(i-1)}[\langle A_{r'}\tilde{f}_{r'}, \tilde{f}_{r'}\rangle],$$

where $\tilde{f}_{r'}(x) = \tilde{f}(r'\cup\{x\})$.

Fix some $r' \neq r$. If $\tilde{f}_{r'}(x) \neq 0$ then $\tilde{f}(r'\cup\{x\}) \neq 0$. In particular, this means that $r \subset r'\cup\{x\}$. Since both r, r' are contained in $r'\cup\{x\}$, this means that $r'\setminus r = \{x\}$. Thus there is at most one vertex $x \in X_{r'}(0)$ such that $\tilde{f}_{r'}(x) \neq 0$. Since $A_{r'}$ is a non-lazy operator, this implies that $\langle A_{r'}\tilde{f}_{r'}, \tilde{f}'_r\rangle = 0$. We remain with

$$\langle M_i^+\tilde{f}, \tilde{f}\rangle = \Pi_{i-1}(r)\langle A_r f, f\rangle. \tag{11}$$

In other words, the upper non-lazy random walk is proportional to the local adjacency operator.

We shall prove below the following claim, which shows that the lower walk scales $\tilde{f}$ by a factor of at most $\frac{i+1}{i}\gamma$:

Claim 5.6. *If $\tilde{f} \in C^i$ is perpendicular to constant functions then $|\langle U_{i-1}D_i\tilde{f},\tilde{f}\rangle| \le \frac{i+1}{i}\gamma$.*

Assuming the above:

$$\left|\frac{\langle A_r f, f\rangle}{\langle f, f\rangle}\right| = |(i+1)\Pi_{i-1}(r)\langle A_r f, f\rangle| = (i+1)|\langle M_i^+\tilde{f},\tilde{f}\rangle| \le$$

$$(i+1)|\langle (M_i^+ - U_{i-1}D_i)\tilde{f},\tilde{f}\rangle| + (i+1)|\langle U_{i-1}D_i\tilde{f},\tilde{f}\rangle| \le (i+1)\left(1+\frac{i+1}{i}\right)\gamma \le 3(i+1)\gamma,$$

where the equalities in the first line use (10) and (11), and the inequalities in the second line use Claim 5.6, our assumption that $\|M_i^+ - U_{i-1}D_i\| \le \gamma$, and the triangle inequality. □

We complete the proof of Theorem 5.5 by proving Claim 5.6:

Proof of Claim 5.6. Since UD is PSD, we have $\langle U_{i-1}D_i\tilde{f},\tilde{f}\rangle \ge 0$, and so we may remove the absolute value and prove

$$\langle U_{i-1}D_i\tilde{f},\tilde{f}\rangle \le \frac{i+1}{i}\gamma.$$

Consider the inner product $\langle D_i\tilde{f}, D_i\tilde{f}\rangle$ — this is the expectation upon choosing $r' \sim \Pi_{i-1}$, and then choosing two i-faces $s_1, s_2 \in X(i)$ containing it independently. Hence we decompose to the cases where $r' = r$ and $r' \ne r$:

$$\langle D_i\tilde{f}, D_i\tilde{f}\rangle = \mathop{\mathbb{E}}_{(r',s_1,s_2)}[f(s_1)f(s_2)] =$$

$$\Pi_{i-1}(r)\mathop{\mathbb{E}}_{(r',s_1,s_2)}[\tilde{f}(s_1)\tilde{f}(s_2)|r'=r] + (1-\Pi_{i-1}(r))\mathop{\mathbb{E}}_{(r',s_1,s_2)}[\tilde{f}(s_1)\tilde{f}(s_2)|r'\ne r]. \tag{12}$$

The first term is 0, since from independence of s_1, s_2:

$$\mathop{\mathbb{E}}_{(r',s_1,s_2)}[\tilde{f}(s_1)\tilde{f}(s_2)|r'=r] = \mathop{\mathbb{E}}_{s_1}[f(s_1)|r \subset s_1]^2 = 0,$$

since by assumption $\tilde{f}$ is perpendicular to constant functions.

We saw above that for any $r' \ne r$, there is at most one i-face containing r' (which is $s = r \cup r'$) such that $\tilde{f}(s) \ne 0$. For any $r' \ne r$, the value $\tilde{f}(s_1)\tilde{f}(s_2)$ is non-zero only when $s_1 = s_2 = r \cup r'$. For every $s_1 \in X(i)$, we define the event E_{s_1} to hold when $s_2 = s_1$. Then

$$(12) = (1-\Pi_{i-1}(r))\mathop{\mathbb{E}}_{s_1}\Big[\tilde{f}^2(s_1)\Pr_{r',s_2}[E_{s_1}|r'\ne r]\Big].$$

Note that if s_1 doesn't contain r then $\tilde{f}^2(s_1) = 0$, hence we continue taking expectation over all $s_1 \in X(i)$, even though some of them are unnecessary terms.

If we prove that for every $s \in X(i)$ we have $\Pr_{r',s_2}[E_{s_1}|r'\ne r] \le \frac{i+1}{i}\gamma$ then

$$(1-\Pi_{i-1}(r))\mathop{\mathbb{E}}_{s_1}[\tilde{f}(s_1)^2\Pr_{r',s_2}[E_{s_1}]] \le \frac{i+1}{i}\gamma\mathop{\mathbb{E}}_{s_1}[\tilde{f}^2(s_1)] = \frac{i+1}{i}\gamma\langle\tilde{f},\tilde{f}\rangle = \frac{i+1}{i}\gamma.$$

Thus we are left with proving the following statement: for all $s_1 \in X(i)$,

$$\Pr_{r',s_2}[E_{s_1}|r'\ne r] \le \frac{i+1}{i}\gamma.$$

We first bound the unconditioned probability $\Pr[E_{s_1}] = \Pr_{r',s_2\in X(i)}[s_2 = s_1|r' \subset s_1, s_2]$. Fix some $s_1 \in X(i)$, and let $\mathbb{1}_{s_1}: X(i) \to \mathbb{R}$ be its indicator. Notice that $U_{i-1}D_i\mathbb{1}_{s_1}(s_1) = \Pr_{r',s_2}[s_2 = s_1]$, and so

$$\langle U_{i-1}D_i\mathbb{1}_{s_1}, \mathbb{1}_{s_1}\rangle = \Pi_i(s_1)U_{i-1}D_i\mathbb{1}_{s_1}(s_1) = \Pi_i(s_1)\Pr[E_{S_1}].$$

We again use the non-laziness property of M_i^+ to assert that $\langle M_i^+\mathbb{1}_{s_1}, \mathbb{1}_{s_1}\rangle = 0$. Since X is a γ-HDX,

$$\langle U_{i-1}D_i\mathbb{1}_{s_1}, \mathbb{1}_{s_1}\rangle = \langle (U_{i-1}D_i - M_i^+)\mathbb{1}_{s_1}, \mathbb{1}_{s_1}\rangle \le \left\|U_{i-1}D_i - M_i^+\right\|\|\mathbb{1}_{s_1}\|^2 = \gamma\Pi_i(s_1).$$

Hence $\Pr[E_{s_1}] \le \gamma$.

Consider now any $s_1 \in X(i)$ containing r, and let r' be a random $(i-1)$-face contained in s_1. The probability that $r' \ne r$ is $\frac{i}{i+1}$, and so

$$\Pr_{r',s_2}[E_{s_1}|r'\ne r] \le \frac{i+1}{i}\Pr_{r',s_2}[E_{s_1}] \le \frac{i+1}{i}\gamma.$$ □

6 Boolean degree 1 functions

In this section we characterize all Boolean degree 1 functions in nice complexes.

Definition 6.1. *Let X be a simplicial complex. The* 1-skeleton *of X is the graph whose vertices are the 0-faces of X and whose edges are the 1-faces of X.*

Theorem 6.2. *Suppose that X is a proper k-dimensional simplicial complex, for $k \geq 2$, whose 1-skeleton is connected. A function $f \in C^k$ is a Boolean degree 1 function if and only if there exists an independent set I such that f is the indicator of intersecting I or of not intersecting I.*

Proof. If f is the indicator of intersecting an independent set I then $f = \sum_{v \in I} y_v$, and so $\deg f \leq 1$. If f is the indicator of not intersecting an independent set I then $f = \sum_{v \in X(0)} y_v/(k+1) - \sum_{v \in I} y_v$, and so again $\deg f \leq 1$.

Suppose now that f is a Boolean degree 1 function. If $|X(0)| \leq 2$ then the theorem clearly holds, so assume that $|X(0)| > 2$. Lemma 3.5 shows that f has a unique representation of the form

$$f = \sum_{v \in X(0)} c_v y_v.$$

Since f is Boolean, it satisfies $f^2 = f$. Note that

$$f^2 = \sum_{\{u,v\} \in X(1)} 2c_u c_v y_{\{u,v\}} + \sum_{v \in X(0)} c_v^2 y_v.$$

Moreover, since every input x to f which contains v contains exactly k other points (elements of $X(0)$), and since $X(1)$ contains all pairs of points from x, we have

$$y_v = \sum_{u \colon \{u,v\} \in X(1)} \frac{y_{\{u,v\}}}{k}.$$

This shows that

$$0 = f^2 - f = \sum_{\{u,v\} \in X(1)} 2c_u c_v y_{\{u,v\}} + \frac{1}{k} \sum_{v \in X(0)} (c_v^2 - c_v) \sum_{u \colon \{u,v\} \in X(1)} y_{\{u,v\}} = \\ \frac{1}{k} \sum_{\{u,v\} \in X(1)} (2kc_u c_v + c_u^2 - c_u + c_v^2 - c_v) y_{\{u,v\}}.$$

Lemma 3.5 shows that the coefficients of all $y_{\{u,v\}}$ must vanish, that is, for all $\{u,v\} \in X(1)$ we have

$$2kc_u c_v = c_u(1 - c_u) + c_v(1 - c_v).$$

Consider now a triple of points u, v, w such that $\{u, v, w\} \in X(2)$, and the corresponding system of equations:

$$\begin{aligned} 2kc_u c_v &= c_u(1 - c_u) + c_v(1 - c_v), \\ 2kc_u c_w &= c_u(1 - c_u) + c_w(1 - c_w), \\ 2kc_v c_w &= c_v(1 - c_v) + c_w(1 - c_w). \end{aligned}$$

Subtracting the second equation from the first, we obtain

$$2kc_u(c_v - c_w) = c_v(1 - c_v) - c_w(1 - c_w) = (c_v - c_w) - (c_v^2 - c_w^2) = (c_v - c_w)(1 - c_v - c_w).$$

This shows that either $c_v = c_w$ or $2kc_u = 1 - c_v - c_w$.

If $c_u \neq c_v, c_w$ then $2kc_w + c_u + c_v = 2kc_v + c_u + c_w = 1$, which implies that $c_v = c_w$. Thus c_u, c_v, c_w can consist of at most two values. If $c := c_u = c_v = c_w$ then $2kc^2 = 2c(1 - c)$, and so $c \in \{0, \frac{1}{k+1}\}$. If $c := c_v = c_w \neq c_u$ then $2kc^2 = 2c(1 - c)$, and so $c \in \{0, \frac{1}{k+1}\}$ as before. We also have $2kc_u c = c_u(1 - c_u) + c(1 - c)$. If $c = 0$ then this shows that $c_u(1 - c_u) = 0$, and so $c_u = 1$. If $c = \frac{1}{k+1}$ then one can similarly check that $c_u = \frac{1}{k+1} - 1$.

Summarizing, one of the following two cases must happen:

1. Two of c_u, c_v, c_w are equal to 0, and the remaining one is either 0 or 1.

2. Two of c_u, c_v, c_w are equal to $\frac{1}{k+1}$, and the remaining one is either $\frac{1}{k+1}$ or $\frac{1}{k+1} - 1$.

Let us say that a vertex $v \in X(0)$ is of type A if $c_v \in \{0, 1\}$, and of type B if $c_v \in \{\frac{1}{k+1}, \frac{1}{k+1} - 1\}$. Since the complex is pure and at least two-dimensional, every vertex must participate in a triangle (two-dimensional face), and so every vertex is of one of the types. In fact, all vertices must be of the *same* type. Otherwise, there would be a vertex v of type A incident to a vertex w of type B (since the link of $\emptyset$ is connected). However, since the complex is pure, $\{v, w\}$ must participate in a triangle, contradicting the classification above.

Suppose first that all vertices are type A, and let $I = \{v : c_v = 1\}$. Note that f indicates that the input face intersects I. Clearly I must be an independent set, since otherwise f would not be Boolean. When all vertices are type B, the function $1 - f = \sum_{v \in X(0)} (\frac{1}{k+1} - c_v) y_v$ is of type A, and so f must indicate not intersecting an independent set. $\square$

If X is a γ-HDX for $0 < \gamma < 1/(k+1)$ then the link of $\emptyset$ has positive spectral gap, and in particular it is connected. Thus Theorem 6.2 applies to high-dimensional expanders.

When the 1-skeleton of X contains r connected components $C_1, \ldots, C_r$, the same argument shows that the Boolean degree 1 functions on X are of the form $f = f_1 + \cdots + f_r$, where each f_i is the indicator of intersecting or not intersecting an independent set of C_i.

7 FKN theorem on high dimensional expanders

In this section, we prove an analog of the classical result of Friedgut, Kalai and Naor [FKN02] for high-dimensional expanders. The FKN theorem states that any Boolean function F on the hypercube that is close to a degree 1 function f (not necessarily Boolean) in the L_2^2-sense must agree with some Boolean degree 1 function (which must be a dictator) on most points. This result for the Boolean hypercube can be easily extended to functions on k-slices of the hypercube, provided $k = \Theta(n)$.

Theorem 7.1 (FKN theorem on the slice [Fil16a])**.** *Let $n, k \in \mathbb{Z}_{\geq 0}$ and $\varepsilon \in (0, 1)$ such that $n/4 \leq k + 1 \leq n/2$. Let $F \colon \binom{[n]}{k+1} \to \{0, 1\}$ be a Boolean function such that $\mathbb{E}[(F - f)^2] < \varepsilon$ for some degree 1 function $f \colon \binom{[n]}{k+1} \to \{0, 1\}$. Then there exists a degree 1 function $g \colon \binom{[n]}{k+1} \to \mathbb{R}$ such that*

$$\Pr[F \neq g] = O(\varepsilon).$$

Furthermore, $g \in \{0, 1, y_i, 1 - y_i\}$, that is, g is a Boolean dictator (1-junta).

Remark 7.2.

1. The function g promised by the theorem satisfies $\mathbb{E}[(g - F)^2] = \Pr[g \neq F] = O(\varepsilon)$ and hence, by the L_2^2-triangle inequality we have $\mathbb{E}[(f - g)^2] \leq 2\,\mathbb{E}[(f - F)^2] + 2\,\mathbb{E}[(g - F)^2] = O(\varepsilon)$. This is the way that the FKN theorem is traditionally stated, but we prefer the above formulation as this is the one we are able to generalize to the high-dimensional expander setting.

2. The function 1 can also be written as $\frac{1}{k+1} \sum_j y_j$. The function $1 - y_i$ can also be written as $\frac{1}{k+1} \sum_{j \neq i} y_j + (\frac{1}{k+1} - 1) y_i$.

3. The result of Filmus [Fil16a] is quite a bit stronger: for every $k \leq n/2$, it promises the existence of a function $g \colon \binom{[n]}{k+1} \to \mathbb{R}$, not necessarily Boolean, such that $\mathbb{E}[(f - g)^2] = O(\varepsilon)$. Moreover, either g or $1 - g$ is of the form $\sum_{i \in S} y_i$ for $|S| \leq \max(1, \sqrt{\varepsilon} \cdot n/k)$. The bound on the size of S ensures that $\Pr[g \in \{0, 1\}] = 1 - O(\varepsilon)$.

Our main theorem is an extension of the above theorem to k-faces of a two-sided link expander.

Theorem 7.3 (FKN theorem for two-sided link expanders)**.** *Let X be a d-dimensional λ-two-sided link expander, where $\lambda < 1/d$, and let $4k^2 < d$. Let $F \colon X(k) \to \{0, 1\}$ be a function such that $\mathbb{E}[(F - f)^2] < \varepsilon$ for some degree 1 function $f \colon X(k) \to \mathbb{R}$. Then there exists a degree 1 function $g \colon X(k) \to \mathbb{R}$ such that*

$$\Pr[F \neq g] = O_\lambda(\varepsilon).$$

Furthermore, the degree 1 function g can be written as $g(y) = \sum_i d_i y_i$, where $d_i \in \{0, 1, \frac{1}{k+1}, \frac{1}{k+1} - 1\}$.

The high-dimensional analog of the FKN theorem is obtained from the FKN theorem for the slice using the agreement theorem of Dinur and Kaufman [DK17].

Using Theorem 5.5, we formulate the FKN theorem in terms of high dimensional expanders:

Corollary 7.4 (FKN theorem for HDX)**.** *Let X be a d-dimensional γ-high-dimensional expander, where $\gamma < 1/3d^2$, and let $4k^2 < d$. Let $F\colon X(k) \to \{0,1\}$ be a function such that $\mathbb{E}[(F-f)^2] < \varepsilon$ for some degree 1 function $f\colon X(k) \to \mathbb{R}$. Then there exists a degree 1 function $g\colon X(k) \to \mathbb{R}$ such that*

$$\Pr[F \neq g] = O_\gamma(\varepsilon).$$

Furthermore, the degree 1 function g can be written as $g(y) = \sum_i d_i y_i$, where $d_i \in \{0, 1, \frac{1}{k+1}, \frac{1}{k+1} - 1\}$.

7.1 Agreement theorem for high dimensional expanders

Dinur and Kaufman [DK17] prove an agreement theorem for high-dimensional expanders. The setup is as follows. For each k-face s we are given a local function $f_s\colon s \to \Sigma$ that assigns values from an alphabet Σ to each point in s. Two local functions $f_s, f_{s'}$ are said to *agree* if $f_s(v) = f_{s'}(v)$ for all $v \in s \cap s'$. Let $\mathcal{D}_{k,2k}$ be the distribution on pairs (s_1, s_2) obtained by choosing a random $t \sim \Pi_{2k}$ and then independently choosing two k-faces $s_1, s_2 \subset t$. The theorem says that if a random pair of faces $(s, s') \sim \mathcal{D}_{k,2k}$ satisfies with high probability that f_s agrees with $f_{s'}$ on the intersection of their domains, then there must be a global function $g\colon X(0) \to \Sigma$ such that almost always $g|_s \equiv f_s$. Formally:

Theorem 7.5 (Agreement theorem for high-dimensional expanders [DK17])**.** *Let X be a d-dimensional λ-two-sided high-dimensional expander, where $\lambda < 1/d$, let $k^2 < d$, and let Σ be some fixed finite alphabet. Let $\{f_s\colon s \to \Sigma\}_{s \in X(k)}$ be an ensemble of local functions on $X(k)$, i.e. $f_s \in \Sigma^s$ for each $s \in X(k)$. If*

$$\Pr_{(s_1,s_2)\sim\mathcal{D}_{k,2k}}[f_{s_1}|_{s_1\cap s_2} \equiv f_{s_2}|_{s_1 \cap s_2}] > 1 - \varepsilon$$

then there is a $g\colon X(0) \to \Sigma$ such that

$$\Pr_{s\sim\Pi_k}[f_s \equiv g|_s] \geq 1 - O_\lambda(\varepsilon).$$

While Dinur and Kaufman state the theorem for a binary alphabet, the general version follows in a black box fashion by applying the theorem for binary alphabets $\lceil \log_2 |\Sigma| \rceil$ many times.

7.2 Proof of Theorem 7.3

Let $f, F \in C^k$, where F is a Boolean function and f is a degree 1 function, as in the hypothesis of Theorem 7.3. Since f is a degree 1 function, Lemma 3.5 guarantees that there exist $a_i \in \mathbb{R}$ such that $f(y) = \sum_{i \in X(0)} a_i y_i$. Note that here we view the inputs of f as $|X(0)|$-bit strings with exactly $k+1$ ones, the rest being zero.

We begin by defining two ensembles of pairs of local functions $\{(f|_t, F|_t)\}_{t\in X(2k)}$, $\{(f|_u, F|_u)\}_{u \in X(4k)}$, which are the restrictions of (f, F) to the $2k$-face t and $4k$-face u. Formally, for any $t \in X(2k)$ and $u \in X(4k)$, consider the restriction of f to t and u defined as follows:

$$f|_t, F|_t\colon \binom{t}{k} \to \mathbb{R}, \qquad f|_t(y) = f(y) = \sum_{i\in t} a_i y_i, \qquad F|_t(y) = F(y),$$
$$f|_u, F|_u\colon \binom{u}{k} \to \mathbb{R}, \qquad f|_u(y) = f(y) = \sum_{i\in u} a_i y_i, \qquad F|_u(y) = F(y).$$

Observe that the $f|_t$'s are degree 1 functions, while the $F|_t$'s are Boolean functions (similarly for $f|_u$'s and $F|_u$'s).

Now, define the following quantities:

$$\varepsilon_t := \mathop{\mathbb{E}}_{s\colon s\subset t}[(f|_t(s) - F|_t(s))^2], \qquad \delta_u := \mathop{\mathbb{E}}_{s\colon s\subset u}[(f|_u(s) - F|_u(s))^2].$$

Clearly, $\mathbb{E}_t[\varepsilon_t] = \mathbb{E}_u[\delta_u] = \varepsilon$.

Let $\alpha_k = \frac{1}{k+1}$. Applying Theorem 7.1 (along with Remark 7.2) to the functions $(f|_t, F|_t)$ for each $t \in X(2k)$, we have the following claim:

Claim 7.6. *For every $t \in X(2k)$, there exists a Boolean dictator $g_t \colon \binom{t}{k} \to \{0,1\}$ such that*

$$\mathop{\mathbb{E}}_{s\colon s\subset t}[(f|_t - g_t)^2] = O(\varepsilon_t).$$

Furthermore, there exists a function $d_t \colon t \to \{0,1,\alpha_k,\alpha_k-1\}$ such that $g_t(y) = \sum_{i\in t} d_t(i) y_i$.

A similar claim holds for each $u \in X(4k)$:

Claim 7.7. *For every $u \in X(4k)$, there exists a Boolean dictator $h_u \colon \binom{u}{k} \to \{0,1\}$ such that*

$$\mathop{\mathbb{E}}_{s\colon s\subset u}[(f|_u - h_u)^2] = O(\delta_u).$$

Furthermore, there exists a function $e_u \colon u \to \{0,1,\alpha_k,\alpha_k-1\}$ such that $h_u(y) = \sum_{i\in u} e_u(i) y_i$.

We will now prove that functions in the collection of local functions $\{d_t\}_t$ typically agree with each other. We will then be able to use the agreement theorem, Theorem 7.5, to sew these different local functions together, yielding a single function $d\colon X(0) \to \{0,1,\alpha_k,\alpha_k-1\}$. This d will determine a global degree 1 function g defined as follows: $g(y) = \sum_{i\in X(0)} d(i) y_i$.

Claim 7.8. *There exists a function $d\colon X(0) \to \{0,1,\alpha_k,\alpha_k-1\}$ such that $\Pr_t[d_t \equiv d|_t] = 1 - O_\lambda(\varepsilon)$.*

Proof. To sew the various d_t together via the agreement theorem, we would like to first bound the probability

$$\Pr_{(t_1,t_2)\sim\mathcal{D}_{2k,4k}}[d_{t_1}|_{t_1\cap t_2} \not\equiv d_{t_2}|_{t_1\cap t_2}]\,.$$

Recall the definition of the distribution $\mathcal{D}_{2k,4k}$: we first pick a set $u \in X(4k)$ according to Π_{4k} and then two $2k$-faces t_1, t_2 of u uniformly and independently. Consider the three functions d_{t_1}, d_{t_2} and e_u. Clearly, if $d_{t_1}|_{t_1\cap t_2} \not\equiv d_{t_2}|_{t_1\cap t_2}$ then one of $e_u|_{t_1} \not\equiv d_{t_1}$ or $e_u|_{t_2} \not\equiv d_{t_2}$ must hold. Thus,

$$\Pr_{(t_1,t_2)\sim\mathcal{D}_{2k,4k}}[d_{t_1}|_{t_1\cap t_2} \not\equiv d_{t_2}|_{t_1\cap t_2}] \le 2\cdot \Pr_{t,u}[e_u|_t \not\equiv d_t]\,. \tag{13}$$

Thus, it suffices to bound the probability $\Pr_{t,u}[e_u|_t \not\equiv d_t]$, where $u \sim \Pi_{4k}$ and t is a random $2k$-face of u.

For any fixed $t \subset u$, the L_2^2 triangle inequality shows that

$$\mathbb{E}[(h_u|_t - g_t)^2] \le 2\,\mathbb{E}[(h_u|_t - f|_t)^2] + 2\,\mathbb{E}[(f|_t - g_t)^2] = 2\,\mathbb{E}[(h_u|_t - f|_t)^2] + O(\varepsilon_t).$$

Taking expectation over $t \in X(2k)$ conditioned on $t \subset u$, we see that

$$\mathop{\mathbb{E}}_{t\subset u}\mathbb{E}[(h_u|_t - g_t)^2] \le 2\,\mathbb{E}[(h_u - f|_u)^2] + O\left(\mathop{\mathbb{E}}_{t\colon t\subset u}\varepsilon_t\right) = O(\delta_u) + O\left(\mathop{\mathbb{E}}_{t\colon t\subset u}\varepsilon_t\right).$$

Taking expectation over $u \sim \Pi_{4k}$, we now have

$$\mathop{\mathbb{E}}_{u}\mathop{\mathbb{E}}_{t\subset u}\mathbb{E}[(h_u|_t - g_t)^2] = O(\varepsilon).$$

For any fixed $t \subset u$, both $h_u|_t$ and g_t are Boolean dictators. Hence either they agree, or $\mathbb{E}[(h_u|_t - g_t)^2] = \Omega(1)$. This shows that $h_u|_t$ disagrees with g_t with probability $O(\varepsilon)$, and so

$$\Pr_{t,u}[e_{u|_t} \not\equiv d_t] = O(\varepsilon).$$

We now return to (13), concluding that

$$\mathop{\mathbb{E}}_{(t_1,t_2)\sim\mathcal{D}_{2k,4k}}[d_{t_1}|_{t_1\cap t_2} \not\equiv d_{t_2}|_{t_1\cap t_2})] = O(\varepsilon).$$

We have thus satisfied the hypothesis of the agreement theorem (Theorem 7.5). Invoking the agreement theorem, we deduce that $\Pr_{t\sim\Pi_{2k}}[d_t \equiv d|_t] = 1 - O_\lambda(\varepsilon)$. □

The d's guaranteed by Claim 7.8 naturally correspond to a degree 1 function $g\colon X(k) \to \mathbb{R}$ as follows:

$$g(y) := \sum_{i \in X(0)} d(i) y_i.$$

We now show that this g is mostly Boolean.

Claim 7.9. $\Pr_s[g(s) \in \{0,1\}] = 1 - O_\lambda(\varepsilon)$.

Proof. Since g_t is Boolean-valued,

$$\Pr_{s \sim \Pi_k}[g(s) \in \{0,1\}] \geq \Pr_t[g|_t = g_t] = \Pr_t[d|_t \equiv d_t] = 1 - O_\lambda(\varepsilon). \qquad \square$$

We now show that g in fact agrees pointwise with F most of the time.

Claim 7.10. $\Pr_s[g \neq F] = O_\lambda(\varepsilon)$.

Proof. Fix any $t \in X(2k)$. We compute $\Pr_{s\colon s \subset t}[F|_t \neq g_t]$ as follows

$$\begin{aligned}
\Pr[F|_t \neq g_t] &= \|F|_t - g_t\|^2 && [\text{ Since } F|_t \text{ and } g_t \text{ are both Boolean }] \\
&\leq 2 \cdot \|F|_t - f|_t\|^2 + 2 \cdot \|f|_t - g_t\|^2 \\
&= O(\varepsilon_t) + O(\varepsilon_t) = O(\varepsilon_t).
\end{aligned}$$

We can now compute $\Pr_s[F \neq g]$ as follows:

$$\Pr[F \neq g] = \mathop{\mathbb{E}}_t \Pr[F|_t \neq g|_t] \leq \mathop{\mathbb{E}}_t \Pr[F|_t \neq g_t] + \Pr_t[g|_t \neq g_t] = O(\varepsilon) + \Pr_t[d|_t \not\equiv d_t] = O_\lambda(\varepsilon). \qquad \square$$

This completes the proof of Theorem 7.3.

8 Expanding posets (eposets)

In this section, we describe a setting generalizing simplicial complexes, namely *measured posets*. These are partially ordered sets (a set X with a partial order $\leq$ on it) whose elements are partitioned into levels $X(j)$, and that have some additional properties stated below. As in simplicial complexes, we can define C^j as the space of real-valued functions on $X(j)$, and averaging operators $U_j\colon C^j \to C^{j+1}$ and $D_{j+1}\colon C^{j+1} \to C^j$.

We shall generalize the notion of a γ-HDX to a γ-*expanding poset* (eposet) — a measured poset with operators D_j, U_j such that

$$\|D_{j+1}U_j - r_j I - \delta_j U_{j-1} D_j\| \leq \gamma,$$

for $\gamma < 1$, all non-extreme levels j of the poset, and some constants r_j, δ_j.

We begin the section by discussing the formal notion of an eposet. Afterwards, we shall generalize Theorem 4.6 to all eposets, and prove it in the general setting. Finally, we will show that if our measured poset is a simplicial complex, then $r_j \approx \frac{1}{j+2}, \delta_j \approx 1 - \frac{1}{j+2}$, under the assumption that the laziness of the lower walk is small.

8.1 Measured posets

A *graded* (or *ranked*) *poset* is a partially ordered set (poset) $(X, \leq)$ equipped with a *rank* function $\rho\colon X \to \mathbb{N} \cup \{-1\}$ such that:

1. For all $x, y \in X$, if $x \leq y$ then $\rho(x) \leq \rho(y)$.
2. For every $x, y \in X$, if y is minimal with respect to elements greater than x (i.e. y *covers* x), then $\rho(y) = \rho(x) + 1$.

We denote the set of elements of rank j by $X(j)$.

For example, any simplicial complex is a graded poset, if we take $\leq$ to be the containment relation.

We say that a graded poset is *d-dimensional* if the maximal rank of an element in X is d. We say that a d-dimensional graded poset is *pure* if all maximal elements are of rank d, that is, for every $t \in X$ there exists $s \in X(d)$ such that $t \leq s$.

Every simplicial complex is a graded poset. Another useful example to keep in mind is the *Grassmann poset* $\mathrm{Gr}_q(n,d)$, whose elements are subspaces of dimension at most $d+1$ of $\mathbb{F}_q^n$, and the order is by containment. The rank function for the Grassmann poset is $\rho(U) = \dim(U) - 1$, and so $X(j) = \{U \subseteq \mathbb{F}_q^n : \dim(U) = j+1\}$.

Definition 8.1 (Measured poset)**.** *Let X be a finite graded pure d-dimensional poset, with a unique minimum element of rank -1. We say that X is* measured *by a distribution $\vec{\Pi} = (\Pi_d, \Pi_{d-1}, \dots, \Pi_{-1})$ if it satisfies the following properties:*

1. *$\Pi_i \in X(i)$ for all i.*
2. *$\Pi_{i-1} \subset \Pi_i$ for all $i > -1$.*
3. *The sequence $\Pi_d, \dots, \Pi_{-1}$ has the Markov property: Π_{i-1} depends only on Π_i for all $i > -1$.*

We denote the real-valued function spaces on $X(j)$ by C^j. We denote the averaging operators of the steps in the Markov process by $U_j \colon C^j \to C^{j+1}, D_{j+1} \colon C^{j+1} \to C^j$.

The process we defined for the measures $\vec{\Pi}$ in a simplicial complex is an example of a measured poset. For the Grassmann poset mentioned above, we also have a similar probabilistic experiment:

1. Choose a subspace of dimension $d+1$, $s_d \in X(d)$, uniformly at random.
2. Given a subspace s_i of dimension $i+1$, choose $s_{i-1} \in X(i-1)$ to be a uniformly random codimension 1 subspace of s_i.

An analogue for Theorem 3.2 holds for any measured poset. We say that a k-dimensional measured poset X is *proper* if for all $j \leq k-1$, $\ker U_j = \{0\}$. Also, as before we denote

$$H^{-1} = C^{-1}, \quad H^i = \ker D_i, \quad V^i = U^{k-i} H^i.$$

Theorem 8.2. *If X is a proper k-dimensional measured poset then we have the following decomposition of C^k:*

$$C^k = V^k \oplus V^{k-1} \oplus \dots \oplus V^{-1}.$$

In other words, for every function $f \in C^k$ there is a unique choice of $h_i \in H^i$ such that the functions $f_i = U^{k-i} h_i$ satisfy $f = f_{-1} + f_0 + \ldots + f_k$.

The proof for this is identical to the proof of Theorem 3.2 and is left out.

Sequentially differential posets were first defined and studied (in a slightly different form) by Stanley in [Sta88, Sta90].

Definition 8.3 (Sequentially differential posets)**.** Sequentially differential posets *are measured posets whose averaging operators U, D satisfy an equation*

$$D_{j+1} U_j - \delta_j U_{j-1} D_j - r_j I = 0, \tag{14}$$

for some $r_j, \delta_j \in \mathbb{R}_{\geq 0}$ and all $0 \leq j \leq k-1$.

For example, the complete complex satisfies this definition with parameters

$$\delta_i = \left(1 - \frac{1}{i+2}\right)\left(1 - \frac{1}{n-i}\right)^{-1} \text{ and } r_i = 1 - \delta_i.$$

In other words,

$$DU - \left(1 - \frac{1}{i+2}\right)\left(1 - \frac{1}{n-i}\right)^{-1} UD - \left(1 - \left(1 - \frac{1}{i+2}\right)\left(1 - \frac{1}{n-i}\right)^{-1}\right) I = 0.$$

The Grassmann poset $\mathrm{Gr}_q(n,d)$ is also a sequentially differential poset with

$$\delta_i = 1 - \left(1 - \frac{q-1}{q^{i+2}-1}\right)\left(1 - \frac{q-1}{q^{n-i}-1}\right)^{-1} \text{ and } r_i = 1 - \delta_i. \tag{15}$$

To see this, observe the following claim that the reader can verify by direct calculation:

Claim 8.4. *Let X be a measured poset, and suppose we can decompose:*

$$D_{i+1}U_i = \alpha_i I + (1-\alpha_i)M_i,$$

$$U_{i-1}D_i = \beta_i I + (1-\beta_i)M_i,$$

where $0 \le \alpha_i, \beta_i \le 1$ are constants and M_i is some operator. Then

$$D_{i+1}U_i - r_i I - \delta_i U_{i-1}D_i = 0,$$

where

$$\delta_i = (1-\alpha_i)(1-\beta_i)^{-1} \textit{ and } r_i = 1 - \delta_i.$$

In both the complete complex and the Grassmann poset $\mathrm{Gr}_q(n,d)_q$, the non-lazy upper walk and the non-lazy lower walk are *the same* — given $t_1 \in X(i)$, our choice for $t_2 \in X(i)$ is a set (or subspace in the Grassmann case) that shares an intersection of size (resp. dimension) i with t_1 (with uniform probability). The only difference between DU and UD is the probability to stay in place. Thus we can decompose:

$$D_{i+1}U_i = \alpha_i I + (1-\alpha_i)M_i,$$

$$U_{i-1}D_i = \beta_i I + (1-\beta_i)M_i,$$

where M^i is the non-lazy upper (or lower) random walk. In the simplicial complex case

$$\alpha_i = \frac{1}{i+2},\ \beta_i = \frac{1}{n-i},$$

and in the Grassmann case

$$\alpha_i = \frac{q-1}{q^{i+2}-1},\ \beta_i = \frac{q-1}{q^{n-i}-1}.$$

We relax Definition 8.3 to an almost sequentially differential poset — a measured poset that approximately satisfies such an identity:

Definition 8.5 (Expanding Poset)**.** *Let $\vec{r}, \vec{\delta} \in \mathbb{R}^k_{\ge 0}$, and let $\gamma < 1$. We say that X is an $(\vec{r}, \vec{\delta}, \gamma)$-*expanding poset *(or $(\vec{r}, \vec{\delta}, \gamma)$-*eposet*) if for all $j \le k-1$:*

$$\|D_{j+1}U_j - r_j I - \delta_j U_{j-1}D_j\| \le \gamma. \tag{16}$$

A sequentially differential poset is an eposet with $\gamma = 0$. As we saw in (3), a γ-HDX is an $(\vec{r}, \vec{\delta}, \gamma)$-eposet, where $r_j = \frac{1}{j+2}$ and $\delta_j = 1 - \frac{1}{j+2}$.

We can use (15) to assert that $\mathrm{Gr}_q(n,d)$ is an $(\vec{r}, \vec{\delta}, \gamma)$-eposet for $r_i = \frac{q-1}{q^{i+2}-1}, \delta_i = 1 - r_i$, and $\gamma = O(1/q^{n-d})$. While this only gives an *eposet* (even though it is truly sequentially differential), the parameters are much simpler, thus calculations regarding the random walks are easier.

8.2 Almost orthogonality of decomposition

In this section we show that in an eposet, the spaces V_i are almost orthogonal to one another. Moreover, we show that these spaces are "almost eigenspaces" of the operator DU.

Theorem 8.6. *Let X be a k-dimensional $(\vec{r}, \vec{\delta}, \gamma)$-eposet. For every function f on C^ℓ for $\ell \leq k$, the decomposition $f = f_{-1} + \cdots + f_\ell$ of Theorem 3.2 satisfies the following properties, when γ is small enough (as a function of k and the eposet parameters):*

- *For $i \neq j$, $|\langle f_i, f_j \rangle| = O(\gamma)\|f_i\|\|f_j\|$.*
- *$\|f\|^2 = (1 \pm O(\gamma))(\|f_{-1}\|^2 + \cdots + \|f_\ell\|^2)$, and for all i, $\|f\|^2 = (1 \pm O(\gamma))(\|f_{\leq i}\|^2 + \|f_{>i}\|^2)$.*
- *If $\ell < k$, the f_i (for $i \geq 0$) are approximate eigenvalues of DU: $\|DUf_i - r^\ell_{\ell-i+1} f_i\| = O(\gamma)\|f_i\|$, where*

$$r^\ell_i = r_\ell + \sum_{j=\ell-i}^{\ell-1} \left(\prod_{t=j+1}^{\ell} \delta_t \right) r_j. \tag{17}$$

  (Note $DUf_{-1} = r^\ell_{\ell+2} f_{-1}$, where $r^\ell_{\ell+2} = 1$.)

- *If $\ell < k$ then $\langle DUf, f \rangle = \sum_{i=-1}^{\ell} r^\ell_{\ell-i+1}\|f_i\|^2 \pm O(\gamma)\|f\|^2$.*

The hidden constant in the O notations depends on k and the eposet parameters only (but not on $|X|$). In particular, the last item implies that if $\vec{r} > 0$ then for a small enough γ, the poset is proper.

In a measured poset, the decomposition of Theorem 8.2 is not necessarily orthogonal. However, this theorem shows that for an eposet, the decomposition is *almost* orthogonal.

Remark 8.7. In the special case of a *sequentially differential poset*, i.e. $\gamma = 0$, we *do* get that the decomposition in Theorem 8.2 is orthogonal, and that the decomposition $C^\ell = V_{-1} \oplus \cdots \oplus V_\ell$ is a decomposition to eigenspaces of DU: for all $f_i \in V_i$,

$$DUf_i = r^\ell_i f_i,$$

for the r^ℓ_i given in (17).

Recall our convention that for $f \in C^{\ell-j}$,

$$U^j f = U_{\ell-1} \cdots U_{\ell-j+1} U_{\ell-j} f \in C^\ell.$$

We start with a technical claim that generalizes the approximate relation between D and U, namely

$$\|DU - rI - \delta UD\| = O(\gamma),$$

to an approximate relation between D and U^j:

$$\|DU^j - rU^{j-1} - \delta U^j D\| = O(\gamma),$$

for appropriate constants $r, \delta \in \mathbb{R}$.

Claim 8.8. *Let X be a k-dimensional $(\vec{r}, \vec{\delta}, \gamma)$-eposet, $1 \leq j \leq \ell+1 \leq k$, and $DU^j : X(\ell-(j-1)) \to X(\ell)$. There exist constants r^ℓ_j, δ^ℓ_j such that*

$$\left\|DU^j - r^\ell_j U^{j-1} - \delta^\ell_j U^j D\right\| = O(\gamma), \tag{18}$$

where the hidden constant depends on $k, \vec{\delta}, \vec{r}$ only.

Furthermore, δ^ℓ_j and r^ℓ_j are given by the following formulas: $\delta^\ell_0 = 1$ and

$$\delta^\ell_j = \prod_{t=\ell-(j-1)}^{\ell} \delta_t, \qquad r^\ell_j = \sum_{t=0}^{j-1} r_{\ell-t}\delta^\ell_t.$$

While this claim seems technical, the proof consists of simply inductively substituting DU with $rI + UD$ in the terms, until the formula is obtained. The proof is given in more detail at the end of this section.

Regarding the constants r, δ, notice the following:

1. $r_1^\ell = r_\ell$ and $\delta_1^\ell = \delta_\ell$.
2. If for all $0 \le j \le \ell$, $r_j + \delta_j = 1$, then for all $0 \le j \le \ell$, $r_j^\ell + \delta_j^\ell = 1$. In this case, we have a better formula for r_j^ℓ:
$$r_j^\ell = 1 - \prod_{t=\ell-(j-1)}^{\ell} \delta_t.$$
3. In a γ-HDX, we get $r_j^\ell = \frac{j}{\ell+2}$ and $\delta_j^\ell = 1 - \frac{j}{\ell+2}$.

This claim directly implies the third item in Theorem 8.6. In other words, the decomposition in Theorem 8.2 is a decomposition of "approximate eigenspaces" of UD:

Corollary 8.9. *Let X be an $(\vec{r}, \vec{\delta}, \gamma)$-eposet, and let $h \in H^{\ell-j}$. Then $U^j h \in V^{\ell-j}$ is an approximate eigenvector of $D_{\ell+1}U_\ell$ with eigenvalue r_{j+1}^ℓ:*
$$\|DU(U^j h) - r_{j+1}^\ell (U^j h)| = O(\gamma)\|h\|.$$

We proceed by showing that these approximate eigenspaces V^j are approximately orthogonal.

Lemma 8.10. *Suppose that X is a k-dimensional $(\vec{r}, \vec{\delta}, \gamma)$-eposet, let $\ell < k$, let $i \neq j$, and let $f_i = U^{\ell-i}h_i, f_j = U^{\ell-i}h_j$ for $h_i \in H^i, h_j \in H^j$, as in Theorem 3.2. Then*
$$\langle f_i, f_j\rangle = O(\gamma)\|h_i\|\|h_j\|,$$
where the hidden constant depends on $k, \vec{\delta}, \vec{r}$ only.

Proof. Recall that $h_i \in H^i = \ker D_i$. Given this, it is easy to see that h_ℓ is orthogonal to $f_{\ell-j}$, for any $j \ge 1$. Indeed, $\langle f_\ell, f_{\ell-j}\rangle = \langle h_\ell, U^j h_{\ell-j}\rangle = \langle D^j h_\ell, h_{\ell-j}\rangle = 0$, since $Dh_\ell = 0$.

To prove the statement in general we use Claim 8.8 and induction on ℓ. The base case where $\ell = 0$ (and thus i or j are 0) is clear from above.

For the induction step, assume without loss of generality that $i > j$ (or $\ell - j > \ell - i$). Then
$$\langle f_i, f_j\rangle = \langle U^{\ell-i}h_i, U^{\ell-j}h_j\rangle = \langle DU^{\ell-i}h_i, U^{(\ell-1)-j}h_j\rangle.$$
By the use of the relation in Claim 8.8,
$$DU^{\ell-i}h_i = r_i^\ell U^{\ell-i-1}h_i + \delta_i^\ell U^{\ell-i}Dh_i + \Gamma_i,$$
where $\|\Gamma_i\| = O(\gamma)\|h_i\|$. The term $\delta_i^\ell U^{\ell-i}Dh_i$ vanishes as $h_i \in \ker D_i$. Thus
$$\langle DU^{\ell-i}h_i, U^{(\ell-1)-j}h_j\rangle = r_i^\ell\langle U^{(\ell-1)-i}h_i, U^{(\ell-1)-j}h_j\rangle + \langle\Gamma_i, U^{(\ell-1)-j}h_j\rangle.$$
The term $r_i^\ell\langle U^{(\ell-1)-i}h_i, U^{(\ell-1)-j}h_j\rangle$ is bounded by $O(\gamma)\|h_i\|\|h_j\|$ by the induction hypothesis, and by the Cauchy–Schwartz inequality,
$$|\langle\Gamma_i, U^{(\ell-1)-j}h_j\rangle| \le \|\Gamma_i\|\|U^{(\ell-1)-j}h_j\| = O(\gamma)\|h_i\|\|h_j\|.$$
The claim follows. ∎

The preceding lemma gives an error estimate in terms of the norms $\|h_i\|$. The following lemma will enable us to express the error in terms of the norms $\|f_i\|$.

Lemma 8.11. *For any k-dimensional $(\vec{r}, \vec{\delta}, \gamma)$-eposet, let $\ell < k$ and let $f_i = U^{\ell-i}h_i$ for $h_i \in H^i$, as in Theorem 3.2. Then*
$$\|f_i\| = (1 \pm O(\gamma))\rho_{\ell-i}^\ell\|h_i\|,$$
where $\rho_j^\ell = \prod_{t=0}^{j} r_{j-t}^{\ell-t}$, and the hidden constant depends only on $k, \vec{r}, \vec{\delta}$.

Proof. The proof of this lemma is similar to the proof of Lemma 8.10. By direct calculation with Claim 8.8 we obtain that for any $h \in \ker D$:

$$D^j U^j h = r_j^\ell D^{j-1} U^{j-1} h + \Gamma_1 = \cdots = \rho_j^\ell h + \sum_{t=1}^{j} \Gamma_t,$$

where Γ_t is the remainder, and $\|\Gamma_t\| = O(\gamma)\,\|h\|$ for all t. Thus

$$\left\|D^j U^j h - \rho_j^\ell h\right\| = O(\gamma)\,\|h\|\,.$$

Hence using Cauchy–Schwartz,

$$\|f_i\|^2 = \langle U^{\ell-i} h_i, U^{\ell-i} h_i\rangle = \langle D^{\ell-i} U^{\ell-i} h_i, h_i\rangle = \rho_{\ell-i}^\ell \|h_i\|^2 \pm O(\gamma)\|h_i\|^2.$$ □

Combining both lemmas, we obtain the following corollary, which proves the first item of Theorem 8.6.

Corollary 8.12. *Suppose that X be a k-dimensional $(\vec{r}, \vec{\delta}, \gamma)$-eposet, let $\ell < k$, and let $f \in C^\ell$ have the decomposition $f = f_{-1} + \cdots + f_\ell$, as in Theorem 3.2. Then for small enough γ,*

$$\langle f_i, f_j\rangle = O(\gamma)\|f_i\|\|f_j\|,$$

where the hidden constant depends only on $k, \vec{r}, \vec{\delta}$.

As a consequence, we obtain an *approximate L_2 mass formula*, constituting the second item of Theorem 8.6:

Corollary 8.13. *Under the conditions of Corollary 8.12, for every $i \le j$ we have*

$$\|f_i + \cdots + f_j\|^2 = (1 \pm O(\gamma))(\|f_i\|^2 + \cdots + \|f_j\|^2),$$

where the hidden constant depends only on $k, \vec{r}, \vec{\delta}$.

In particular,

$$\|f\|^2 = (1 \pm O(\gamma))(wt_{\le i}(f) + wt_{>i}(f)) = (1 \pm O(\gamma))(\|f_{\le i}\|^2 + \|f_{>i}\|^2).$$

Proof. Expanding $\|f_i + \cdots + f_j\|^2$, we obtain

$$\left|\|f_i + \cdots + f_j\|^2 - \|f_i\|^2 - \cdots - \|f_j\|^2\right| \le 2 \sum_{i \le a < b \le j} |\langle f_a, f_b\rangle| = O(\gamma) \sum_{i \le a < b \le j} \|f_a\|\|f_b\| \le$$
$$O(\gamma)\left(\|f_i\| + \cdots + \|f_j\|\right)^2 \le O(\gamma)(\|f_i\|^2 + \cdots + \|f_j\|^2),$$

swallowing a factor of $j - i + 1$ in the last inequality. □

The fourth item of Theorem 8.6 follows from the preceding ones:

Corollary 8.14. *Under the conditions of Corollary 8.12,*

$$\langle DUf, f\rangle = (1 \pm O(\gamma)) \sum_{i=-1}^{\ell} r_{\ell-i+1}^\ell \|f_i\|^2.$$

Proof. Let $DUf_i = r_{\ell-i+1}^\ell f_i + g_i$, where $\|g_i\| = O(\gamma)\|f_i\|$ according to the third item. Then

$$\langle DUf, f\rangle = \sum_{i=-1}^{\ell} r_{\ell-i+1}^\ell \langle f_i, f\rangle + \sum_{i=-1}^{\ell} \langle g_i, f\rangle. \tag{19}$$

We can bound the magnitude of the second term using Cauchy–Schwartz:

$$\sum_{i=-1}^{\ell} |\langle g_i, f\rangle| \le \sum_{i=-1}^{\ell} \|g_i\|\|f\| = O(\gamma) \sum_{i=1}^{\ell} \|f_i\|\|f\| = O(\gamma)\|f\|^2,$$

using the second item.

For every i, we can bound $\langle f_i, f\rangle$ by

$$\langle f_i, f\rangle = \|f_i\|^2 + \sum_{j\neq i}\langle f_i, f_j\rangle = \|f_i\|^2 \pm O(\gamma)\|f_i\|\|f\|,$$

using the first two items.

Substituting both bounds in (19) and using the second item again, we get

$$\langle DUf, f\rangle = \sum_{i=-1}^{\ell} r^{\ell}_{\ell-i+1}\|f_i\|^2 \pm O(\gamma)\|f\|^2. \qquad \square$$

Finally, we prove Claim 8.8:

Proof of Claim 8.8. We prove the claim by induction on j. The base case $j = 1$ follows by the definition of an eposet: $\delta^{\ell}_1 = \delta_{\ell}$, $r^{\ell}_1 = r_{\ell}$, and

$$\left\|DU - r^{\ell}_1 I - \delta^{\ell}_1 UD\right\| \leq \gamma.$$

For the induction step on $j+1$, note that $DU^{j+1} = DU^jU$. We add and subtract:

$$DU^jU = \left[DU^jU - (r^{\ell}_j U^{j-1}U + \delta^{\ell}_j U^j DU)\right] + (r^{\ell}_j U^j + \delta^{\ell}_j U^j DU). \tag{20}$$

The term inside the square brackets is $O(\gamma)\|U\|$ due to the induction hypothesis. Since $\|U\| \leq 1$,

$$\|DU^jU - r^{\ell}_j U^{j-1}U + \delta^{\ell}_j U^j DU\| = O(\gamma).$$

We consider next the term $\delta^{\ell}_j U^j DU$, and substitute the DU in it with

$$(r_{\ell-j} I + \delta_{\ell-j} UD) + \Gamma,$$

where $\Gamma = DU - (r_{\ell-j} I + \delta_{\ell-j} UD)$ has norm at most γ. We get that

$$\left\|\delta^{\ell}_j U^j DU - \delta^{\ell}_j U^j (r_{\ell-j} I + \delta_{\ell-j} UD)\right\| = O(\gamma).$$

We rearrange the left-hand side of the equation to get

$$\delta^{\ell}_j U^j DU - \delta^{\ell}_j U^j (r_{\ell-j} I + \delta_{\ell-j} UD) = \delta^{\ell}_j U^j DU - r_{\ell-j}\delta^{\ell}_j U^j - \delta^{\ell}_{j+1} UD.$$

Plugging this term back in (20), we get

$$\|DU^{j+1} - r^{\ell}_{j+1} U^j - \delta^{\ell}_{j+1} U^{j+1} D\| = O(\gamma). \qquad \square$$

8.3 Decomposition in the Grassmann poset

Applying Theorem 8.2, we obtain the following properties on the decomposition of $\mathrm{Gr}_q(n, d)$. These properties are well-known in the literature, but we rederive them to show the versatility of Theorem 8.6:

Claim 8.15. *Fix some $d, n \in \mathbb{N}$, let $X = \mathrm{Gr}_q(n, d)$, and let $\ell < d$. Let $f \colon X(\ell) \to \mathbb{R}$ be an arbitrary function. Then we can decompose $f = f_{-1} + \cdots + f_{\ell}$, where $f_i \in V^i$:*

1. *For $i \neq j$, $\langle f_i, f_j\rangle = 0$.*
2. *$\|f\|^2 = \|f_{-1}\|^2 + \cdots + \|f_{\ell}\|^2$.*
3. *The f_i's are eigenvectors of DU. The eigenvalues are*

$$r^{\ell}_i = 1 - \prod_{j=\ell-i+1}^{\ell}\left(1 - \frac{q-1}{q^{j+2}-1}\right) + \Theta(1/q^{n-\ell}).$$

4. *In particular, DU has a constant spectral gap, that is, all its eigenvalues are bounded by a constant strictly smaller than* 1 *when* n *is large enough compared to* ℓ*:*

$$r_i^\ell \leq \frac{q}{q^2-1} + O\left(\frac{1}{q^{n-\ell}}\right) < 1.$$

Proof. The first two items are by invoking Theorem 8.6 and using (15), which shows the Grassmann poset is a sequentially differential poset.

The third item is by invoking Theorem 8.6, and using the fact that $\mathrm{Gr}_q(n, \ell)$ is also an *expanding poset*, with $r_i = \frac{q-1}{q^{i+2}-1}$, $\delta_i = 1 - r_i$, and $\gamma = O(1/q^{n-\ell})$. The fourth item is by direct calculation: one may show using induction that the approximate formula for r_i^ℓ is

$$1 - \prod_{j=i}^{\ell}\left(1 - \frac{q-1}{q^{j+2}-1}\right) \leq \sum_{j=i}^{\ell} \frac{q-1}{q^{j+2}-1}.$$

By taking ℓ to infinity and rearranging, we obtain

$$\sum_{j=i}^{\ell} \frac{q-1}{q^{j+2}-1} \leq (q-1) \sum_{j=i+2}^{\infty} \frac{q^j}{q^j(q^j-1)} \leq (q-1)\frac{q^{i+2}}{q^{i+2}-1} \sum_{j=i+2}^{\infty} \frac{1}{q^j}.$$

The infinite sum converges to $\frac{1}{q^{i+1}(q-1)}$, and so

$$(q-1)\frac{q^{i+2}}{q^{i+2}-1} \sum_{j=i+2}^{\infty} \frac{1}{q^j} = (q-1)\frac{q^{i+2}}{q^{i+2}-1}\frac{1}{q^{i+1}(q-1)} = \frac{q}{q^{i+2}-1}.$$

Hence $r_i^\ell \leq \frac{q}{q^2-1} + O\left(\frac{1}{q^{n-\ell}}\right)$. □

Remark 8.16.
1. The actual values for r_i^ℓ can also be calculated by the formula devised in Theorem 8.6. The calculations are omitted, as they don't add any additional insight.
2. This result is also analogous to the decomposition of the complete complex, see for example [FM16].

8.4 Is there a bounded degree Grassmann poset?

A high dimensional expander, as constructed in [LSV05a], is a simplicial complex that is an eposet and a bounded-degree sub-complex of the complete complex. Is there an analogous construction of an eposet that is a bounded-degree subcomplex of the Grassmann poset? We conjecture the existence of such posets:

Conjecture 8.17. For any prime power q, $d \in \mathbb{N}$, and any $0 < \gamma < 1$, there exists an infinite sequence of natural numbers $n_1 < n_2 < n_3 < \dots$ such for all $n = n_j$ there exists a d-dimensional measured poset X with the following properties:

1. X is sparse, that is $|X(0)| = n$ and $X(d) = O(n)$ (the O-notation hides a constant that may depend on q, d, but not on n).
2. X may be embedded (as a poset) into $\mathrm{Gr}_q(n, d)$. In addition, for all $i < d$, Π_i is obtained by the same probabilistic experiment described for the Grassmann poset:
   (a) Choose a subspace of dimension $d+1$, $s_d \in X(d)$.
   (b) Given a space s_i of dimension $i+1$, choose s_{i-1} to be a uniformly random codimension 1 subspace of s_i.

   In particular, X is downward closed, that is, if $s \in X$ then every subspace $s' \subset s$ also belongs to X.
3. X is an $(\vec{r}, \vec{\delta}, \gamma)$-eposet for $r_i = \frac{1}{q^{i+2}-1}$ and $\delta_i = 1 - r_i$.

This existence of sub-posets as above is the vector-space analogue of the existence of γ-HDX simplicial complexes. Moreover, it would be interesting to construct such a poset such that Π_0, Π_d are uniform. Note however that even in the known constructions for γ-HDX simplicial complexes, Π_d is not uniform (but Π_0 is uniform).

Moshkovitz and Raz [MR08] gave a construction that can be viewed as an interesting step in this direction. They constructed, towards a derandomized low degree test, a small set of planes by choosing only planes spanned by directions coming from a smaller field $\mathbb{H} \subset \mathbb{F}_q$.

8.5 Eposet parameters in a simplicial complex

Although the definition of (approximately) sequentially differential poset allows a range of parameters $\vec{r}$ and $\vec{\delta}$, these parameters turn out to be determined by the laziness of the upper and lower walks. The lemma below shows that any family of simplicial complexes which are eposets, have parameters $\vec{r}, \vec{\delta}$ approaching $r_j = \frac{1}{j+2}$ and $\delta_j = 1 - \frac{1}{j+2}$ as γ goes to zero.

Lemma 8.18. *Let $X^{(m)}$ be a sequence on k-dimensional $(\vec{r}^{(m)}, \vec{\delta}^{(m)}, \gamma^{(m)})$-eposets, where $\lim_{m\to\infty} \gamma^{(m)} = 0$. Then for all $j \leq k-1$:*

$$\lim_{m\to\infty} r_j^{(m)} + \delta_j^{(m)} = 1. \tag{21}$$

Furthermore, suppose that the following two conditions hold:

1. *For all $j \leq k-1$, the laziness of $U_{j-1}D_j$, goes to 0 as m goes to infinity:*
$$\lim_{m\to\infty} \Pr_{(t_1,t_2)\sim UD}[t_1 = t_2] = 0.$$
2. *There exists $\vec{\alpha}$ such that for all $j \leq k-1$, $D_{j+1}U_j = \alpha_j I + (1-\alpha_j)M^+$, where M^+ is a non-lazy averaging operator.*

Then

$$\lim_{m\to\infty} r_j^{(m)} = \alpha_j \text{ and } \lim_{m\to\infty} \delta_j^{(m)} = 1 - \alpha_j,$$

In particular, if $X^{(m)}$ are k-dimensional simplicial complexes, then $\alpha_j = \frac{1}{j+2}$ and we get

$$\lim_{m\to\infty} r_j^{(m)} = \frac{1}{j+2} \text{ and } \lim_{m\to\infty} \delta_j^{(m)} = 1 - \frac{1}{j+2},$$

under the mild assumption that the laziness probability of UD goes to zero. In other words, the interesting eposets are γ-HDXs.

Proof. To prove both assertions, we use the definition of an eposet to get the following inequality:

$$\left| \frac{\langle (DU - r_j^{(m)} I - \delta_j^{(m)} UD) f, f\rangle}{\langle f, f\rangle} \right| \leq \gamma^{(m)}, \tag{22}$$

for any function $f \in C^k$. We use this inequality on specific functions f we choose: the constant function, and indicator functions.

To show that

$$\lim_{m\to\infty} r_j^{(m)} + \delta_j^{(m)} = 1,$$

we apply $DU - r_j^{(m)} I - \delta_j^{(m)} UD$ to the constant vector $\mathbb{1}$, which is fixed by all of DU, I, UD:

$$\left| \frac{\langle (DU - r_j^{(m)} I - \delta_j^{(m)} UD)\mathbb{1}, \mathbb{1}\rangle}{\langle \mathbb{1}, \mathbb{1}\rangle} \right| \leq \gamma^{(m)} \Longrightarrow |1 - r_j^{(m)} - \delta_j^{(m)}| \leq \gamma^{(m)},$$

thus $\lim_{m\to\infty} r_j^{(m)} + \delta_j^{(m)} = 1$.

To show that $\lim_{m\to\infty} r_j^{(m)} = \alpha_j$, we fix j and take a sequence of $\sigma^{(m)} \in X(j)$ such that probability of laziness given that $t_1 = \sigma^{(m)}$ goes to zero:

$$\lim_{m\to\infty} \Pr_{(t_1,t_2)\sim UD}[t_2 = \sigma^{(m)} | t_1 = \sigma^{(m)}] = 0.$$

Denote by $\mathbb{1}_{\sigma^{(m)}}$ the indicator of $\sigma^{(m)}$. Then

$$\frac{\langle UD\mathbb{1}_{\sigma^{(m)}}, \mathbb{1}_{\sigma^{(m)}}\rangle}{\langle \mathbb{1}_{\sigma^{(m)}}, \mathbb{1}_{\sigma^{(m)}}\rangle} = \Pr_{(t_1,t_2)\sim UD}[t_2 = \sigma^{(m)} | t_1 = \sigma^{(m)}].$$

Moreover,

$$\frac{\langle (DU - r_j^{(m)} I)\mathbb{1}_{\sigma^{(m)}}, \mathbb{1}_{\sigma^{(m)}}\rangle}{\langle \mathbb{1}_{\sigma^{(m)}}, \mathbb{1}_{\sigma^{(m)}}\rangle} = \alpha_j - r_j^{(m)}.$$

Plugging $f = \mathbb{1}_{\sigma^{(m)}}$ into (22), we get

$$\left|\alpha_j - r_j^{(m)} - \delta_j^{(m)} \Pr_{(t_1,t_2)\sim UD}[t_2 = \sigma^{(m)} | t_1 = \sigma^{(m)}]\right| \leq \gamma^{(m)}.$$

Since the $\delta_j^{(m)}$ are bounded, this shows that $\lim_{m\to\infty} r_j^{(m)} = \alpha_j$. The analogous statement for $\delta_j^{(m)}$ follows from (21). □

References

[BGH$^+$15] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. *Making the long code shorter*. SIAM J. Comput., 44(5):1287–1324, 2015. (Preliminary version in *53rd FOCS*, 2012). arXiv:1111.0405, eccc:2011/TR11-142. 1

[BKS19] Boaz Barak, Pravesh Kothari, and David Steurer. *Small-set expansion in shortcode graph and the 2-to-2 conjecture*. In Avrim Blum, ed., *Proc.* 10*th Innovations in Theor. Comput. Sci. (ITCS)*, volume 124 of *LIPIcs*. Schloss Dagstuhl, 2019. (To appear). arXiv:1804.08662, eccc:2018/TR18-077. 1

[DDFH18] Yotam Dikstein, Irit Dinur, Yuval Filmus, and Prahladh Harsha. *Boolean function analysis on high-dimensional expanders*. In Eric Blais, Klaus Jansen, José D. P. Rolim, and David Steurer, eds., *Proc.* 20*th International Workshop on Randomization and Computation (RANDOM)*, volume 116 of *LIPIcs*, pages 38:1–38:20. Schloss Dagstuhl, 2018. arXiv:1804.08155, eccc:2018/TR18-075. 1

[DFH19] Irit Dinur, Yuval Filmus, and Prahladh Harsha. *Analyzing Boolean functions on the biased hypercube via higher-dimensional agreement tests*. In *Proc.* 30*th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 2124–2133. 2019. arXiv:1711.09428,1711.09426, eccc:2017/TR17-180,2017/TR17-181. 5

[DHK$^+$19] Irit Dinur, Prahladh Harsha, Tali Kaufman, Inbal Livni Navon, and Amnon TaShma. *List decoding with double samplers*. In *Proc.* 30*th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 2134–2153. 2019. eccc:2018/TR18-198. 1

[DK17] Irit Dinur and Tali Kaufman. *High dimensional expanders imply agreement expanders*. In *Proc.* 58*th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 974–985. 2017. eccc:2017/TR17-089. 1, 2, 5, 6, 9, 10, 11, 12, 17

[DKK$^+$18a] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. *On non-optimally expanding sets in Grassmann graphs*. In *Proc.* 50*th ACM Symp. on Theory of Computing (STOC)*, pages 940–951. 2018. eccc:2017/TR17-094. 1, 4

[DKK$^+$18b] ———. *Towards a proof of the 2-to-1 games conjecture?* In *Proc.* 50*th ACM Symp. on Theory of Computing (STOC)*, pages 376–389. 2018. eccc:2016/TR16-198. 1, 4

[Dun76] Charles Dunkl. *A Krawtchouk polynomial addition theorem and wreath products of symmetric groups*. Indiana Univ. Math. J., 25:335–358, 1976. 3

[EFF15a] David Ellis, Yuval Filmus, and Ehud Friedgut. *A quasi-stability result for dictatorships in S_n*. Combinatorica, 35(5):573–618, 2015. arXiv:1209.5557. 1

[EFF15b] ———. *A stability result for balanced dictatorships in S_n*. Random Structures Algorithms, 46(3):494–530, 2015. arXiv:1210.3989. 1

[EFF17] ———. *Low degree Boolean functions on s_n, with an application to isoperimetry*. Forum of Mathematics, Sigma, 5:e23, 2017. arXiv:1511.08694. 1

[EK16] Shai Evra and Tali Kaufman. *Bounded degree cosystolic expanders of every dimension*. In *Proc. 48th ACM Symp. on Theory of Computing (STOC)*, pages 36–48. 2016. arXiv:1510.00839. 2, 11

[Fil16a] Yuval Filmus. *Friedgut-Kalai-Naor theorem for slices of the Boolean cube*. Chic. J. Theoret. Comput. Sci., 2016(14), 2016. arXiv:1410.7834. 1, 5, 16

[Fil16b] ———. *An orthogonal basis for functions over a slice of the Boolean hypercube*. Electron. J. Combin., 23(1):P1.23, 2016. arXiv:1406.0142. 1

[FKMW16] Yuval Filmus, Guy Kindler, Elchanan Mossel, and Karl Wimmer. *Invariance principle on the slice*. In *Proc. 31st Comput. Complexity Conf.*, volume 50 of *LIPIcs*, pages 15:1–15:10. Schloss Dagstuhl, 2016. arXiv:1504.01689. 1, 3

[FKN02] Ehud Friedgut, Gil Kalai, and Assaf Naor. *Boolean functions whose Fourier transform is concentrated on the first two levels and neutral social choice*. Adv. Appl. Math., 29(3):427–437, 2002. 5, 16

[FM16] Yuval Filmus and Elchanan Mossel. *Harmonicity and invariance on slices of the Boolean cube*. In *Proc. 31st Comput. Complexity Conf.*, volume 50 of *LIPIcs*, pages 16:1–16:13. Schloss Dagstuhl, 2016. arXiv:1507.02713. 1, 3, 26

[KKL14] Tali Kaufman, David Kazhdan, and Alexander Lubotzky. *Ramanujan complexes and bounded degree topological expanders*. In *Proc. 55th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 484–493. 2014. arXiv:1408.6351. 2

[KM18] Tali Kaufman and David Mass. *Good distance lattices from high dimensional expanders*, 2018. (manuscript). arXiv:1803.02849. 1, 2

[KMS17] Subhash Khot, Dor Minzer, and Muli Safra. *On independent sets, 2-to-2 games, and Grassmann graphs*. In *Proc. 49th ACM Symp. on Theory of Computing (STOC)*, pages 576–589. 2017. eccc:2016/TR16-124. 1, 4

[KMS18] ———. *Pseudorandom sets in Grassmann graph have near-perfect expansion*. In *Proc. 59th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 592–601. 2018. eccc:2018/TR18-006. 1, 4

[KO18] Tali Kaufman and Izhar Oppenheim. *High order random walks: Beyond spectral gap*. In Eric Blais, Klaus Jansen, José D. P. Rolim, and David Steurer, eds., *Proc. 20th International Workshop on Randomization and Computation (RANDOM)*, volume 116 of *LIPIcs*. Schloss Dagstuhl, 2018. arXiv:1707.02799. 2, 9, 11

[LSV05a] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. *Explicit constructions of Ramanujan complexes of type $\tilde{A}_d$*. European J. Combin., 26(6):965–993, 2005. arXiv:math/0406217. 1, 2, 12, 26

[LSV05b] ———. *Ramanujan complexes of type $\tilde{A}_d$*. Israel J. Math., 149(1):267–299, 2005. arXiv:math/0406208. 1, 2

[MO10] Ashley Montanaro and Tobias Osborne. *Quantum boolean functions*. Chic. J. Theoret. Comput. Sci., 2010, 2010. arXiv:0810.2435. 1

[MR08] Dana Moshkovitz and Ran Raz. *Sub-constant error low degree test of almost-linear size*. SIAM J. Comput., 38(1):140–180, 2008. (Preliminary version in *38th STOC*, 2006). eccc:2005/TR05-086. 27

[Opp18] Izhar Oppenheim. *Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps*. Discrete Comput. Geom., 59(2):293–330, 2018. arXiv:1709.04431. 10

[OW13] Ryan O'Donnell and Karl Wimmer. *KKL, Kruskal-Katona, and monotone nets*. SIAM J. Comput., 42(6):2375–2399, 2013. (Preliminary version in *50th FOCS*, 2009). 1

[Pla15] Rafael Plaza. *Stability for intersecting families in $PGL(2, q)$*. Electron. J. Combin., 22(4):P4.41, 2015. 1

[Sta88] Richard P. Stanley. *Differential posets*. J. Amer. Math. Soc., 1(4):919–961, 1988. 4, 20

[Sta90] ———. *Variations on differential posets*. In Dennis Stanton, ed., *Invariant Theory and Tableaux*, volume 19 of *IMA Vol. Math. Appl.*, pages 145–165. Springer, 1990. 20

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il