Chernoff Bounds and Reverse Hypercontractivity on HDX

Yotam Dikstein[*] and Max Hopkins[†]

April 16, 2024

Abstract

We prove optimal concentration of measure for lifted functions on high dimensional expanders (HDX). Let X be a k-dimensional HDX. We show for any $i \leqslant k$ and function $f : X(i) \to [0,1]$:

$$\mathbb{P}_{s \in X(k)} \left[\left| \mathbb{E}_{t \subseteq s} [f(t)] - \mu \right| \geqslant \varepsilon \right] \leqslant \exp\left(-\varepsilon^2 \frac{k}{i} \right).$$

Using this fact, we prove that high dimensional expanders are *reverse hypercontractive*, a powerful functional inequality from discrete analysis implying that for any sets $A, B \subset X(k)$, the probability a ρ-correlated pair passes between them is at least

$$\mathbb{P}_{s,s' \sim T_\rho} [s \in A, s' \in B] \geqslant \mathbb{P}[A]^{O(1)} \mathbb{P}[B]^{O(1)}.$$

Our results hold under weak spectral assumptions on X. Namely we prove exponential concentration of measure for any complex below the 'Trickling-Down Threshold' (beyond which concentration may be arbitrarily poor), and optimal concentration for $\sqrt{k}$-skeletons of such complexes. We also show optimal bounds for the top dimension of stronger HDX among other settings.

We leverage our inequalities to prove several new agreement testing theorems on high dimensional expanders, including a new 99%-regime test for *subsets*, and a variant of the 'Z-test' achieving *inverse exponential soundness* under the stronger assumption of ℓ_∞-expansion. The latter gives rise to the first optimal testers beyond the complete complex and products, and a stepping stone toward strong soundness the use of HDX in strong soundness PCPs.

We also give applications within expansion, analysis, combinatorics, and coding theory, including a proof that two-sided HDX have optimal geometric overlap (giving the first explicit bounded-degree construction), near-optimal double samplers, new super-exponential degree lower bounds for certain HDX, distance-amplified list-decodable and locally testable codes, a Frankl–Rödl Theorem, and more.

[*]Institute for Advanced Study, USA. email: yotam.dikstein@gmail.com.

Contents

1 Introduction

Recent years have seen the emergence of *high dimensional expanders* (HDX) as a core tool in theoretical computer science, with breakthrough applications in approximate sampling [Ana+19; ALG21], coding theory [Din+22a; PK22], and quantum complexity [ABN23], and strong promise toward longstanding problems in hardness of approximation and probabilistically checkable proofs [DK17; DD19; KM20; GK22; HL22; BLM24; DDL24]. One central force behind the success of HDX in application is their *concentration of measure*. Consider (as a warmup) the following fundamental question: given a k-uniform hypergraph X and a function $f : X(1) \to [0,1]$ on its vertices, how concentrated is f to its mean across a random hyperedge?

$$\mathbb{P}_{\{v_1,\ldots,v_k\} \in X(k)} \left[\left| \frac{1}{k} \sum_{i=1}^{k} f(v_i) - \mu \right| > \varepsilon \right] \overset{?}{\leqslant} \beta(\varepsilon, k) \tag{1.1}$$

When X is the complete hypergraph, this was classically resolved by Chernoff and Hoeffding [Che52; Hoe94] who showed $\beta(\varepsilon, k) \leqslant \exp(-\varepsilon^2 k)$. Over the years, near-matching bounds have been shown for more general hypergraph families (see e.g. [Pau14; KS18; KKS21]), and even for certain bounded degree systems like walks on expanders [AKS87; Gil98]. These objects, called *sampler graphs* or *extractors*, are by now a core tool in the field with applications in de-randomization, complexity, and cryptography [Gol97; Wig09].

High dimensional expanders are known to satisfy a 'Chebyshev-type' bound $\beta(\varepsilon, k) \leqslant \frac{1}{k\varepsilon^2}$ for (1.1) [DK17], but this is of course not what makes them powerful. Rather, many modern applications in complexity require concentration of X against *broader classes of functions*. Of particular interest is the extension of (1.1) to functions $f : X(i) \to [0,1]$ sitting on *i-sets* of X. Complexes satisfying such a bound are called *inclusion samplers*; they arise naturally in the study of agreement and PCPs (forming consistency checks), and play a critical role in low soundness constructions toward the Sliding Scale Conjecture [IKW12; DM11; Mos17; DL17]. Unfortunately, inclusion samplers are notoriously difficult to construct, and no bounded degree families were known for many years.

This changed with the advent of spectral HDX [Gar73; LSV05b; EK16; KM17; DK17; Opp18]. A key motivation behind the modern incarnation of these objects [KM17; DK17; Opp18] was the fact that their 'inclusion graphs' (the bipartite adjacency matrix between k-sets and i-sets) are spectral expanders, and therefore satisfy 'Chebyshev-type' concentration $\beta(\varepsilon, i, k) \leqslant \frac{i}{\varepsilon^2 k}$. This observation played a central role in recent progress in agreement testing [DK17; Din+18a; DD19; KM20; GK22; DD23a; BM23a], but left an exponential gap from the Chernoff-type concentration needed for other applications.

In this work we take a major step toward closing this gap, resolving the problem completely in many regimes of interest. In particular, drawing on ideas from the concentration of measure [BLM13], PCP [IKW12], and HDX [DD19; AJT19] literature, we prove HDX are optimal samplers between every two levels.

Theorem 1.1 (HDX are Optimal Inclusion Samplers (Informal))**.** *Let X be a d-dimensional HDX. Then for any $i \leqslant k \leqslant d$ and any $f : X(i) \to [0,1]$ of expectation μ:*

$$\mathbb{P}_{s \in X(k)} \left[\left| \underset{t \subset s}{\mathbb{E}}[f(t)] - \mu \right| \geqslant \varepsilon \right] \leqslant \exp\left(-\varepsilon^2 \frac{k}{i} \right).$$

Moreover, this is essentially tight—no inclusion sampler achieves $\exp\left(-\omega\left(\varepsilon^2 \frac{k}{i} \right) \right)$ error.

Theorem 1.1 is powerful when $i = o(k)$, but our corresponding lower bounds show non-trivial inclusion sampling is essentially impossible in the critical regime $i = \Theta(k)$. In fact this exact issue stood as a barrier for many years on the complete complex, where Impagliazzo, Kabanets, and Wigderson [IKW12] proved Theorem 1.1 and used it to construct sub-optimal testers and PCPs by taking $i = \Theta(\sqrt{k})$. In 2017, Dinur and Livni-Navon [DL17] resolved this problem by appealing to a somewhat surprising player: an underutilized tool from boolean function analysis called *reverse hypercontractivity* (RHC).

Reverse hypercontractivity is a classic functional inequality due to Borell [Bor82] that lower bounds the correlation between sets on the noisy hypercube. Roughly speaking, it implies that for any two subsets $A, B \subseteq \{0,1\}^d$, the probability a ρ-correlated edge crosses A and B is at least some power of their measure:

$$\mathbb{P}_{(s,s') \sim T_\rho}[s \in A, s' \in B] \geqslant \mathbb{P}[A]^{O_\rho(1)} \mathbb{P}[B]^{O_\rho(1)}.$$

Reverse hypercontractivity is an even rarer phenomenon than inclusion sampling—while there are some moderately de-randomized inclusion graphs (namely the Grassmann [IKW12]) satisfying Chebyshev-type variants of Theorem 1.1, none are known to admit reverse hypercontractivity. At outset this may even seem necessary. The inequality is closely tied to tensorization and (modified) Log-Sobolev Inequalities [MOS13], which inherently fail on sparse hypergraphs.

In this work, we circumvent this issue by a new combinatorial argument showing any hypergraph satisfying Theorem 1.1 'locally' is also reverse hypercontractive. This leads to the following corollary for HDX:

Theorem 1.2 (Reverse Hypercontractivity for HDX (Informal)). *let X be a d-dimensional HDX. Then for any $\rho \in (0, 1)$, $k \leqslant d$, and $A, B \subset X(k)$:*

$$\underset{(s,s') \sim T_\rho(X)}{\mathbb{P}}[s \in A, s' \in B] \geqslant \mathbb{P}[A]^{O_\rho(1)} \mathbb{P}[B]^{O_\rho(1)}.$$

Here $T_\rho(X)$ is the natural generalization of the noise operator to hypergraphs which, roughly speaking, generates s' by re-sampling each vertex in s with probability $1 - \rho$. We also prove a more standard analytical version of reverse hypercontractivity for all functions by reduction to the above, as well as analog statements for the heavily studied 'down-up' walks which play a critical role in almost all HDX applications.

Applications. The remainder of our work is devoted to proving a variety of applications of Theorem 1.1 and Theorem 1.2 including new families of agreement tests, bounded degree complexes with optimal geometric overlap, double samplers with near-optimal overhead, new degree lower bounds for HDX, large distance list-decodable and locally-testable codes, and several extensions of classic combinatorial and analytic results to HDX. Here we choose to highlight just two of these applications, agreement testing and geometric overlap. We refer the reader to Section 2.5 and Section 2.6 for a more in-depth discussion of our applications.

Agreement testing is a general property testing paradigm that generalizes classical subroutines in PCP theory like the plane vs. plane test (e.g. [AS97; GS97; Din07]). Roughly speaking, an agreement test consists of a universe U and a family of overlapping subsets $S \subset \mathbb{P}(U)$ (think of U as the vertices of X, and S as its hyperedges). Given a family of *local* assignments $\{f_s : s \to \{0, 1\}\}_{s \in S}$, we'd like to test whether f_s actually agrees with a *global* function $g : U \to \{0, 1\}$ in the sense that $f_s = g|_s$ for many s. These systems occur in PCP reductions to ensure the provers (who answer on subsets) can't 'cheat' by using answers not corresponding to a real solution of the initial problem (corresponding to global functions on U).

Agreement testing has two major regimes, the '99%-regime' (where we'd like to infer global structure only if the test passes with high probability) and the '1%-regime' (where we'd like to infer global structure even when the test only passes with *non-trivial* probability). Leveraging reverse hypercontractivity, we give new tests in both regimes. In the 99%-regime, we give a new 2-query test where the universe U corresponds to *i-sets* of a high dimensional expander, extending a similar result of [DFH19] on the complete complex. Our test, along with Theorem 1.1, has since appeared as an important sub-routine in the construction of the first bounded-degree 1% agreement testers in [DD23b; DDL24].

In the 1%-regime, we give a new 3-query tester for a variant of ℓ_∞-expanders (a strengthened notion of HDX [KKS21]) with *optimal soundness*, meaning that if the test passes with probability asymptotically better than a random function ($2^{-\Theta(k)}$ for a k-dimensional ℓ_∞-expander), we are able to infer global structure. In the context of PCPs, such a soundness guarantee is critical to ensure the alphabet size stays polynomial in the soundness. Our result gives rise to the first families of optimal testers beyond the complete complex and products [DL17], including (dense) random complexes, skeletons of many well-studied spin systems, the full linear matroid, and more. We remark that while all such examples are dense, we prove an optimal 'local' agreement theorem only under the assumption of local spectral expansion. We lift this guarantee to the general result using ℓ_∞-expansion, but it is possible this requirement could be relaxed using recent techniques based on coboundary expansion [GK22; BM23b; DD23b; BLM24; DDL24] to give sparse optimal testers.

The *Geometric Overlap Property* is one of the earliest notions of high dimensional expansion [Gro10; Lub18]. A d-dimensional complex X has c-geometric overlap if for any embedding of X into $\mathbb{R}^d$, there is a point $p \in \mathbb{R}^d$ that is covered by a c-fraction of X's hyperedges. Geometric overlap was first proven for the complete complex in 2-dimensions by Boros and Füredi [BF84], and later extended to all dimensions by Bárány [Bár82]. Gromov [Gro10] asked whether there are bounded degree complexes satisfying geometric

overlap. This was resolved in [Fox+12], who gave both an optimal random construction and several explicit constructions with bounded but sub-optimal overlap (including one from high dimensional expanders). Several later works [PRT16; Evr17; Opp18] continued to build the connection between high dimensional expanders and geometric overlap, but it remained open whether a construction achieving the best of both worlds (explicit and optimal) could be achieved. Leveraging a variant of Theorem 1.1, we resolve this problem, showing under mild assumptions that *any* sufficiently strong high dimensional expander has near-optimal geometric overlap.

The Trickling-Down Threshold. High dimensional expanders exhibit a *phase transition* at a certain expansion parameter called the 'Trickling-Down (TD)-Threshold' [Opp18]. Any complex breaking this barrier immediately has 'local-to-global' structure, meaning properties such as global expansion and fast-mixing can be inferred just from *local* structure of the complex [Gar73; KM17; Opp18; AL20]. This barrier shows up as a major point in constructions of HDX as well. It is relatively easy to build complexes *at* the threshold by tensoring an expander with the complete complex [LMY20; Gol21], but *breaking* the barrier requires complicated algebraic machinery and is considered the 'gold standard' for HDX—indeed only three such constructions are known [LSV05b; KO18; Dik22].

Theorem 1.1 and Theorem 1.2 hold for $\sqrt{d}$-dimensional skeletons of any complex breaking the TD-Threshold. In fact, we prove that the top level of such complexes satisfy exponential concentration for all Lipschitz functions, corresponding to $\beta(\varepsilon, i, k) \leqslant \exp(-\varepsilon \frac{\sqrt{k}}{i})$ in the inclusion sampling setting. On the other hand, we give examples of complexes *at* the TD-threshold that exhibit arbitrarily poor concentration. We view our results in this sense as a significant strengthening of HDX's 'local-to-global' phase transition, and a partial explanation to the great difficulty of constructing hypergraphs beyond the TD-Threshold. Whereas graph expansion and mixing are 'common' phenomenon, Theorem 1.1 and Theorem 1.2 imply as soon as one breaks the TD-Threshold X not only satisfies exponential concentration of measure, its skeletons satisfy optimal concentration and reverse hypercontractivity—rare properties not known for any other sparse systems.

2 Main Results and Proof Overview

2.1 Background

Before stating our results in somewhat more formality, we cover some basic background on high dimensional expanders, random walks, and sampling. We refer the reader to Section 4 for formal details and discussion.

Simplicial Complexes. A d-**uniform simplicial complex** X consists of a d-uniform hypergraph $X(d)$ together with its downward closure
$$X = X(1) \cup \ldots \cup X(d),$$
where $X(i) \subseteq \binom{[n]}{i}$, called the '$i$-faces', are all i-size subsets that sit in some hyperedge in $X(d)$. Given a face $t \in X(i)$, the **link** of t is the sub-complex induced by localizing to faces that include t, that is $X_t := \{s : s \cup t \in X\}$. We say X is **connected** if the base graph $(X_t(1), X_t(2))$ of every link is connected. A d-uniform complex is **partite** if its vertices can be partitioned into d 'parts' such that each top-level face has one part from each component. Finally, the **inclusion graph** $(X(k), X(i))$ is the bipartite graph between k-sets and i-sets of X where edges are given by inclusion.

We emphasize that we have changed convention here from dimension, where $X(d)$ refers to sets of size $d + 1$, to uniformity where $X(d)$ consists of sets of size d. We will use the latter notation throughout the rest of the paper to simplify both the statements and proofs of our results.

High dimensional expanders. A simplicial complex X is a λ-**one-sided HDX** (resp. two-sided) if for every $s \in X$ of co-dimension[1] at least 2 $(X_t(1), X_t(2))$ is a λ-one-sided spectral expander (resp. two sided).

The weakest form of high dimensional expanders are λ-**Trickling-Down (TD) Complexes**. We call X λ-TD if it is connected and every co-dimension 2 link is a $\frac{\lambda}{d-1}$-one-sided expander.

Our results hold for a broad class of HDX we call **Nice Complexes**. We defer the formal definition to Section 4.8, and here give three basic examples of nice complexes for which our results hold:

---

[1] Here by co-dimension j we mean sets of size $d - j$.

   – X is the $\sqrt{d}$-skeleton of a λ-TD complex

   – X is the $\frac{d}{\log(d)}$-skeleton of a partite 2^{-d}-one-sided HDX

   – X is a 2^{-d}-two-sided HDX

High Order Random Walks. Simplicial complexes admit a variety of 'high order random walks' generalizing the standard random walk on a graph. There are three natural walks critically important to our results. The first are the classical **down-up walks** of [KM17], denoted $U_{k,d}D_{d,k}$, which walk between $s, s' \in X(d)$ through a shared k-face $t \subset s, s'$. The second are the **swap walks** of [DD19; AJT19], denoted $S_{i,j}$, which walk between *disjoint* faces $s \in X(i)$ and $s' \in X(j)$ via $s \cup s' \in X(i+j)$. The final walk is the **noise operator** T_ρ, which walks between $s, s' \in X(d)$ via a random i-subface of s for $i \sim Bin(1 - \rho, d)$.[2]

Sampler Graphs. A bipartite graph $G = (L, R, E)$ is an (ε, β)-function additive sampler if $\forall f : R \to [0, 1]$

$$\mathbb{P}_{v \in L}\left[\left|\mathop{\mathbb{E}}_{u \in R, u \sim v}[f(u)] - \mathop{\mathbb{E}}_{u \in R}[f(u)]\right| \geqslant \varepsilon\right] < \beta.$$

G is called an (α, β, δ)-multiplicative sampler if for all $f : R \to [0, 1]$ of density at least $\mathbb{E}_{u \in R}[f(u)] \geqslant \alpha$:

$$\mathbb{P}_{v \in L}\left[\left|\mathop{\mathbb{E}}_{u \in R, u \sim v}[f(u)] - \mathop{\mathbb{E}}_{u \in R}[f(u)]\right| \geqslant \delta \mathop{\mathbb{E}}_{u \in R}[f(u)]\right] < \beta.$$

It is a well-known fact (c.f. [IKW12]) that one can move between additive and multiplicative samplers without much parameter loss, and moreover that one can 'flip' L and R in the sense that if $G = (L, R, E)$ is an (α, β, δ)-sampler, then $G^{op} = (R, L, E)$ is roughly a (β, α, δ)-sampler up to slight decay in parameters. We refer the reader to Section 4 for details.

2.2 Sampling and Concentration

We can now state and overview the proof of optimal sampling on HDX in somewhat more formality.

Theorem 2.1 (Sampling on HDX (Informal Theorem 5.1))**.** *Let X be a nice, k-uniform HDX and $i \leqslant k$. The containment graph $(X(k), X(i))$ is a (ε, β)-function sampler for*

$$\beta \leqslant \exp\left(-\Omega\left(\varepsilon^2 \frac{k}{i}\right)\right).$$

In other words, for any $f : X(i) \to [0, 1]$ of expectation μ:

$$\mathbb{P}_{s \in X(k)}\left[\left|\mathop{\mathbb{E}}_{t \subseteq s}[f(t)] - \mu\right| \geqslant \varepsilon\right] < \exp\left(-\Omega\left(\varepsilon^2 \frac{k}{i}\right)\right).$$

Since $\sqrt{d}$-skeletons of λ-TD complexes are nice, we get the following immediate corollary.

Corollary 2.2 (Sampling for TD-Complexes (Informal Claim 4.28))**.** *Fix $\lambda < 1$ and let X be a d-uniform λ-TD complex. Then for any $i < k \leqslant \sqrt{d}$, the containment graph of $(X(k), X(i))$ is a (ε, β)-sampler for*

$$\beta \leqslant \exp\left(-c_\lambda \varepsilon^2 \frac{k}{i}\right).$$

for some $c_\lambda > 0$ dependent only on λ.

We remark that this (and all following) results also hold in some variation under the popular notion of *spectral independence*, which for our purposes act similar to TD complexes. We discuss in the main body.

En route to Theorem 2.1, we also show the *top level* of any λ-TD complex satisfies exponential concentration for Lipschitz functions, which may be of independent interest.

---

[2]We remark that while definition this may seem odd at outset, it is the natural extension of the noise operator to hypergraphs and recovers the standard notion on products. See [Baf+22b] or Section 4 for further details.

Theorem 2.3 (Exponential Concentration of λ-TD Complexes (Informal Corollary 7.7))**.** *For any $\lambda < 1$, $\nu > 0$, and $k \in \mathbb{N}$, let X be a k-uniform λ-TD complex. For any ν-Lipschitz function $f : X(k) \to \mathbb{R}$:*

$$\mathbb{P}_{s \in X(d)} [|f(s) - \mathbb{E}[f]| \geqslant t] \leqslant \exp\left(-c_\lambda \frac{t}{\sqrt{\nu}}\right)$$

for some $c_\lambda > 0$ dependent only on λ.

We refer the reader to Section 4.6 for the definition of ν-Lipschitz, and here just note that specialized to inclusion sampling this implies $(X(k), X(i))$ is roughly a $(\varepsilon, \exp(-\varepsilon \frac{\sqrt{k}}{i}))$-sampler. A similar statement holds under spectral independence, albeit for a more restricted class than Lipschitz functions (see Corollary 7.8).

Proof Overview. The proof of Theorem 2.1 is broken into two main components: the 'Chernoff' ($i = 1$) setting, and a bootstrapping argument lifting Chernoff to general i. Here we focus just on the simplest version of the proof, the case of *two-sided HDX*, and refer the reader to Section 7 for the general case.

Part I: Chernoff-Hoeffding. Toward Theorem 2.1, we first prove a fairly general "k vs 1" concentration theorem for any HDX X whose strength depends on X's underlying quantitative expansion. Under strong assumptions (2^{-k}-HDX), we recover true Chernoff. Perhaps more surprisingly, even under very weak assumptions ($\frac{1}{k}$-HDX) we still recover strong concentration of the form $\exp(-\varepsilon^2 \sqrt{k})$. Taking a $\sqrt{k}$-skeleton of the latter recovers Chernoff (see Section 4.6). We give a simplified statement here:

Theorem 2.4 (Chernoff-Hoeffding for HDX (Informal Theorem 5.2))**.** *Let X be a k-uniform $\frac{1}{k}$-two-sided HDX. Then for any $f : X(k) \to [0, 1]$:*

$$\mathbb{P}_{\{v_1, \ldots, v_k\} \in X(k)} \left[\left| \frac{1}{k} \sum_{i=1}^{k} f(v_i) - \mu \right| > \varepsilon \right] \leqslant \exp(-\Omega(\varepsilon^2 \sqrt{k}))$$

Moreover, if X is a 2^{-k}-two-sided HDX stronger concentration holds:

$$\mathbb{P}_{\{v_1, \ldots, v_k\} \in X(k)} \left[\left| \frac{1}{k} \sum_{i=1}^{k} f(v_i) - \mu \right| > \varepsilon \right] \leqslant \exp(-\Omega(\varepsilon^2 k))$$

For simplicity let $f = 1_A$ be the indicator of some $A \subset X(1)$. The proof of Theorem 2.4 follows the standard Chernoff-style method of bounding the moment generating function (MGF). Namely for any $r > 0$:

$$\mathbb{P}_{\{v_1, \ldots, v_k\} \in X(k)} \left[\sum_{i=1}^{k} 1_A(v_i) \geqslant k(\mu + \varepsilon) \right] = \mathbb{P}\left[\exp\left(\sum_{i=1}^{k} r 1_A(v_i) \right) \geqslant \exp(rk(\mu + \varepsilon)) \right] \leqslant \frac{\mathbb{E}\left[\prod_{i=1}^{k} e^{r 1_A(v_i)} \right]}{e^{rk(\mu + \varepsilon)}}.$$

In the classical proof of Chernoff, the variables v_i are independent, so one can bound the MGF by

$$\mathbb{E}_{\{v_1, \ldots, v_k\} \in X(k)} \left[\prod_{i=1}^{k} e^{r 1_A(v_i)} \right] = \mathbb{E}_{v \in X(1)} \left[e^{r 1_A(v)} \right]^k \lesssim e^{rk(\mu + r)} \tag{2.1}$$

and set $r = \Theta(\varepsilon)$ to get the desired bound.

When X is an HDX, the variables $\{v_1, \ldots, v_k\}$ may be extremely correlated, so the above approach breaks down naively. Instead, we take inspiration from Healy's proof [Hea08] of the Expander-Chernoff Theorem and recursively bound the MGF by 'splitting' it into components on lower levels of X using the swap walks. Toward this end, for any $\ell \leqslant k$, define the partial MGF $z_\ell : X(\ell) \to \mathbb{R}$ on a face $t = \{v_1, v_2, \ldots, v_\ell\}$ as

$$z_\ell(v_1, v_2, \ldots, v_\ell) = \prod_{i=1}^{\ell} e^{r 1_A(v_i)} = e^{r|t \cap A|}.$$

Our goal is to bound $\mathbb{E}[z_k] \approx \mathbb{E}[z_1]^k$. To set up a recursion, we appeal to the elementary observation that a k-set $s \in X(k)$ can be sampled by first drawing a $\frac{k}{2}$-set $t_1 \in X(k/2)$, then $t_2 \in X_{t_1}(k/2)$ conditionally from its link. We set $s = t_1 \uplus t_2$. This allows us to 'split' z_k into two correlated copies of $z_{k/2}$:

$$\mathbb{E}_{s\in X(k)}[z_k(s)] = \mathbb{E}_{s\in X(k)}[e^{r|s\cap A|}]$$
$$= \mathbb{E}_{t_1\uplus t_2}[e^{r|t_1\cap A|}e^{r|t_2\cap A|}]$$
$$= \mathbb{E}_{t_1\uplus t_2}[z_{k/2}(t_1)z_{k/2}(t_2)].$$

The trick is now to observe that (t_1, t_2) is distributed exactly as an edge in the swap walk $S_{\frac{k}{2},\frac{k}{2}}$. Since swap walks on HDX have excellent spectral expansion [AJT19; DD19; GLL22; AP23], we can 'de-correlate' the above and write the relation

$$\mathbb{E}[z_k] \leqslant (1-\gamma)\,\mathbb{E}[z_{k/2}]^2 + \gamma\mathbb{E}[z_{k/2}^2] \tag{2.2}$$

where γ is the expansion of the swap walk.

We could now recurse if not for the error term $\gamma\mathbb{E}[z_{k/2}^2]$, which could very well be the dominating factor. The key is show $\mathbb{E}[z_{k/2}^2]$ actually can't be too much bigger than $\mathbb{E}[z_k]$ itself:

$$\mathbb{E}[z_{k/2}^2] \lesssim e^{r^2k}\,\mathbb{E}[z_k]. \tag{2.3}$$

Setting r such that $\gamma e^{r^2k} \ll 1$, (2.2) becomes

$$\mathbb{E}[z_k] \lesssim \frac{1-\gamma}{1-e^{r^2k}\gamma}\,\mathbb{E}[z_{k/2}]^2 \approx \mathbb{E}[z_{k/2}]^2,$$

allowing us to recurse to show an upper bound similar to (2.1). For 'weak' HDX, this strategy requires setting $r = \Theta(\varepsilon k^{-1/2})$, while for strong HDX we may set $r = \Theta(\varepsilon)$. This results in the gap in the stated bounds.

To prove (2.3) we in some sense 'reverse' the sampling process above, and draw a $\frac{k}{2}$-set t by first sampling a k-set $s \in X(k)$, then take $t \subset s$ uniformly at random. This means we can write the error term as:

$$\mathbb{E}_{t\in X(k/2)}[z_{k/2}(t)^2] = \mathbb{E}_{s\in X(k)}\left[\mathbb{E}_{t\subseteq s}\left[z_{k/2}(t)^2\right]\right] = \mathbb{E}_{s\in X(k)}\left[\mathbb{E}_{t\subseteq s}\left[e^{2r|t\cap A|}\right]\right].$$

Crucially, the inner expectation is now over the $\frac{k}{2}$-uniform *complete complex* (whose vertices are the k-set s), so standard concentration implies the exponent, $2r|t \cap A|$, is close to its expectation $r|s \cap A|$ up to a factor of roughly r^2k with high probability. Thus:

$$\mathbb{E}_{s\in X(k)}\left[\mathbb{E}_{t\subseteq s}\left[e^{2r|t\cap A|}\right]\right] \lesssim \mathbb{E}_{s\in X(k)}\left[e^{r|s\cap A|+r^2k}\right] = e^{r^2k}\mathbb{E}_{s\in X(k)}[z_k].$$

Part II: Bootstrapping. We now argue one can lift Chernoff for HDX to optimal sampling for their inclusion graphs $(X(k), X(i))$. The idea, based on the method in [IKW12] for the complete complex, is to try to reduce to a $\frac{k}{i}$-uniform complex Y whose vertices are i-sets in $X(i)$, and whose $\frac{k}{i}$-faces correspond in some way to the k-sets in X. We could then hope to apply Chernoff to this system to bound the k vs. i sampling behavior in the original complex.

Somewhat more formally, we'll consider the $\frac{k}{i}$-uniform complex generated by the following process: draw a random k-face $\{v_1, \ldots, v_k\} \in X(k)$, and partition the face randomly into $\frac{k}{i}$ subsets of size i denoted $\{I_1, \ldots, I_{\frac{k}{i}}\}$. The resulting complex, called the *faces complex* of X, inherits local-spectral expansion from X's swap walks so we can apply Chernoff for HDX to get concentration of the form $\exp(-\varepsilon^2\frac{k}{i})$.

Unfortunately, this does not prove concentration for quite the right object. In particular, given a function $f : X(i) \to [0,1]$, the above really states that a random partitioning of a k-face s into $\frac{k}{i}$ sub i-faces satisfies

$$\mathbb{P}_{s\in X(k),\, \cup_j I_j=s}\left[\left|\frac{i}{k}\sum_{j=1}^{k/i}f(I_j)-\mu\right|>\varepsilon\right] \leqslant \exp\left(-\varepsilon^2\frac{k}{i}\right),$$

whereas we'd like to bound $\underset{s \in X(k)}{\mathbb{P}} \left[\left| \underset{t \subset s}{\mathbb{E}}[f(t)] - \mu \right| > \varepsilon \right]$. Here we are saved by the fact that each I_j is *marginally* distributed as a random i-set $t \subset s$. We prove in such cases it is possible to inherit sampling from the correlated bound up to a small loss in parameters, completing the proof.

2.3 Reverse Hypercontractivity

Leveraging Theorem 2.1 *locally* in X, we prove reverse hypercontractivity for high dimensional expanders.

Theorem 2.5 (Reverse Hypercontractivity (Informal Theorem 6.1)). *Fix $\rho \in (0,1)$ and let X be a k-uniform nice complex for k sufficiently large. Then there exist constants C, q (depending only on ρ) such that*

$$\langle f, T_\rho g \rangle \geqslant C \|f\|_q \|g\|_q.$$

In other words T_ρ is "(C, q, q')-reverse hypercontractive.[3] "

The core of Theorem 2.5 is really a version of the result for the down-up walk which simply states any two subsets $A, B \subset X(k)$ of non-trivial size must have correspondingly many edges between them.

Lemma 2.6 (Reverse Hypercontractivity of the Down-Up Walk (Informal Theorem 6.4)). *Fix $\gamma \in (0,1)$ and let X be a k-uniform nice complex for k sufficiently large. There exist constants c, q (depending only on γ) such that for any $A, B \subset X(k)$ of measure at most $\exp(-ck)$:*

$$\underset{s, s' \sim U_{\gamma k, k} D_{k, \gamma k}}{\mathbb{P}} [s \in A, s' \in B] \geqslant \mathbb{P}[A]^q \, \mathbb{P}[B]^q.$$

Note that the assumption on the set size in Lemma 2.6 is essentially tight. Even if one walks from s to s' while going down to a single vertex it is possible to find sets A and B of size roughly $\exp(-k)$ which are totally disconnected, even on the complete complex.

Proof Overview. The proof of Lemma 2.6 relies on the following simple observation: the down-up walk samples an edge by first sampling $t \in X(\gamma k)$ and then *independently* sampling $s, s' \in X(k)$ containing t, so we can express the γ-correlated mass of A and B as:

$$\underset{s, s' \sim U_{\gamma k, k} D_{k, \gamma k}}{\mathbb{P}} \left[s \in A, s' \in B \right] = \underset{t \in X(\gamma k)}{\mathbb{E}} \left[\underset{s}{\mathbb{P}} \left[A \mid s \supseteq t \right] \underset{s}{\mathbb{P}} \left[B \mid s \supseteq t \right] \right].$$

Now the connection with inclusion sampling becomes clear—if even a non-negligible fraction of $t \in X(\gamma k)$ see a reasonable portion of A and B simultaneously, the right-hand side should be large!

Naturally our first thought might be to appeal to Theorem 2.1 which bounds this type of behavior, but when $i = \gamma k$ the resulting bound is far too weak to be useful. In particular, after swapping from additive to multiplicative sampling and 'flipping' the graph, Theorem 2.1 roughly states that $(X(i), X(k))$ is an (α, β, δ)-sampler for $\alpha \approx \exp(-\delta^2 \beta \frac{k}{i})$. When $i = \Theta(k)$, this only gives a guarantee for sets of constant size.

The key is to realize that because X expands *locally*, we don't have to sample $t \in X(\gamma k)$ all at once. Instead, we'll sample t in 'steps' as $t = t_1 \cup \ldots \cup t_m$ where each t_j is sampled conditionally (i.e. from the link of) its predecessors, and apply Theorem 2.1 to each t_j individually. Toward this end, assume $\mathbb{P}[A] \leqslant \mathbb{P}[B]$ and fix the 'step size' $\ell = |t_j|$ roughly such that $\exp(-c' \frac{k}{\ell}) = \mathbb{P}[A]$ for some sufficiently small $c' > 0$.

Drawing our first component t_1, observe for small enough c' we have set our parameters such that Theorem 2.1 promises $(X(\ell), X(k))$ is an $(\alpha, \frac{1}{4}, \frac{1}{2})$-sampler for $\alpha \leqslant \min\{\mathbb{P}[A], \mathbb{P}[B]\}$, so with probability at least $\frac{1}{2}$ the conditional measures of A and B in t_1 are at least $\frac{1}{2} \mathbb{P}[A]$ and $\frac{1}{2} \mathbb{P}[B]$ respectively. Moreover, moving into the link of t_1 we can apply exactly the same process to t_2 (and so forth), so by the i-th step one can inductively show that with probability at least 2^{-i}, the conditional mass of A and B are at least $2^{-i} \mathbb{P}[A]$ and $2^{-i} \mathbb{P}[B]$ respectively. On the other hand, we only take $m = \frac{\gamma k}{\ell} = O(\log(\mathbb{P}[A]^{-1}))$ total steps by our choice of ℓ, so at the end of the procedure we have that with probability at least $2^{-m} = \mathbb{P}[A]^{O(1)}$, the density of A and B in t is at least $\mathbb{P}[A]^{O(1)}$ as well. Putting everything together, we have

$$\underset{s, s' \sim U_{\gamma k, k} D_{k, \gamma k}}{\mathbb{P}} \left[s \in A, s' \in B \right] = \underset{t \in X(\gamma k)}{\mathbb{E}} \left[\underset{s}{\mathbb{P}} \left[A \mid s \supseteq t \right] \underset{s}{\mathbb{P}} \left[B \mid s \supseteq t \right] \right]$$

---

[3] Here q' is the Hölder Conjugate of q. See Section 4 for a more detailed background on reverse hypercontractivity.

$$\geqslant 2^{-m} \cdot (2^{-m} \, \mathbb{P}[A]) \cdot (2^{-m} \, \mathbb{P}[B])$$
$$\geqslant \mathbb{P}[A]^{O(1)} \, \mathbb{P}[B]$$

as desired.

Once one has Lemma 2.6, Theorem 2.5 follows by passing to discrete approximations of f and g that carefully balance the quality of approximation (namely regarding the mass of f and g and pointwise closeness) and coarseness (number of discrete values in the approximation). Done correctly, one can then divide the functions into boolean level-sets and apply Lemma 2.6 to achieve full reverse hypercontractivity at the cost of a slight decrease in the corresponding norm and an additional constant factor.

2.4 Optimality and the Trickling-Down Threshold

Oppenheim's Trickling Down Theorem is among the most fundamental results in the theory of spectral high dimensional expansion [Opp18]. It states that any connected complex whose top links have expansion *strictly* better than $\frac{1}{d-1}$ exhibits 'local-to-global' behavior: one can immediately give quantitative bounds on the spectral expansion of *all* links of X (including its 'global' 1-skeleton), as well as on the down-up walks [AL20]. Our results are in some sense a strengthening of the local-to global behavior of Oppenheim: essentially as soon as one passes this threshold, not only can one infer 'Chernoff'-type concentration on skeletons of X, X itself actually satisfies exponential concentration for any Lipschitz function. This is in stark contrast to complexes *at* the TD-Barrier, which we observe may have no concentration properties whatsoever:

Theorem 2.7 (Lower Bounds at the TD-Barrier (Proposition 11.7)). *For every $\beta < 1$ and $k \in \mathbb{N}$, there exists a family of k-uniform 1-TD complexes $\{X_n\}$ such that $(X_n(k), X_n(1))$ is not a $(\frac{1}{2}, \beta)$-additive sampler.*

Theorem 2.1 also gives essentially the best possible parameters for inclusion sampling on any complex.

Theorem 2.8 (Inclusion Sampling Lower Bounds (Informal Theorem 11.1)). *Let $i < k \in \mathbb{N}$, $\varepsilon \in (0, 0.1)$, and let X be a k-uniform complex, then one of the following holds:*

1. *$(X(k), X(i))$ is not an (ε, β)-sampler for $\beta \leqslant \exp\left(-O\left(\varepsilon^2 \frac{k}{i}\right)\right)$.*

2. *$(X(k), X(i))$ is not an $(\frac{1}{10i}, \frac{\varepsilon}{6i})$-sampler.*

Classical lower bounds on samplers (e.g. [CEG95]) are based on *degree* and are far from tight in our setting where inclusion structure is the main barrier. Nevertheless, we are able to argue our bounds are optimal by reducing to $i = 1$ where the degree bound is tight. The idea is to show any set A which is a counter-example to sampling of $(X(k), X(1))$ can be 'lifted' to a counter example to $(X(k), X(i))$ by taking A' to be the family of i-sets that hit A. This works when the measure of A' is roughly $i\mu(A)$, and therefore when 1) A is small, and 2) $(X(k), X(i))$ is a sampler. We show the original technique of [CEG95] can be used to analyze sets of any density, thereby giving the desired counter-example.

2.5 Agreement Testing

Leveraging Theorem 2.5, we prove several new agreement theorems in both the 99% and 1%-regimes. In the former we consider a general setup of [DFH19]. Given a complex X and an ensemble of functions $\mathcal{F} = \{f_s : \binom{s}{i} \to \mathbb{F}_2\}_{s \in X(k)}$, we'd like to test whether $\mathcal{F}$ 'comes from' a global function $g : X(i) \to \mathbb{F}_2$. To be concrete, consider $i = 1$. Here we are given $\mathcal{F} = \{f_s : s \to \mathbb{F}_2\}$ and can interpret each f_s as the indicator of some subset $r_s \subseteq s$. We want to test whether there is a 'global' subset $R \subseteq X(1)$ that approximately determines most local subsets in the sense that $r_s \approx R \cap s$. For $i = 2$, a similar interpretation holds but for *subgraphs* of X. Namely, think of $f_s : \binom{s}{2} \to \mathbb{F}_2$ now as a subgraph of s, i.e. corresponding to the indicator of an edge-set $e_s \subseteq \binom{s}{2}$. Now 'coming from' a global function means there is a global subgraph $E \subseteq X(2)$ such that for most local graphs, $e_s \approx E \cap s$. For $i > 2$ one interprets this similarly as a question about local vs. global sub-complexes of X. Such 'subcomplex' tests have proved useful both in boolean analysis [DFH19] and as a tool for analyzing the more challenging 1%-regime [DD23b; DDL24].

We study the classical V-test in this context, which draws a pair of k-faces $(s, s') \sim U_{k/2,k} D_{k,k/2}$ intersecting on $k/2$ vertices and checks if s and s' agree on most i-faces in the intersection. Let $Agree_\eta^V(\mathcal{F})$

denote the probability (s, s') agree on at least a $1 - \eta$ fraction of their shared i-faces. Similarly, for any global function $g : X(i) \to \mathbb{F}_2$, write $f_s \overset{\eta}{\approx} g|_s$ if f_s agrees with g on a $1 - \eta$ fraction of its i-faces. Then:

Theorem 2.9 (The Subcomplex V-Test, 99%-Regime (Informal Theorem 8.1)). *Let X be a k-uniform nice complex. For any $\eta, \varepsilon > 0$ if $Agree_\eta^V(\mathcal{F}) \geqslant 1 - \varepsilon$:*

$$\exists g : X(1) \to \mathbb{F}_2, \quad \underset{s \in X(k)}{\mathbb{P}}[f_s \overset{\eta'}{\approx} g|_s] \geqslant 1 - \varepsilon'.$$

where $\eta' \leqslant O(\eta + \varepsilon)$ and $\varepsilon' \leqslant \varepsilon^{O_\eta(1)} + e^{-\Omega_\eta(k)}$.

In the '1%'-regime, our goal is to infer global structure of $\mathcal{F}$ even when the test passes with small (but non-negligible) probability. We focus on the $i = 1$ setting and aim to construct a test $\mathcal{T}$ that infers global structure whenever $Agree^\mathcal{T}(\mathcal{F}) \geqslant \exp(-\Omega(k))$, the best possible bound since a random function passes any such test with this probability. To this end, we consider a variant of [IKW12]'s Z-test on high dimensional expanders which samples a triple $(\sigma, \sigma', \sigma'') \in X(k)$ roughly such that (σ, σ') is distributed as the V-test, and σ'' is drawn from the link of $\sigma' \setminus \sigma$. The Z-test passes if (σ, σ') and (σ', σ'') are consistent on their intersections. We prove that the Z-test is sound under the stronger assumption that the complex is "λ-global", meaning the swap walk $S_{\frac{k}{2}, \frac{k}{2}}$ is λ-close to its stationary distribution in ℓ_∞-norm (see Definition 8.6).

Theorem 2.10 (The Z-Test, 1%-Regime (Informal Theorem 8.7)). *$\forall \lambda, \eta > 0$ and large enough k, let X be a λ-global k-uniform nice complex. Then for any $\delta \in (8\lambda + e^{-\Omega(\eta k)}, \frac{1}{8})$ if $Agree_0^Z(\mathcal{F}) \geqslant \delta$:*

$$\exists g : X(1) \to \mathbb{F}_2, \quad \underset{s \in X(k)}{\mathbb{P}}[f_s \overset{\eta}{\approx} g|_s] \geqslant \delta/8.$$

We will not give a true proof overview of our testing results (though it should be mentioned that the high level strategy closely follows [IKW12; DL17]), but wish to highlight that the core of Theorem 8.7 is really a 'weak' *local* agreement theorem for the V-test on nice HDX of independent interest. The statement is quite technical, so we give a very informal version here that captures it in spirit.

Theorem 2.11 (The Local Agreement Theorem (Informal Theorem 8.12)). *Let X be a nice complex. Then for any $\delta \geqslant \exp(-\Omega(k))$ if $Agree_0^V \geqslant \delta$, for an $\Omega(\delta)$-fraction of $t \in X(k/2)$ there is a 'smoothing' of $\mathcal{F}$ which maintains its structure and has $(1 - \delta^2)$-agreement on the link X_t.*

In other words, any non-trivial agreement of the V-test *must* come from the fact that $\mathcal{F}$ is locally consistent with a global function. It is possible Theorem 2.11 could be propogated to a true Z-test under much weaker conditions than λ-globality, e.g. under the recent topological notions of [GK22; DD23b; BM23b]. We leave this as an open question for the 1%-regime.

2.6 Further Applications

Finally, we discuss our applications of Theorem 2.1 and Theorem 2.5 beyond agreement testing.

Geometric Overlap. A complex X has (d, c)-geometric overlap if for every embedding $X(1)$ into $\mathbb{R}^{d-1}$, there is a point $q \in \mathbb{R}^{d-1}$ that lies in the convex hulls of at least a c-fraction of X's embedded d-faces. [Fox+12] proved every d-uniform bounded degree complex has geometric overlap at best c_d, where c_d is the overlap of the complete complex. We prove sufficiently strong high dimensional expanders match this bound.

Theorem 2.12 (Geometric Overlap (Informal Theorem 9.13)). *For every $d \in \mathbb{N}$ and $\varepsilon > 0$, there exists $\lambda > 0$ such that any λ-two-sided HDX with uniform vertex weights has $(d, c_d - \varepsilon)$-geometric overlap.*

We remark that while geometric overlap is a classical problem in mathematics (dating back to [BF84]), we are not aware of any applications in computer science. Nevertheless other instances of overlap theorems have proven powerful (see e.g. [Cha+23]), so it seems plausible such results may be of future use.

Double Samplers. Double samplers are a powerful variant of sampler graphs used in [Din+18a] to construct good list-decodable codes and in [DW22] for the heavy hitters problem. Roughly speaking, double samplers are 'three-wise' inclusion structures $(X(k), X(j), X(1))$ such that

1. $(X(k), X(j))$ is a $(1/2, \beta)$-additive sampler

2. For all $s \in X(k)$, the restriction of $(X(j), X(1))$ to vertices in s is a $(1/2, \beta)$-additive sampler

The best known prior construction of double samplers [Din+18a] had overhead $\frac{|X(k)|}{|X(1)|} = \exp(\text{poly}(\beta^{-1}))$, leading to poor rate of their resulting list-decodable codes. One of the main open questions asked in [Din+18a] was to determine the optimal overhead of a double sampler.

We nearly resolve this problem for 'typical' complexes X which are non-contracting (i.e. $|X(j)| > |X(i)|$ whenever $j > i$), and satisfy a weak 'hitting-set' type guarantee. In particular, Theorem 2.1 gives double samplers with only *quasi-polynomial overhead*, while for such typical complexes Theorem 2.8 implies a corresponding near-matching lower bound.

Theorem 2.13 (Near-Optimal Double Samplers (Informal Theorem 9.3))**.** *For every $\beta > 0$, there exists an infinite family of double samplers with $\exp(O(\log^6(\beta^{-1})))$ overhead. Moreover, any 'typical' double sampler has overhead at least $\exp(\Omega(\log^2(\beta^{-1})))$.*

We refer the reader to Section 9.1 for the exact restrictions on X for which the lower bound holds.

Locally Testable and List-Decodable Codes. While it is likely applying our double sampler machinery to the arguments of [Din+18a] would give list-decodable codes with substantially improved rate, this was already achieved subsequent to [Din+18a]'s work using walks on expanders [Ale+20]. Instead, using a variant of our machinery we give a closely related application to distance amplification for locally testable codes with near-optimal distance-alphabet trade-off. The resulting codes are list-decodable and preserve testability up to a log factor in the alphabet.

Theorem 2.14 (Large Distance List-Decodable LTCs (Informal Theorem 10.21))**.** *For all large enough $k \in \mathbb{N}$ and $\varepsilon > 0$, there exists an explicit family of $\mathbb{F}_2$-linear codes with*

1. *Distance: $1 - 2^{-k} - \varepsilon$*

2. *Alphabet Size: 2^k*

3. *Rate: $\varepsilon^{O(k)}$*

Moreover, the codes are efficiently list decodable up to distance $1 - 2^{-\Omega(k)}$ and locally testable in $O(1)$ queries with soundness $\Omega(\frac{1}{k \log(1/\varepsilon)})$.

Prior distance amplification techniques for LTCs [Kop+17] have exponentially worse distance-alphabet trade-off, but improved rate. To our knowledge the above are the first large distance list-decodable LTCs. We also give a candidate technique for amplifying distance while maintaining *constant* soundness (removing the k factor in the denominator above) using HDX, and prove the method works on the complete complex. We refer the reader to Section 10 for further details on these notions.

Degree Lower Bounds for HDX. One of the formational results in the study of expansion is the Alon-Boppana theorem [Nil91], roughly stating any family of bounded-degree λ-expanders have degree at least $\frac{2}{\lambda^2}$. Despite degree being a critical parameter in application (degree controls the 'blow-up' induced by using an HDX as a gadget), in high dimensions we understand very little about optimal degree. The best known constructions of λ-TD complexes have degree $\lambda^{\tilde{O}(d^2)}$. Is this optimal? Could $\lambda^{o(d^2)}$ be achieved?

We take the first step toward answering this question. Combining our sampling results with degree lower bounds of [CEG95], we show super-exponential lower bounds for certain special classes of HDX, including polynomial skeletons of any hyper-regular λ-TD complex.

Theorem 2.15 (Degree Lower Bounds (Informal Theorem 9.8)). *Fix $\lambda < 1$, $d \in \mathbb{N}$, and $k \leqslant \sqrt{d}$. Let X be the k-skeleton of a d-uniform hyper-regular λ-TD complex. Then*

$$deg(X) \geqslant 2^{\Omega_\lambda(k^2)}$$

Theorem 2.15 exhibits yet another threshold phenomenon at the TD-barrier, since there exist 1-TD complexes [Gol21] which have exponential degree at any level. While the latter are not hyper-regular, they are still reasonably balanced and our techniques extend to this regime.[4]

If one could prove a true Chernoff bound for λ-TD complexes (rather than our concentration of $\exp(-\sqrt{k})$), Theorem 2.15 could be improved to showing degree $2^{\Omega(d^2)}$ at the *top level* of X, albeit still under the assumption of hyper-regularity. While bounded-degree hyper-regular HDX do exist [FI20], their degree is substantially worse and removing this constraint from our lower bound remains an important open problem.

Separating MLSI and Reverse Hypercontractivity. The Modified Log-Sobolev Inequality (MLSI) is a powerful analytic inequality used to bound the mixing time of Markov chains and previous weakest sufficient condition for reverse hypercontractivity [MOS13]. To the best of our knowledge, it was open until this work whether MLSI was *necessary*. We resolve this question for the setting allowing a leading constant $C > 1$.

Corollary 2.16 (Separating RHC and MLSI (Informal Corollary 9.17)). *There exist constants C, q, q' such that for infinitely many $N \in \mathbb{N}$, there exist (C, q, q')-reverse hypercontractive operators on N vertices with vanishing MLSI constant:*

$$\rho_{MLSI} \leqslant \tilde{O}\left(\frac{1}{\log(N)}\right)$$

This separation is essentially the strongest possible for operators with constant expansion, since any such operator has LSI (and therefore MLSI) at least $\Omega(\frac{1}{\log(N)})$ [DS96].

A Frankl–Rödl Theorem. The Frankl–Rödl Theorem [FR87] is a broadly applied result in extremal combinatorics bounding the independence number of the graph on $\{0,1\}^n$ connecting strings of fixed intersection size. Lemma 2.6 implies an analogous result for the γk-step down-up walk on HDX.

Corollary 2.17 (Frankl–Rödl for HDX (Informal Claim 9.19)). *Fix $\gamma > 0$ and let X be a k-uniform nice complex for k sufficiently large. The γk-step down-up walk has independence number*

$$\alpha \leqslant \exp(-\Omega_\gamma(k)) \cdot |X(k)|$$

In comparison to Frankl–Rödl, the above bounds the size of sets avoiding intersections of size *at most* instead of *exactly* $(1 - \gamma)k$. A similar issue appears when using reverse hypercontractivity toward this end in the cube and is handled by a separate Fourier analytic argument in [Ben12]. It would be interesting to extend our results to the exact case to give a true analog of Frankl–Rödl.

'It Ain't Over Till It's Over'. Reverse hypercontractivity was used to resolve Friedgut and Kalai's 'It Ain't Over Till It's Over' Conjecture [MOO05], which bounds the tail of random restrictions in the hypercube. The core of their proof was an analog for the closely related noise operator previously used in analysis of non-interactive correlation distillation protocols [Mos+06]. We generalize the latter theorem to HDX:

Theorem 2.18 (Noise Operator Tail Bounds (Informal Theorem 9.18)). *Fix $\rho \in (0, 1)$ and $k \in \mathbb{N}$ sufficiently large. For any k-uniform nice complex X and $f : X(k) \to [0, 1]$ of density μ:*

$$\mathbb{P}\left[T_\rho f \notin [\delta, 1 - \delta]\right] \leqslant \delta^{O_{\mu,\rho}(1)}$$

---

[4]Namely, we only really require that the underlying graph of every link does not have a (roughly) $\exp(-\sqrt{d})$ fraction of vertices making up $1/2$ the measure. Unfortunately, constructions such as the Ramanujan complex have extremely unbalanced links, and actually do fail this property. See Section 9.2 for further discussion.

3   Related Work

High Dimensional Expanders.   Spectral high dimensional expansion was developed over a series of works [EK16; KM17; Opp18; DK17] building on prior notions of high dimensional expansion in topology [Gar73; LSV05b; LM06; Gro10; KL14; KKL14]. Basic higher order random walks were introduced by Kaufman and Mass [KM17], with the swap operators later introduced independently in [DD19; AJT19]. A great deal of work has been done at the intersection of high dimensional expanders and analysis, including determining the optimal spectral gap of the down-up walks [KM17; KO20; Dik+18; AL20; KT21], developing a general theory of Fourier analysis [KM20; Dik+18; Gai+22; Baf+22b; GLL22; GK23a], and applications thereof to agreement testing [DK17; DD19; KM20; GK22], mixing of Markov Chains [Ana+19; ALG21; CLV21; Che+21; Liu21; Bla+22; Fen+22; Ana+23] (among many others), and algorithms [DW22; Baf+22a; Bea+23].

Two works [KS20; KKS21] in the HDX literature also study Chernoff-style bounds, albeit in very different regimes from our results. Kaufman and Sharakanski [KS20] study concentration for global functions over repeated random walks on k-faces, which is incomparable to our setting. Kaufman, Kyng, and Song [KKS21] prove scalar and matrix Chernoff bounds under ℓ_∞-*independence*, a strengthening of spectral HDX that only holds for dense complexes. While in this work we are primarily interested in proving concentration for *bounded-degree* complexes, it is interesting to ask whether their stronger result (Chernoff for matrix-valued functions) extends to the sparse regime.

In [Baf+22b; GLL22], the authors prove a notion called "global" hypercontractivity for high dimensional expanders. While it is known that reverse hypercontractivity follows from the standard hypercontractive inequality [MOS13], we are not aware of a reduction from the weaker global variant. Our work also differs substantially from [Baf+22b; GLL22] in its tools and regime of application.

Samplers and Chernoff-Hoeffding.   Samplers are among the most classical tools in theoretical computer science, see e.g. [Gol97; Sha04; Wig09] for surveys of their many constructions and applications. De-randomized Chernoff bounds are an important sub-family of sampler graphs that exhibit optimal tails in the degree of the graph. Sparse examples were constructed in [AKS87; Gil98] using walks on expanders, and there is a great deal of literature toward understanding what general families of hypergraphs admit such concentration, e.g. under limited independence [SS89; SS90; SSS95; Pem01], negative correlation [DR96], for edge colorings [PS97], under Dobrushin uniqueness [Pau14], for strongly Raleigh distributions [KS18], and most recently under ℓ_∞-independence [KKS21]. Such bounds have many interesting applications beyond those discussed in this work.

Inclusion samplers were first introduced in Impagliazzo, Kabanets, and Wigderson's [IKW12] work on agreement testers and PCPs with strong soundness. Prior to Theorem 2.1, there were only two known systems of inclusion samplers with optimal tails: the complete complex [IKW12], and curves [Mos17]. Both examples were used toward the construction of PCPs with strong soundness, but failed to achieve optimal parameters due to their blow-up in size. Several sparse variants of inclusion samplers were known to admit weaker 'Chebyshev' type bounds, including high dimensional expanders [DK17] and the Grassmann [IKW12]. The latter also lead to PCPs with strong soundness [DM11], but failed to achieve optimal parameters due to its polynomial tail. Theorem 2.1 is the first to achieve the best of both worlds, though there remain substantial barriers toward its use for PCPs.

Reverse Hypercontractivity.   Reverse hypercontractivity was first shown by Borell [Bor82] and first utilized in theoretical computer science by [Mos+06] who observed the inequality implies fine-control of mixing between sets (akin to Lemma 2.6). The result has since found great use in hardness of approximation [MOO05; FKO07; She09; Ben12; Kau+14], social choice [MR12; Kel12; Mos12], and extremal combinatorics [Ben12; Kau+14]. In 2013, [MOS13] extended known bounds on reverse hypercontractivity to general product spaces and settings with bounded modified Log-Sobolev Inequalities, a surprising result given the failure of standard hypercontractivity in these settings. Dinur and Livni-Navon [DL17] were the first to apply these results to agreement testing, where they resolved the conjecture of [IKW12] regarding exponential soundness for the Z-test and built the first 'combinatorial' PCPs with optimal soundness.

We note that MLSI bounds (and therefore RHC) are known for many 'dense' settings of HDX studied in the sampling literature (see e.g. [CLV21]), with the most general prior condition being the notions of entropic

independence and fractional log-concavity [Ana+22] which are significant strengthenings of spectral HDX. These methods all necessarily rely on density and cannot capture our theory.

Agreement Tests. Agreement testing, also known as direct product testing, is a powerful tool in the construction of PCPs and locally testable codes [RS96; AS97; RS97; GS97; DR06; Din07; Dik+20; Din+22b]. Agreement tests in the 99%-regime were studied by [GS97] and [DS14] in the complete complex, and extended to high dimensional expanders in [DK17; DD19; KM20]. [DFH19] gave the first 'subcomplex' agreement theorem in the 99%-regime (the setting of Theorem 2.9) in the complete complex. Agreement testing in the 1%-regime was studied in the complete complex in [DG08a; IKW12; DL17], with the lattermost work giving optimal bounds. [IKW12] also gave a de-randomized test over the Grassmann complex, which is polynomial size, but suffers from exponentially worse soundness. A different 1%-regime test on the Grassmann was also used to resolve the 2-2 Games Conjecture [KMS17; Din+18b; BKS19].

Finally, [GK22; BM23b; DD23b; BLM24; DDL24] studied the 1% (or closely related *list* agreement) regime on high dimensional expanders. These works identified key *topological* properties any complex must exhibit to be a good tester, and subsequently constructed complexes satisfying these notions, giving the first bounded degree 1%-agreement testers. The methods of [DD23b; DDL24] in particular rely on Theorem 2.1 and Theorem 2.9. The soundness achieved by the above works, however, is inverse *logarithmic* in the dimension instead of inverse exponential (or even polynomial). Achieving inverse exponential soundness as in Theorem 2.10 for families of bounded degree complexes remains an important open question.

Open questions

1. While we are able to show optimal concentration for a fairly broad class of high dimensional expanders, in the weakest settings (namely under spectral independence and at the TD-Threshold), we are only able to prove exponential concentration at the top level. It is unclear whether this is a fundamental or purely technical barrier: do such complexes satisfy a true Chernoff bound? As discussed above, a resolution of this question in the positive leads to better degree lower bounds for HDX, and in particular a $2^{\Omega(d^2)}$ lower bound for the top level of hyper-regular λ-TD complexes.

2. The best known constructions of high dimensional expanders are not hyper-regular. Is it possible to remove this constraint from our degree lower bound? While our technique holds even for 'reasonably balanced' complexes, it cannot handle objects like the Ramanujan complexes have extremely unbalanced links. It seems likely this is a technical rather than inherent barrier, and we conjecture some finer notion of concentration or application thereof may be able to remove this constraint.

3. In Section 7, our proof of inclusion sampling on partite high dimensional expanders requires moving to a slightly lower dimensional skeleton. In fact, we are unable to prove optimal inclusion sampling at the top level even for product spaces. In Appendix B, we show this (and the general bound for HDX) would be implied by proving optimal entropy contraction of the Kneser graphs. Does such a bound hold? Are product spaces optimal inclusion samplers?

4. We show that 'standard' simplicial complexes used in the construction of double samplers cannot achieve better than quasipolynomial overhead. Are there non-simplicial objects better served for this purpose?

5. Our 1%-regime test only holds on dense complexes due to the assumption of ℓ_∞-expansion. However, the main argument only requires reverse hypercontractivity and spectral gap of the down-up walk. Can the argument be completed without the assumption of ℓ_∞-expansion to give new sparse agreement testers in the low acceptance regime?

Roadmap

In Section 4 we give more detailed and formal preliminaries on high dimensional expanders, random walks, samplers, and reverse hypercontractivity. In Section 5 we give the formal statement of our optimal sampling theorem and prove the two-sided case, deferring the remaining cases to Section 7. In Section 6 we prove

reverse hypercontractivity. In Section 8 and Section 9 we give applications of reverse hypercontractivity to agreement testing, and analysis and combinatorics respectively. In Section 10 we give basic sampling lemmas for splitting trees and construct high distance list-decodable LTCs. Finally in Section 11 we show optimality of our sampling results in terms of requirements on X and strength of sampling.

4   Preliminaries and Notation

4.1   Graphs and Spectral Expansion

A weighted un-directed graph $G = (V, E, \mathbb{P})$ consists of a finite vertex set V, a set of edges $E \subseteq \binom{V}{2}$, and a distribution on the edges $\mathbb{P} : E \to (0, 1]$. For a vertex v we denote by $\mathbb{P}[v] = \frac{1}{2} \sum_{uv \in E} \mathbb{P}[uv]$. In this paper all graphs are assumed to be weighted (the weight is implicit in notation). We also assume that there are no isolated vertices.

Let $\ell_2(V) = \{f : V \to \mathbb{R}\}$. The graph G induces an inner product $\langle f, g \rangle = \mathbb{E}_v[f(v)g(v)]$ on $\ell_2(V)$, as well as a normalized adjacency operator $A : \ell_2(V) \to \ell_2(V)$ defined as

$$Af(v) := \mathop{\mathbb{E}}_{u \sim v}[f(u)] = \sum_{uv \in E} \frac{\mathbb{P}[uv]}{\sum_{w \sim v} \mathbb{P}[wv]} f(u).$$

It is well known that A is diagonalizable and has eigenvalues

$$1 = \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_{|V|} \geqslant -1,$$

where the first 'trivial' eigenvalue corresponds to the space of constant functions. A graph is called a spectral expander if all non-trivial eigenvalues are small.

Definition 4.1 (Spectral expander)**.** A graph $G = (V, E)$ is called a λ-*one-sided spectral expander* if $\lambda_2 \leqslant \lambda$. We say that G is a λ-*two-sided spectral expander* if $\lambda_2 \leqslant \lambda$ and $\lambda_{|V|} \geqslant -\lambda$.

We record the following inequality for λ-two-sided spectral expanders, a variant of the classical *expander-mixing lemma*. For every $f, g \in \ell_2(V)$,

$$|\langle f, Ag \rangle - \mathbb{E}[f]\,\mathbb{E}[g]| \leqslant \lambda \|f\|_2 \|g\|_2. \tag{4.1}$$

4.1.1   Bipartite Graphs and Bipartite Expanders

A bipartite graph is a graph where the vertex set can be partitioned to two independent sets $V = L \cup R$, called sides. We sometimes denote such graphs by $G = (L, R, E)$.

The Bipartite Adjacency Operator.   In a bipartite graph, we view each side as a separate probability space, where for any $v \in L$ (resp. R), $\mathbb{P}[v] = \sum_{w \sim v} \mathbb{P}[wv]$. We can define the *bipartite adjacency operator* as the operator $B : \ell_2(L) \to \ell_2(R)$ by

$$\forall f \in \ell_2(L), v \in R, \ Bf(v) = \mathop{\mathbb{E}}_{w \sim v}[f(u)]$$

where the expectation is taken with respect to the probability space L, conditioned on being adjacent to v. There is a similar operator $B^* : \ell_2(R) \to \ell_2(L)$ as the bipartite operator for the opposite side. As the notation suggests, B^* is adjoint to B with respect to the inner products of $\ell_2(L), \ell_2(R)$.

We denote by $\lambda(B)$ the spectral norm of B when restricted to $\ell_2^0(L) = \{\mathbf{1}\}^\perp$, the orthogonal complement of the constant functions (according to the inner product the measure induces on L). Namely

$$\lambda(B) = \sup \{\langle Bf, g \rangle \mid \|g\|, \|f\| = 1, f \perp \mathbf{1}_L \}.$$

Definition 4.2 (Bipartite Expander)**.** Let G be a bipartite graph, let $\lambda < 1$. We say G is a λ-*bipartite expander*, if $\lambda(B) \leqslant \lambda$.

It is easy to show that a bipartite graph is a λ-bipartite expander if and only if it is a λ-one-sided spectral expander. So we use these terms interchangeably on bipartite graphs.

We record the following inequality for λ-bipartite expanders similar to (4.1). For every $f \in \ell_2(L), g \in \ell_2(R)$,

$$|\langle f, Ag\rangle - \mathbb{E}[f]\,\mathbb{E}[g]| \leq \lambda\|f\|_2\|g\|_2. \tag{4.2}$$

4.2 Reverse Hypercontractivity

Hypercontractivity and reverse hypercontractivity are powerful analytic inequalities from boolean function analysis that bound the contraction behavior of operators between normed spaces. Recall that for $p \neq 0$ and a function f on a probability space, we denote by $\|f\|_p = \mathbb{E}[|f|^p]^{1/p}$ (for $p \leq 0$ this is only defined for f that are non-zero almost everywhere). We note that $p \mapsto \|f\|_p$ is monotone increasing.

Let V be a finite probability space and let $\ell_2(V) = \{f : V \to \mathbb{R}\}$. A monotone operator is an operator such that for any non-negative function $f : V \to [0, \infty)$, Af is also non-negative (i.e. for any $x \in X$, $Af(x) \geq 0$). For example, adjacency operators of graphs are always monotone.

Let A be a monotone operator. Let $1 \leq p < q$ and let $C > 0$. A is typically called (p, q, C)-*hypercontractive* if for every $f : V \to \mathbb{R}$,

$$\|Af\|_q \leq C\|f\|_p.$$

It is well know that this is equivalent to 'two-function hypercontractivity', that for every $f, g : V \to \mathbb{R}$

$$\langle Af, g\rangle \leq C\|f\|_p\|g\|_{q'}$$

where $q' = \frac{q}{q-1}$ is q's Hölder conjugate [ODo14]. *Reverse hypercontractivity* 'flips' this inequality in multiple ways: both in the direction of the inequality, and in the relation of p and q.

Definition 4.3 (Reverse hypercontractivity). Let $q < p < 1$ such that $p, q \neq 0$. Let $C > 0$. Let A be a monotone operator. We say that A is (p, q, C)-reverse hypercontractive if for every $f : V \to \mathbb{R}^{\geq 0}$,

$$\|Af\|_q \geq C\|f\|_p$$

or equivalently if for every $f, g : V \to \mathbb{R}^{\geq 0}$ it holds that

$$|\langle Af, g\rangle| \geq C\|f\|_p\|g\|_{q'}$$

where $q' = \frac{q}{q-1}$ is q's Hölder conjugate (the equivalence is similar to standard hypercontractivity).

It is sometimes convenient to substitute $f' := f^p, g' = g^{q'}$ and write the two-function reverse hypercontractivity inequality as

$$\left|\langle Af'^{1/p}, g'^{1/q'}\rangle\right| \geq \mathbb{E}[f']^{1/p}\,\mathbb{E}[g']^{1/q'}.$$

Probably the simplest (and most useful) interpretation of this inequality is when f and g are indicators of sets $M, N \subseteq V$ respectively, and that A is an adjacency operator of a graph (whose vertices are the points in V). Then this inequality has a combinatorial interpretation as a type of 'mixing' lemma. Namely that the probability a random edge hits M and N:

$$\mathbb{P}_{uv \in E}[u \in M, v \in N] \geq \mathbb{P}[M]^{1/p}\,\mathbb{P}[N]^{1/q'}.$$

This becomes a powerful tool in the regime where the sets are smaller than can be controlled by the expander-mixing lemma.

4.3 Simplicial Complexes

Simplicial complexes. A pure d-uniform simplicial complex X is a finite set system (hypergraph) consisting of an arbitrary collection of sets of size d together with all their subsets (individually called 'faces'). The sets of size i in X are denoted by $X(i)$, and in particular, the vertices of X are denoted by $X(1)$ (we do not distinguish between vertices and the singletons of vertices). We write $X^{\leq j}$ to denote faces of X up to size j.

We note that this notation departs somewhat from the typical convention in the literature which denotes faces of size $i+1$ by $X(i)$, and what we have called a d-uniform simplicial complex is more typically referred to as a $(d-1)$-*dimensional* simplicial complex. We have chosen to adopt the former notation since most theorems in this work require a substantial amount of arithmetic on set sizes, and the latter convention quickly becomes overly cumbersome.

Probability over simplicial complexes. Let X be a simplicial complex and $\mathbb{P}_d : X(d) \to (0,1]$ a density function on $X(d)$. This density function induces densities on lower level faces $\mathbb{P}_k : X(k) \to (0,1]$ by $\mathbb{P}_k(t) = \frac{1}{\binom{d}{k}} \sum_{s \in X(d), s \supset t} \mathbb{P}_d(s)$. Equivalently $\mathbb{P}_k$ is the density induced by drawing a d-face from $s \sim \mathbb{P}_d$, and a uniformly random k-face $t \subset s$.

When clear from the context, we omit the level of faces and just write $\mathbb{P}[T]$ or $\mathbb{P}_{t \in X(k)}[T]$ for an event $T \subseteq X(k)$.

Links of faces. Let X be a d-uniform simplicial complex. Let $k < d$ and $s \in X(k)$. The link of s (called a k-link) is a $(d-k)$-uniform simplicial complex defined by $X_s = \{t \setminus s \mid t \in X, t \supseteq s\}$. Note the link of the emptyset is X itself $X_\emptyset = X$.

Let $s \in X(k)$ for some $k \leqslant d$. The density function $\mathbb{P}_d$ on X induces a density function $\mathbb{P}^s_{d-k} : X(d-k) \to (0,1]$ on the link where $\mathbb{P}^s_{d-k}[t] = \frac{\mathbb{P}[t \cup s]}{\mathbb{P}[s]\binom{d}{k}}$. We usually omit s in the probability, and for $T \subseteq X_s(k)$ write $\mathbb{P}_{t \in X_s(k)}[T]$.

Connected Complexes. We call a complex connected if the graph underlying every link is connected.

Partite complexes. A d-partite simplicial complex is a d-uniform complex whose vertices can be partitioned into d disjoint sets (typically called "parts" or "colors")

$$X(0) = X[1] \uplus X[2] \uplus \ldots \uplus X[d]$$

such that every $s \in X(d)$ has exactly one vertex from each part.

Let X be a d-partite simplicial complex. The *color* of a face $t \in X(k)$ is $col(t) = \{i \in [d] \mid t \cap X[i] \neq \emptyset\}$. Let $F \subseteq [d]$. We denote by $X[F] = \{s \in X \mid col(s) = F\}$ the faces of X with color F. The projection of the *complex* X onto F is the sub-simplicial complex $X^F = \{s \in X \mid col(s) \subseteq F\}$. When X is endowed with a density $\mathbb{P}_d$, X^F has a naturally induced density $\mathbb{P}^F$ by sampling $\sigma \sim \mathbb{P}_d$, and outputting the projection σ_F. Quantitatively, the induced distribution can be written as

$$\mathbb{P}^F(s) = \sum_{t \in X(d), s \subseteq t} \mathbb{P}[t].$$

Finally, when $F = \{i\}$ is a singleton we just write $X[i]$, X^i, and x_i for brevity.

Given a generic complex X, it will often be useful to consider X's 'partitification' $P = P(X)$, which simply includes every possible ordering of the faces of X as tuples. Formally, endow the faces of X with an arbitrary order and define

$$P(k) := \left\{ (s, \pi) := \left\{ (s_{\pi(1)}, 1), (s_{\pi(2)}, 2), \ldots, (s_{\pi(k+1)}, k) \right\} \ \middle| \ s \in X(k), \pi \in S_k \right\},$$

where S_k is the group of permutations on k-letters and the measure of a face (s, π) is inherited naturally as $\frac{1}{k!} \mathbb{P}_X[s]$. Note that $P(X)$ does not depend on the choice of ordering, which is simply a notational convenience.

Degree of Simplicial Complexes. Degree is a critical parameter of simplicial complexes capturing the (local) 'blow-up' incurred by moving to higher uniform faces. We define the (max) degree of a complex with respect to i-faces as:

$$\deg^i(X) = \max_{v \in X(1)} |\{s \in X(i) : v \in s\}|$$

16

We write just $\deg(X) := \deg^d(X)$ to denote degree with respect to top level faces of X, and $\deg^{(i)}(v)$ for the degree of a specific vertex. An infinite family of d-uniform simplicial complexes $\{X_n\}$ is called *bounded-degree* if there exists a (dimension-dependent) constant C such that $\deg^i(X) \leqslant C$ for all $i \leqslant d$.

Hitting Set. It will occasionally be useful for us to use a common variant of sampling on simplicial complexes called *hitting set*. We call a complex (γ, i)-hitting if for any $A \subset X(1)$:

$$\mathbb{P}_{\sigma \in X(i)}[\sigma \subset A] \leqslant \mathbb{P}[A]^i + \gamma$$

In other words, if the probability $\sigma \in X(i)$ *hits* $X(1) \setminus A$ is at least $1 - \mu(A)^i - \gamma$. If a complex is (γ, i)-hitting for all sizes i, we just call it γ-hitting.

4.4 Higher Order Random Walks

Simplicial complexes come equipped with several natural families of random walks generalizing the standard random walk on a graph. Toward this end, let X be a d-uniform simplicial complex and $\ell < k \leqslant d$ we define the standard 'averaging' or 'random walk' operators that move up and down the complex:

Down and up operators. The *down operator* $D_{k,\ell} : \ell_2(X(k)) \to \ell_2(X(\ell))$ is the bipartite operator of the containment graph of $(X(k), X(\ell))$, that is:

$$D_{k,\ell}f(s) = \mathbb{E}_{t \supseteq s}[f(t)].$$

The adjoint of $D_{k,\ell}$ is the *up operator*, $U_{\ell,k} : \ell_2(X(\ell)) \to \ell_2(X(k))$, given by:

$$U_{\ell,k}g(t) = \mathbb{E}_{s \subseteq t}[g(s)].$$

We also write $U_k := U_{k-1,k}$ and $D_k := D_{k,k-1}$ as shorthand throughout.

The composition of the up and down operators $U_{\ell,k}D_{k,\ell}$, called the "down-up" walk, is the normalized adjacency matrix of the graph whose vertices are $X(k)$ and whose edge distribution is defined by sampling $t \sim \mathbb{P}_\ell$, then $s, s' \sim \mathbb{P}_k$ conditioned on containing t. We denote the corresponding (bipartite) graph by $(X(k), X(\ell))$. For brevity, we sometimes just write $UD_{k,\ell}$ instead of $U_{\ell,k}D_{k,\ell}$.

Noise operator. The noise operator is one the most classical objects of study in Boolean function analysis. Given $\rho \in [0, 1]$, the standard noise operator $T_{k,\rho}$ operates on the hypercube $\mathbb{F}_2^k$ by re-sampling each bit uniformly with probability $1 - \rho$.

The noise operator has a natural extension to simplicial complexes [Baf+22b; GLL22]. It is convenient to define $T_{k,\rho} : \ell_2(X(k)) \to \ell_2(X(k))$ by its action on a face $\sigma \in X(k)$ by the following process:

1. Sample $\ell \sim Bin(k, \rho)$

2. Sub-sample $\tau \subset \sigma$ of size ℓ

3. sample $\sigma' \in X(k)$ conditioned on containing τ.

It is easy to check that when X is a product space,[5] this exactly recovers the standard noise operator. We note it is also possible to write the noise operator as a convex combination of down-up walks:

$$T_{k,\rho} = \sum_{\ell=0}^{k} \binom{k}{\ell} \rho^\ell (1 - \rho)^{k-\ell} UD_{k,\ell}.$$

---

[5]One can always express a product space $\otimes_i \Omega_i$ as a partite simplicial complex by defining each coordinate as a part, see [Baf+22b] for further details.

Swap walks. Let X be a d-uniform simplicial complex. Let i, j be so that $i + j \leqslant d$. The swap walk $S_{i,j} = S_{i,j}(X)$ is the bipartite adjacency operator of the graph $(X(i), X(j), E)$. An edge $\{s_i, s_j\}$ is chosen in this graph by first selecting a face $t \in X(i + j)$, and then partitioning it to $t = s_i \uplus s_j$ where $s_i \in X(i)$ and $s_j \in X(j)$ uniformly at random. This walk was defined and studied independently by [AJT19] and by [DD19].

Colored swap walks. The standard swap walks are not well behaved on partite complexes, but there is a useful analog for this setting defined in [DD19]. Let X be a d-uniform partite simplicial complex, and $F_1, F_2 \subseteq [d]$ be two disjoint subsets. The colored swap walk $S_{F_1, F_2} = S_{F_1, F_2}(X)$ is the bipartite adjacency operator of the graph $(X[F_1], X[F_2], E)$. An edge $\{s_1, s_2\}$ is chosen in this graph by first selecting a face $t \in X[F_1 \uplus F_2]$, and then partitioning it to $t = s_1 \uplus s_2$ according to its colors.

A useful observation is that the spectral expansion of swap walks (and color swap walks) is monotone in the following sense.

Observation 4.4. Let X be a d-uniform simplicial complex. Then for every i and $j' < j$, $\lambda(S_{i,j'}) \leqslant \lambda(S_{i,j})$. The same holds for partite complexes. For every disjoint F_1, F_2 and $F_2' \subseteq F_2$, $\lambda(S_{F_1, F_2'}) \leqslant \lambda(S_{F_1, F_2})$.

Proof. The proof follows from the fact that one can factor

$$S_{i,j'}(X) = D_{j,j'} S_{i,j}(X)$$

and the observation that $D_{j,j'}$ contracts 2-norms. A similar argument is true for colored swap walks in the partite case, replacing $D_{j,j'}$ with the corresponding bipartite graph between the F_2-colored faces and the F_2'-colored faces they contain. $\square$

4.5 Sampler Graphs

Sampler graphs are bipartite graphs $G = (L, R, E)$ where a random vertex $v \in L$ "sees" any large enough set in R with approximately the correct probability. Here we discuss a few classical variants of samplers and their relations, and refer the reader to [Gol97] for further discussion.

Definition 4.5 (Multiplicative sampler). Let $G = (L, R, E)$ be a bipartite graph and $\alpha, \beta, \delta > 0$. G is an (α, β, δ)-*multiplicative sampler* if for every set $A \subseteq R$ of size $\mathbb{P}[A] \geqslant \alpha$ it holds that

$$\mathbb{P}_{v \in L} \left[\left| \mathbb{P}_{u \sim v}[u \in A] - \mathbb{P}[A] \right| > \delta \, \mathbb{P}[A] \right] \leqslant \beta.$$

Note that the definition of a sampler is not a priori symmetric, L and R have different roles. We will also study a related *additive* notion of samplers.

Definition 4.6 (Additive sampler). Let $G = (L, R, E)$ be a bipartite graph and $\beta, \varepsilon > 0$. G is an (ε, β)-*additive sampler* if for every set $A \subseteq R$ it holds that

$$\mathbb{P}_{v \in L} \left[\left| \mathbb{P}_{u \sim v}[u \in A] - \mathbb{P}[A] \right| > \varepsilon \right] \leqslant \beta.$$

Finally, sometimes it is also useful to sample *functions* instead of sets. toward this we introduce the following definition in additive notation.

Definition 4.7. Let $G = (L, R, E)$ be a bipartite graph and $\varepsilon, \beta > 0$. G is a (ε, β)-*function additive sampler* if for every $f : R \to [0, 1]$ with expectation $\mathbb{E}_{v \in R}[f(v)] = \mu$, it holds that

$$\mathbb{P}_{v \in L} \left[\left| \mathbb{E}_{u \sim v}[f(v)] - \mu \right| > \varepsilon \right] < \beta,$$

and similarly for a (α, β, δ)-*function multiplicative sampler*.

18

4.5.1 Basic Sampler Properties

While the definition of samplers are not symmetric with respect to L and R, a near-tight correspondence (L, R, E) and (R, L, E) is given in [IKW12, Lemma 2.5]. We repeat their proof in Appendix C in a more general setup.

Claim 4.8. Let $\beta, \delta > 0$, let $\delta' > \delta$ and $\alpha < \frac{\min\{\delta, 0.5\}}{1+\delta}$. Then for every (α, β, δ)-sampler $G = (L, R, E)$, it holds that $G_{op} := (R, L, E)$ is a $(\frac{1-\alpha(1+\delta)}{\alpha(\delta'-\delta)}\beta, 2\alpha, \delta')$-sampler.

We also note that additive samplers and multiplicative are equivalent, at least in the weak sense as below. We prove this claim in Appendix C.

Claim 4.9. Let $G = (L, R, E)$ be a bipartite graph.

1. If G is a (β, δ)-additive sampler then G is a $(C\delta, \beta, \frac{1}{C})$-multiplicative sampler for any $C > 1$.

2. If G is a (α, β, δ)-multiplicative sampler for $\alpha \leqslant \frac{1}{2}$. Then G is a (β, δ)-additive sampler, where $\delta = \max\{\delta, (1+\delta)(\alpha+p)\}$ and $p = \max_{v \in R} \mathbb{P}[v]$.

Under a slightly stronger assumption one can also remove the dependence on δ in the second item.

Claim 4.10. Let $\beta, \alpha_0 > 0$. If for every $\alpha > \alpha_0$ it holds that G is an $(\alpha, \beta, \frac{\alpha_0}{\sqrt{\alpha}})$-multiplicative sampler, then G is a$(\beta, 2(\alpha+p))$-additive sampler where $p = \max_{v \in R} \mathbb{P}[v]$.

Below we show that every additive sampler is also a function additive sampler (albeit with worse parameters). We did not try to optimize parameters and it could be the case that a better reduction exists.

Claim 4.11. Let $G = (L, R, E)$ be an (ε, β)-additive sampler such that every $r \in R$ has degree at least k and every $v \in L$ has probability at most $\frac{1}{k}$. Assume that $\exp(-0.01\varepsilon^2 k)) < \frac{1}{4}$. Then G is also a $(4\beta, 2\varepsilon)$-function additive sampler.

We prove the claim in Appendix C. We note that [Gol97] also presents a reduction from an additive sampler to a function additive sampler is presented. The reduction there changes the underlying graph, which is why we prove a reduction more suitable to our needs.

4.6 Concentration of Measure

Sampler graphs are a special case of the powerful concept of concentration of measure, a viewpoint we will also take throughout this work. We follow the standard notation of [BLM13] adapted in the natural way from product measures to simplicial complexes. Given a simplicial complex X and a function $f : X(k) \to \mathbb{R}$, let $Z = f(x_1, \ldots, x_k)$ denote the random variable distributed as $f(x)$ where $x \in X(k)$. We are interested in understanding the concentration of Z around its mean. Z is said to satisfy *subexponential* concentration if there exists a constant $c > 0$ such that

$$\mathbb{P}[|Z - \mathbb{E}[Z]| > t] \leqslant \exp(-ct),$$

and *subgaussian* concentration if

$$\mathbb{P}[|Z - \mathbb{E}[Z]| > t] \leqslant \exp(-ct^2).$$

Classical concentration bounds (e.g. on the cube, products) typically hold for functions f satisfying certain Lipschitz-type conditions. Toward this end, define the Z-dependent variable $Z'_{(i)} = f(x_1, \ldots z_i, \ldots, x_k)$ where z_i is sampled conditional on x_{-i}.[6] Abusing notation slightly, we will call a function ν-Lipschitz if its squared difference with respect to the down-up walk is bounded:

Definition 4.12 (ν-Lipschitz function). Let X be a simplicial complex and $\nu > 0$. We call a function $f : X(d) \to \mathbb{R}$ ν-Lipschitz if with probability 1:

$$\sum_{i=1}^{d} (Z - Z'_{(i)})_+^2 \leqslant \nu,$$

where $(z)_+ = \max\{z, 0\}$.

---

[6]Recall that even in the non-partite setting, we assign each k-face an arbitrary order, making $Z'_{(i)}$ well defined.

We will occasionally rely on a stricter variant that requires the difference of f on any two neighboring k-faces be bounded.

Definition 4.13 (ν-bounded difference). We say a function $f : X(d) \to \mathbb{R}$ has ν-bounded difference if for every $s \in X(d)$ and neighboring s' of the down-up walk:

$$(f(s) - f(s'))^2 \leqslant \frac{\nu}{d}.$$

Most of our work focuses on a special case of functions with ν-bounded difference we call *lifted* functions.

Definition 4.14 (Lifted functions). Let X be a d-uniform simplicial complex and $f : X(k) \to \mathbb{R}$ any function. For any $k < k' \leqslant d$, the k'-lift of f is the function:

$$U_{k,k'} f(s) = \mathbb{E}_{s' \subset s}[f(s')]$$

We remark that lifted functions (sometimes called 'degree' or 'level'-k functions) are fundamental objects in boolean analysis and heavily studied in the HDX literature [Dik+18; Baf+22a; Gai+22]. As discussed in the introduction, inclusion sampling is in one-to-one correspondence with concentration bounds for lifted functions. More generally, bipartite sampling guarantees on $G = (L, R, E)$ correspond to concentration for functions of the form $A_G f$, where $f : R \to [0, 1]$ and $A_G : \mathbb{R}^R \to \mathbb{R}^L$ is the normalized bipartite adjacency matrix of G.

On simplicial complexes, we will typically be interested in understanding concentration at various levels of the complex. One frequently useful fact is that once concentration has been established for k-faces, it can typically be 'raised' or 'lowered' to functions on other levels via simple reductions to concentration of the complete complex. In cases, this can even be used to *improve* bounds on lower levels by bootstrapping from stronger bounds from larger set sizes. We record the following basic results to this effect here and use them throughout. Their proofs are given in [Appendix C].

Given integers $k \leqslant d$, a d-uniform simplicial complex X, and a function $f : X(k) \to \mathbb{R}$, denote

$$\pi_{up}^{d,k,f}(t) = \max_{s \in X(d)} \left\{ \Pr_{r \subset s} \left[f(r) - \mathbb{E}_{r \subset s}[f] > t \right] \right\}$$

and similarly

$$\pi_{low}^{d,k,f}(t) = \max_{s \in X(d)} \left\{ \Pr_{r \subset s} \left[f(r) - \mathbb{E}_{r \subset s}[f] < -t \right] \right\}$$

to be the worst-case concentration of f restricted to the induced complete complex on d-faces of X.

Lemma 4.15 (Raising Concentration). *Let X be a d-uniform simplicial complex, $k \leqslant d$, and $f : X(k) \to [0, 1]$ a function satisfying*

1. **Upper Tail**: $\Pr_{X(k)}[f - \mathbb{E}[f] > t] \leqslant up(t)$

2. **Lower Tail**: $\Pr_{X(k)}[f - \mathbb{E}[f] < -t] \leqslant low(t)$.

for some functions $up, low : \mathbb{R}_+ \to [0, 1]$. Then the d-lift $U_{k,d} f : X(d) \to \mathbb{R}$ satisfies:

1. **Upper Tail**: $\Pr_{X(d)}[U_{k,d}f - \mathbb{E}[f] > t] \leqslant up(\frac{t}{2})(1 - \pi_{low}^{d,k,f}(\frac{t}{2}))^{-1}$

2. **Lower Tail**: $\Pr_{X(d)}[U_{k,d}f - \mathbb{E}[f] < -t] \leqslant low(\frac{t}{2})(1 - \pi_{up}^{d,k,f}(\frac{t}{2}))^{-1}$.

We remark that since concentration in the complete complex is typically quite good, for most function classes of interest the latter term is close to 1 and very little is lost lifting the concentration to level d. We will also use a variant of this result for 'lowering' concentration bounds specified to Lipschitz functions.

Lemma 4.16 (Lowering Concentration). *Let X be a d-uniform simplicial complex and $k \leqslant d$. Assume there exist functions $up(t, \nu)$ and $low(t, \nu)$ such that any ν-Lipschitz $f : X(d) \to \mathbb{R}$ satisfies:*

1. **Upper Tail**: $\mathbb{P}[f - \mathbb{E}[f] > t] \leqslant up(t, \nu)$

2. **Lower Tail**: $\mathbb{P}[f - \mathbb{E}[f] < -t] \leqslant low(t, \nu)$.

Then any function $f' : X(k) \to \mathbb{R}$ with ν-bounded difference satisfies:

1. **Upper Tail**: $\mathbb{P}[f' - \mathbb{E}[f'] > t] \leqslant up(\frac{t}{2}, \frac{k}{d}\nu) + e^{-\frac{t^2}{4\nu}}$

2. **Lower Tail**: $\mathbb{P}[f' - \mathbb{E}[f'] < -t] \leqslant low(\frac{t}{2}, \frac{k}{d}\nu) + e^{-\frac{t^2}{4\nu}}$.

4.7 High Dimensional Expanders

In this section we define the notions of high dimensional expansion used throughout this work and dicsuss their relation with high order random walks.

Local Spectral Expanders. The most standard notion of spectral high dimensional expansion is 'local-spectral expansion', due to [DK17; Opp18].

Definition 4.17 (Local-spectral expander). Let X be a d-uniform simplicial complex. We say that X is a λ-one-sided (two sided) high dimensional expander if for every $i \leqslant d - 2$ and $s \in X(i)$, the graph underlying X_s is a λ-one-sided (two sided) spectral expander.

A key property of local-spectral expanders is that they imply the expansion of associated high order random walks [KM17; DK17; KO20; Dik+18; AL20]. We first state such a bound for the single and multi-step down-up walks (or more accurately their corresponding bipartite operator).

Theorem 4.18 ([AL20]). *Let X be a λ-one-sided d-uniform high dimensional expander. Then for every $\ell < k \leqslant d$ it holds that $\lambda(D_{k,\ell}) \leqslant \sqrt{\frac{\ell}{k}} \cdot (1 + \lambda)^{\frac{k-\ell}{2}}$.*

In the special case of the single step operator, we will occasionally use the following refined bound.

Theorem 4.19 ([AL20]). *Let X be a λ_i-one-sided d-uniform high dimensional expander. Then for every $\ell < k \leqslant d$ it holds that $\lambda(D_k) \leqslant \sqrt{\frac{1}{k} \prod_{i=0}^{k-2} (1 - \lambda_i)}$*

A critical observation of [AJT19; DD19] is that by removing the laziness inherent in the down-up walk, one can substantially improve expansion. We state an improvement of their bounds for these swap walks by Gur, Lifshitz, and Liu [GLL22]:[7]

Theorem 4.20 ([GLL22]). *Let X be a λ-two-sided high dimensional expander. Then*

$$\lambda(S_{\ell_1, \ell_2}) \leqslant \sqrt{\ell_1 \ell_2} \lambda$$

Recently, [AP23] proved tighter quantitative bounds that are non-trivial for any underlying local-spectral expansion. We will not need such fine-grain control in this work.

In the partite case, the colored swap walks also have known expansion bounds assuming partite local-spectral expansion [DD19; GLL22].

Theorem 4.21 ([DD19]). *Let X be a partite λ-one-sided high dimensional expander. Then*

$$\lambda(S_{F_1, F_2}) \leqslant |F_1| \, |F_2| \, \lambda.$$

We will also need a bound on the spectrum of the averaging operators based on local-spectral expansion. This has been the subject of intense study [KM17; DK17; Dik+18; KO20; AL20], with the best known bounds given by Alev and Lau [AL20].

---

[7]Formally this result is stated in [GLL22] for λ-products (defined below), but follows for λ-two-sided HDX since their partifications are λ-products.

λ-Products. Another useful notion of partite high dimensional expansion is to look directly at the expansion of the bipartite graphs between different components. We say that X is a λ-product (a term coined in [GLL22]), if for every face $s \in X$ and every two sides $X_s[i], X_s[j]$, the bipartite graph between the two sides is a λ-one-sided spectral expander. Dikstein and Dinur showed any partite local-spectral expander is a λ-product.

Claim 4.22 ([DD19]). Let X be a k-partite $\frac{\lambda}{1+\lambda}$-one-sided local-spectral expander. Then X is also a λ-product.

λ-Trickling-Down Complexes. Oppenheim's trickling-down theorem [Opp18] shows that in any connected complex, expansion in the top links 'trickles down' to expansion at lower levels.

Theorem 4.23 (The Trickling-Down Theorem [Opp18]). *Let X be a connected d-uniform simplicial complex. Assume that for every $s \in X(d-2)$, it holds that all non-trivial eigenvalues of X_s are in $[-\tau, \lambda]$ for some $\tau, \lambda \geqslant 0$. Then all non-trivial eigenvalues of X are in*

$$\left[-\frac{\tau}{1+(d-2)\tau}, \frac{\lambda}{1-(d-2)\lambda}\right].$$

As discussed in the introduction, the Trickling-Down Theorem introduces a 'phase transition' in the local-to-global behavior of HDX when co-dimension 2 links have expansion $\frac{1}{d-1}$. Past this point, expansion of the 1-skeleton of X can be inferred from expansion of the links. Before it, no guarantee is possible. This threshold-type behavior naturally suggests studying following definition of complexes 'approaching' the TD-barrier.

Definition 4.24 (λ-Trickling-Down Complex). We call a d-uniform simplicial complex λ-Trickling-Down (λ-TD) if it is connected, and all co-dimension 2 links have (one-sided) expansion $\frac{\lambda}{d-1}$.

Spectral Independence. Spectral independence is a closely related notion to local-spectral expansion introduced in the sampling literature [ALG21]. Here we give an equivalent version of the definition in terms of link expansion, and refer the reader to [ALG21] for the original definition and their equivalence.

Definition 4.25 (Spectral independence). For $\eta > 0$, a d-uniform complex X is called η-spectrally independent (η-SI) if the graph underlying every co-dimension j link is a (one-sided) $\frac{\eta}{j}$-expander.

4.8 Nice Complexes

In the following sections of the paper we prove results for various standard notions of high dimensional expansion including two sided high dimensional expanders, $\Omega(\frac{d}{\log d})$-skeletons of partite high dimensional expanders and $\sqrt{d}$-skeletons of one-sided high dimensional expanders. To compactify our statements, we bundle these notions into one definition henceforth dubbed 'nice complexes'.

Definition 4.26 (Nice complex). Let $c > 0$. A k-uniform complex X is *c-nice* if it meets one of the following conditions:

1. X is a 2^{-cd}-two-sided HDX.

2. X is the k-skeleton of a 2^{-cd}-one-sided d-partite HDX for $d \geqslant 3k \log k$.

3. X is the k-skeleton of a d-uniform complex Y with $\lambda(U_{d-1}D_d) = 1 - \frac{c}{d}$ and $d \geqslant 2k^2$.

We call X *c-locally nice* if every link of the complex is also c-nice.

Remark 4.27. We note that the two-sided condition of Definition 4.26 could be relaxed to instead assuming that for some $\alpha \in [0,1]$, X is the k-skeleton of a $\frac{1}{d}2^{-cd^\alpha}$-two-sided d-uniform HDX for $d \geqslant k^{\frac{2}{1+\alpha}}$. We prove a Chernoff bound in this regime which we believe is of independent interest, but for simplicity focus just on the $\alpha = 1$ case for the rest of the paper.

Note that skeletons of nice complexes are nice (see [Lee23, Theorem 3.5] for the third case), and that if X is c-nice for one of the first two items, it is also c-locally nice. It is simple computation to verify that skeletons of λ-TD and η-SI complexes are locally nice.

Claim 4.28. Let $\lambda < 1$ and $\eta > 0$. For any d-uniform complex Y:

1. If Y is λ-TD, then its $\sqrt{\frac{d}{2}}$-skeleton is $e^{-\frac{\lambda}{1-\lambda}}$-locally nice.

2. If Y is η-SI, then its $\sqrt{\frac{d}{\max\{2,\eta\}}}$-skeleton is e^{-2}-locally nice.

Proof. We first prove niceness. The first item follows from a simple computation combining Theorem 4.23 and Theorem 4.19 to show λ-TD complexes have $\lambda(U_{d-1}D_d) = 1 - \frac{e^{-\frac{\lambda}{1-\lambda}}}{d}$.

For the second item, let $k' = \frac{d}{\max\{2,\eta\}}$. By definition one can check the k'-skeleton of Y is a $\min\{\frac{3}{d}, \frac{\eta}{(1-1/\eta)d}\}$-one-sided local-spectral expander. Then Theorem 4.19 implies $\lambda(U_{k'-1}D_{k'}) \leqslant 1 - \frac{e^{-2}}{k'}$.

For local niceness, note that η-SI and λ-TD complexes have η-SI and λ-TD links respectively. Further, one can check directly that if X is a k-skeleton of d-uniform Y, then X_s is a $(k - |s|)$-skeleton of Y_s, which is $(d - |s|)$-uniform. The rate $\frac{d-|s|}{k-|s|} \geqslant \frac{d}{k}$ thus we can apply the same proof of niceness to every link as well. $\square$

5 Concentration on (Two Sided) HDX

We first give the formal statement of our optimal inclusion sampling theorem in the general case.

Theorem 5.1 (Sampling on HDX). *Fix $c > 0$. There are constants $c_1, c_2 > 0$ such that for any $k \in \mathbb{N}$, $i < k$, and k-uniform c-nice complex X, the containment graph $(X(k), X(i))$ is an (ε, β)-function sampler for*

$$\beta = \frac{c_1}{\varepsilon} \exp\left(-c_2 \varepsilon^2 \frac{k}{i}\right).$$

Moreover, if X is a k-skeleton of a d-uniform complex Y, then for every $k < k' \leqslant d$, the graph $(Y(k'), Y(i))$ is also an $(\varepsilon, O(\beta))$-function sampler.

In this section we will only prove the two-sided case of Theorem 5.1. We start by analyzing the more typical 'Chernoff-Hoeffding'-type setting, then show how to bootstrap this bound to prove the two-sided case of Theorem 5.1.

5.1 Chernoff-Hoeffding

In this section we introduce a new technique to prove 'k vs. 1' concentration on simplicial complexes X whose underlying strength depends on the expansion of the *swap walks* of X. The general form is fairly cumbersome (we refer the reader to Proposition 5.4, which bounds the moment generating function in terms of swap walk expansion), and for readability only state the result here for $\frac{1}{k}2^{-k^\alpha}$-two-sided HDX, interpolating between $\frac{1}{2k}$-HDX at $\alpha = 0$, and $2^{-\Omega(k)}$-HDX at $\alpha = 1$.

Theorem 5.2. *Let $\alpha \in [0, 1]$ and X be a k-uniform $\frac{1}{k}2^{-k^\alpha}$-two-sided local-spectral expander. Then for any $f : X(1) \to [0, 1]$ and $\varepsilon > 0$:*

1. *Upper tail:*

$$\mathbb{P}_{\{v_1,\dots,v_k\}\in X(k)}\left[\frac{1}{k}\sum_{i=1}^{k} f(v_i) - \mu > \varepsilon\right] \leqslant 2\min\{k, c_\alpha\}\exp\left(-\frac{1}{12}\varepsilon^2 k^{\frac{1+\alpha}{2}}\right)$$

2. *Lower tail:*

$$\mathbb{P}_{\{v_1,\dots,v_k\}\in X(k)}\left[\frac{1}{k}\sum_{i=0}^{k} f(v_i) - \mu < -\varepsilon\right] \leqslant 2\min\{k, c_\alpha\}\exp\left(-\frac{1}{12}\varepsilon^2 k^{\frac{1+\alpha}{2}}\right)$$

where $c_\alpha \leqslant e^{\frac{(\frac{3}{\alpha})^{\frac{1}{\alpha}}}{e^\alpha}}$.

We remark it is possible to use this approach to prove similar concentration for any λ-TD complex, but in Section 7.2 we will see a stronger bound for this specific regime based on a variant of the Herbst argument. As such we allow ourselves the additional constant factor here in expansion for convenience. Second, note that by setting α appropriately, Theorem 5.2 shows for any $c > 0$ that the containment graph $(X(k), X(1))$ of any 2^{-ck}-two-sided local spectral expanders is a (ε, β)-sampler for $\beta = \exp(-\Omega_c(-\varepsilon^2 k))$.

The core of Theorem 5.2 centers around the notion of *splittability* [AJT19], which will allow us to de-correlate the variables in our complex to bound the moment generating function of $U_{1,k}f$. We introduce a slight variant of the standard definition convenient for our purposes.

Definition 5.3 (Balanced Splittability). We call a d-uniform complex X $\{\gamma_i\}_{i \in [\lfloor \log(d) \rfloor - 1]}$-balanced-splittable if for every i the 2^i-th swap walk expands:
$$\lambda(S_{2^i, 2^i}) \leqslant \gamma_i$$

All applications of splittability in the literature lose *additive* factors in the γ_i. We introduce a new approach combining splittability with an 'internal' reduction to the complete complex that allows finer control over this error and instead incurs more manageable *multiplicative* loss.

Proposition 5.4. *For any $j \in \mathbb{N}$ and $r \in [0, 1]$, let $k = 2^j$, and X be a k-uniform $\{\gamma_i\}$-balanced-splittable complex with $e^{\frac{5}{4}2^{-i}r^2 k}\gamma_{j-1-i} < 1$. Then for any $f : X(1) \to [0, 1]$:*

$$\mathbb{E}_{\{v_1, \ldots, v_k\} \in X(k)} \left[\prod_{i=0}^{k} e^{rf(v_i)} \right] \leqslant c_{r,\gamma} \mathbb{E}_{X(1)} [e^{rf}]^k$$

where $c_{r,\gamma} \leqslant \left(\prod_{i=0}^{j-1} \left(\frac{1 - \gamma_{j-1-i}}{1 - e^{\frac{5}{4}2^{-i}r^2 k}\gamma_{j-1-i}} \right)^{2^i} \right).$

While the definition of $c_{r,\gamma}$ may be somewhat hard to parse, note at least that for any fixed r, $c_{r,\gamma}$ goes to 1 as γ tends to 0.

Proof. For any even ℓ, denote by $z_\ell : X(\ell) \to \mathbb{R}$ the 'partial' moment generating function:

$$z_\ell(\{v_1, \ldots, v_\ell\}) = \prod_{i=1}^{\ell} e^{rf(v_i)}.$$

Our goal is to bound $\mathbb{E}[z_k] = \mathbb{E}[\exp(r \sum_{i=1}^{k} f(v_i))]$ by induction on z_ℓ. Denote by $z_\ell^0 = \mathbb{E}[z_\ell]\mathbf{1}$ the projection of z_ℓ to the space of constant functions and by $z_\ell^\perp = z_\ell - z_\ell^0$ its orthogonal part. The first key observation is that we may split z_ℓ into two 'copies' of $z_{\ell/2}$ using the swap operator:

$$\mathbb{E}_{X(\ell)} [z_\ell] = \mathbb{E}_{X(\ell/2)} [z_{\ell/2} S_{\ell/2} z_{\ell/2}].$$

Now splitting $z_{\ell/2}$ itself into parallel and perpendicular components, we can bound this as

$$\begin{aligned}
\mathbb{E}_{X(\ell/2)} [z_{\ell/2} S_{\ell/2} z_{\ell/2}] &= \mathbb{E}[z_{\ell/2} S_{\ell/2} (z_{\ell/2}^0 + z_{\ell/2}^\perp)] \\
&= \mathbb{E}[z_{\ell/2}]^2 + \langle z_{\ell/2}^\perp, S_{\ell/2} z_{\ell/2}^\perp \rangle \\
&\leqslant \mathbb{E}[z_{\ell/2}]^2 + \gamma_{\log(\ell)-1} \|z_{\ell/2}^\perp\|_2^2 \\
&= (1 - \gamma_{\log(\ell)-1}) \mathbb{E}[z_{\ell/2}]^2 + \gamma_{\log(\ell)-1} \|z_{\ell/2}\|_2^2.
\end{aligned} \tag{5.1}$$

The first of these two terms can clearly be bounded by induction. The difficulty lies in the error term, as the 2-norm of the partial MGF may be much larger than its expectation. To handle this, we argue via reduction to the complete complex the second moment of $z_{\ell/2}$ is actually quite close to the original expectation of z_ℓ.

Claim 5.5. For any even ℓ and $r \in [0, 1]$:

$$\|z_{\ell/2}\|_2^2 \leqslant e^{\frac{5}{4}r^2 \ell} \mathbb{E}[z_\ell]$$

24

Let's first complete the proof given this fact. Plugging the claim back into (5.1) and applying induction we get the following bound:

$$\mathbb{E}[z_k] \leqslant \frac{1 - \gamma_{j-1}}{1 - e^{\frac{5}{4}r^2\ell}\gamma_{j-1}} \mathbb{E}[z_{k/2}]^2$$

$$\leqslant \frac{1 - \gamma_{j-1}}{1 - e^{\frac{5}{4}r^2k}\gamma_{j-1}} \left(\prod_{i=0}^{j-2} \left(\frac{1 - \gamma_{j-2-i}}{1 - e^{2^{-i}\frac{5}{4}r^2\frac{k}{2}}\gamma_{j-2-i}} \right)^{2^{i+1}} \right) \mathbb{E}[z_1]^k$$

$$= \left(\prod_{i=0}^{j-1} \left(\frac{1 - \gamma_{j-1-i}}{1 - e^{2^{-i}\frac{5}{4}r^2(k+1)}\gamma_{j-1-i}} \right)^{2^i} \right) \mathbb{E}[e^{rf}]^k$$

as desired. $\qquad\square$

It is left to prove the claim relating $\|z_{\ell/2}\|_2^2$ to $\mathbb{E}[z_\ell]$.

Proof of Claim 5.5. Recall our goal is to show

$$\mathbb{E}[z_{\ell/2}^2] = \mathop{\mathbb{E}}_{t \in X(\ell/2)} \left[\prod_{v \in t} e^{2rf(v)} \right] \leqslant e^{\frac{5}{4}r^2\ell}\mathbb{E}[z_\ell]$$

The key is to observe that we can draw $t \in X(\ell/2)$ by first drawing $s \in X(\ell)$, and then drawing $t \subset s$ uniformly at random:

$$\mathop{\mathbb{E}}_{t \in X(\ell/2)} \left[\prod_{v \in t} e^{2rf(v)} \right] = \mathop{\mathbb{E}}_{s \in X(\ell)} \left[\mathop{\mathbb{E}}_{t \subset s} \left[\prod_{v \in t} e^{2rf(v)} \right] \right]$$

The inner expectation is now within the complete complex which is negatively correlated, so may therefore 'pull out' the product (e.g. by [Hoe94, Theorem 4]) and write:

$$\mathop{\mathbb{E}}_{t \subset s} \left[\prod_{v \in t} e^{2rf(v)} \right] \leqslant \mathop{\mathbb{E}}_{v \in s}[e^{2rf(v)}]^{\ell/2}.$$

Let $\mu_s = \mathbb{E}_{v \in s}[f(v)]$ and observe $\ell\mu_s = \sum_{v \in s} f(v)$. Combining the above, we have

$$\mathbb{E}[z_{\ell/2}^2] \leqslant \mathop{\mathbb{E}}_{s \in X(\ell)} \left[\mathop{\mathbb{E}}_{v \in s}[e^{2rf(v)}]^{\ell/2} \right]$$

$$\leqslant \mathop{\mathbb{E}}_{s \in X(\ell)} \left[\mathop{\mathbb{E}}_{v \in s}[1 + 2rf(v) + \frac{5}{2}r^2f(v)]^{\ell/2} \right]$$

$$= \mathop{\mathbb{E}}_{s \in X(\ell)} \left[(1 + 2r\mu_s + \frac{5}{2}r^2\mu_s)^{\ell/2} \right]$$

$$\leqslant \mathop{\mathbb{E}}_{s \in X(\ell)} \left[e^{(r+\frac{5}{4}r^2)\ell\mu_s} \right]$$

$$= \mathop{\mathbb{E}}_{s \in X(\ell)} \left[\prod_{v \in s} e^{(r+\frac{5}{4}r^2)f(v)} \right]$$

$$\leqslant e^{\frac{5}{4}r^2\ell}\mathbb{E}[z_\ell]$$

where in the second inequality we've used that $e^x < 1 + x + \frac{5}{4}x^2$ for $x \in [0,2]$ and that $f^2(v) \leqslant f(v)$. $\quad\square$

The proof of Theorem 5.2 is now essentially immediate from Chernoff's method.

Proof of Theorem 5.2. We prove the upper tail. The lower tail follows from applying the upper tail to $1 - f$. Let k' be the largest power of 2 such that $k' \leqslant k$. Fix $r = \frac{1}{2}\varepsilon k'^{\frac{\alpha-1}{2}}$, and recall by Theorem 4.20 that $\gamma_i \leqslant \frac{2^i}{k'} 2^{-(k')^\alpha}$. By Proposition 5.4, we therefore have:

$$\mathop{\mathbb{P}}_{\{v_1,\ldots,v_{k'}\}\in X(k')}\left[\sum_{i=1}^{k'} f(v_i) \geqslant \mu k' + \varepsilon k'\right] = \mathop{\mathbb{P}}_{\{v_1,\ldots,v_{k'}\}\in X(k')}\left[\exp\left(r\sum_{i=1}^{k} f(v_i)\right) \geqslant \exp\left(rk'\mu + rk'\varepsilon\right)\right]$$

$$\overset{Markov}{\leqslant} \frac{\mathbb{E}\left[\prod_{i=0}^{k'} e^{rf(v_i)}\right]}{\exp(rk'\mu + rk'\varepsilon)}$$

$$\leqslant \left(\prod_{i=0}^{j-1}\left(\frac{1}{1 - 2^{-i-1}e^{-\frac{1}{3}(k')^\alpha}}\right)^{2^i}\right)\exp\left(-\frac{\varepsilon^2}{3}(k')^{\frac{1+\alpha}{2}}\right)$$

where in the final inequality we have applied Proposition 5.4 and bounded $\mathbb{E}[e^{rf}]^{k'}$ by standard manipulations. To bound the leading coefficient first note that for any $\alpha \geqslant 0$ we have

$$\prod_{i=0}^{j-1}\left(\frac{1}{1 - 2^{-i-1}e^{-\frac{1}{3}(k')^\alpha}}\right)^{2^i} = \prod_{i=0}^{j-1}\left(\frac{1}{1 - 2^{-i-1}}\right)^{2^i} \leqslant k'$$

since each individual term in the product is at most $\frac{1}{2}$ and $j = \log_2(k')$. For $\alpha > 0$ we may also write

$$\prod_{i=0}^{j-1}\left(\frac{1}{1 - 2^{-i-1}e^{-\frac{1}{3}(k')^\alpha}}\right)^{2^i} \leqslant \prod_{i=0}^{j-1}\left(\frac{1}{1 - \frac{e^{-\frac{1}{3}(k')^\alpha}}{2}}\right)^{2^i}$$

$$\leqslant \left(1 + e^{-\frac{1}{3}(k'+1)^\alpha}\right)^{k'}$$

$$\leqslant e^{e^{\frac{k'}{\frac{1}{3}(k')^\alpha}}}$$

$$\leqslant e^{\frac{(\frac{3}{\alpha})^{\frac{1}{\alpha}}}{e^\alpha}}$$

where the last inequality follows from observing $e^{e^{\frac{k'}{\frac{1}{3}(k')^\alpha}}}$ is maximized at $k' = \left(\frac{3}{\alpha}\right)^{-\alpha}$.

Finally we apply Lemma 4.15 to lift this guarantee to $X(k)$:

$$\mathop{\mathbb{P}}_{X(k)}[U_{1,k}f - \mu > \varepsilon] \leqslant \min\{k', c_\alpha\}\exp\left(-\frac{\varepsilon^2}{12}(k'+1)^{\frac{1+\alpha}{2}}\right)\left(1 - \pi_{low}^{k,k',U_0^{k'}f}(\varepsilon/2)\right)^{-1}$$

$$\leqslant \min\{k, c_\alpha\}\exp\left(-\frac{\varepsilon^2}{12}(k+1)^{\frac{1+\alpha}{2}}\right)\left(1 - e^{-\frac{\varepsilon^2(k+1)}{6}}\right)^{-1}$$

$$\leqslant 2\min\{k, c_\alpha\}\exp\left(-\frac{\varepsilon^2}{12}k^{\frac{1+\alpha}{2}}\right)$$

where in the final inequality we have assumed $\varepsilon^2 \geqslant \frac{6}{k}$ (else the stated bound is trivial). $\qquad\square$

5.2   From Chernoff to Inclusion Sampling

We now show how to bootstrap Chernoff-type concentration to optimal sampling for the general inclusion graph $(X(k), X(i))$. We focus on the case of two-sided HDX, and give the substantially more involved argument for the full theorem in Section 7. We restate the theorem in the two-sided case here for convenience.

Theorem 5.6 (Sampling on two-sided HDX)**.** *For any $c > 0$ there exists $c' > 0$ such that for any $k \in \mathbb{N}$, $i < k$, and k-uniform 2^{-ck}-two-sided local spectral expander X, $(X(k), X(i))$ is an (ε, β) sampler for*

$$\beta = \frac{c'}{\varepsilon} \exp\left(-\Omega_c\left(\varepsilon^2 \frac{k}{i}\right)\right).$$

Moreover, if X is a k-skeleton of a d-uniform complex Y, then for every $k < k' \leqslant d$, the graph $(Y(k'), Y(i))$ is also an $(\varepsilon, O(\beta))$-function sampler.

The proof of Theorem 5.6 is based on a refinement of an idea of [IKW12], who observed on the complete complex one can prove an analogous bound by partitioning k-sets into i disjoint $\frac{k}{i}$-sets, and treating these as independent variables. In the world of general simplicial complexes, this type of subdivision is called the *faces complex* of X [DD23b].

Definition 5.7 (Faces Complex [DD23b])**.** Let X be a d-uniform simplicial complex. For any $i \leqslant d$, the faces complex $F_X^{(i)}$ is the $\lfloor \frac{d}{i} \rfloor$-uniform complex with vertices $X(i)$ and top level faces:

$$F_X^{(i)}\left(\left\lfloor \frac{d}{i} \right\rfloor\right) = \left\{ \{s_1, \ldots, s_{\lfloor \frac{d}{i} \rfloor}\} \;\middle|\; \bigcup_{j=1}^{\lfloor \frac{d}{i} \rfloor} s_j \in X, \bigcup_{j \neq \ell} s_j \cap s_\ell = \emptyset \right\}$$

The weight of a face is proportional to the weight of its union in X. We drop X from the notation when clear from context.

In our setting, subdivided faces have extreme dependencies so [IKW12]'s analysis for the complete complex fails. Nevertheless, it is in fact possible to generalize this approach to any simplicial complex through somewhat more careful analysis.

Proposition 5.8. *Let X be a d-uniform complex and $i \leqslant k \leqslant d$ be such that $\left(F^{(i)}\left(\lfloor \frac{k}{i} \rfloor\right), F^{(i)}(1)\right)$ is an (ε, β)-function additive sampler. Then $(X(k), X(i))$ is a $(2\varepsilon, \frac{2}{\varepsilon}\beta)$-function additive sampler.*

The proof of Proposition 5.8 itself centers around a lemma for general bipartite samplers stating that sampling over correlated sub-divisions is sufficient to imply sampling for the general graph as long as the components are appropriately marginally distributed.

Lemma 5.9. *Given a graph $G = (L, R, E)$, parameters $\varepsilon, \beta > 0$ and $k \in \mathbb{N}$, and $f : L \to [0, 1]$, suppose there is a distribution $(v_1, v_2, \ldots, v_k, u) \sim D$ over $L^k \times R$ satisfying*

1. *For every $i \in [k]$, the marginal (v_i, u) is distributed according to the edges of G.*

2. $\mathbb{P}_{(v_1, v_2, \ldots, v_k, u) \sim D}\left[\frac{1}{k}\sum_{i=1}^{k} f(v_i) > \mathbb{E}[f] + \varepsilon\right] \leqslant \beta.$

Then

$$\mathbb{P}_{u \sim R}\left[\mathbb{E}_{v \in L, v \sim u}[f(v)] > \mathbb{E}[f] + 2\varepsilon\right] \leqslant \frac{\beta}{\varepsilon}. \tag{5.2}$$

The same statement applies replacing $>$ with $<$ and ε with $-\varepsilon$ inside the probabilities.

We defer the proof to the end of the subsection and first prove Proposition 5.8.

Proof of Proposition 5.8. Let $m = \lfloor \frac{k}{i} \rfloor$ and fix a function $f : X(i) \to [0, 1]$ with expectation $\mu = \mathbb{E}_{s \in X(i)}[f(v)]$. Consider the distribution D which samples $(s_1, s_2, \ldots, s_m, t) \subset X(i)^m \times X(k)$ by drawing $\{s_i\}_{i=1}^{m}$ from the faces complex, and draws $t \supset \bigcup s_i$ from $X(k)$. We claim D satisfies the requirements of Lemma 5.9, namely:

1. $\forall i \in [m]$, (s_i, t) is marginally distributed according to the inclusion graph edges $(X(k), X(i))$

2. $\mathbb{P}_{\{s_1, s_2, \ldots, s_m, t\} \sim D}\left[\left|\frac{1}{k}\sum_{i=1}^{k} f(s_i) - \mu\right| > \varepsilon\right] \leqslant \beta.$

27

The first fact follows from the definition of the faces complex, namely that marginally each s_i is distributed as $X(i)$, and t is drawn from $X(k)$ conditional on containing s_i. The second fact is immediate from our assumption that the faces complex is an (ε, β)-sampler.

$\square$

We are now ready to prove the two-sided case of Theorem 5.1.

Proof of Theorem 5.6 (Theorem 5.1 two sided case). By Theorem 5.2 if the faces complex of X is a 2^{-ck}-two-sided HDX, then it is an (ε, β)-sampler for $\beta \leqslant c_1 e^{\Omega_c\left(\varepsilon^2 \lfloor \frac{k}{i} \rfloor\right)}$ and $c_1 > 0$ some constant dependent only on c. Then by Proposition 5.8, X would be a $(2\varepsilon, \frac{\beta}{\varepsilon})$-sampler. The proof is completed by observing $2^{-\Omega(k)}$-local-spectral expansion of the faces complex is immediate from Theorem 4.20, since the graph underlying any link of the faces complex is a swap walk within (a link of) X. The 'moreover' statement follows directly from Lemma 4.15.

$\square$

It remains to prove Lemma 5.9.

Proof of Lemma 5.9. Let $Z : R \to [0, 1]$ record the probability $u_0 \in R$ ε-mis-samples f under D:

$$Z(u_0) := \mathop{\mathbb{P}}_{(v_1, v_2, \ldots, v_k, u) \sim D} \left[\frac{1}{k} \sum_{i=1}^{k} f(v_i) > \mu + \varepsilon \ \middle| \ u = u_0 \right].$$

By assumption, the expected mis-sampling probability is small

$$\mathop{\mathbb{E}}_{u_0} [Z(u_0)] = \mathop{\mathbb{P}}_{(v_1, v_2, \ldots, v_k, u) \sim D} \left[\frac{1}{k} \sum_{i=1}^{k} f(v_i) > \mu + \varepsilon \right] \leqslant \beta,$$

so Markov's inequality implies the following tail bound on $Z(u_0)$:

$$\mathop{\mathbb{P}}_{u_0 \in R} [Z(u_0) > \varepsilon] \leqslant \frac{\beta}{\varepsilon}.$$

We argue if u_0 is such that $\mathbb{E}_{v \in L, v \sim u_0} [f(v_i)] > \mu + 2\varepsilon$ then $Z(u_0) > \varepsilon$, thus concluding that (5.2) holds.

The key is the simple observation that the (conditional) expectation A under D and G are equivalent:

$$\mathop{\mathbb{E}}_{(v_1, v_2, \ldots, v_k, u) \sim D} [f(v_i) \mid u = u_0] = \frac{1}{k} \sum_{i=1}^{k} \mathop{\mathbb{E}}_{(v_i, u)} [f(v_i) | u = u_0]$$
$$= \mathop{\mathbb{E}}_{v \in L} [f(v) \mid v \sim u_0] \tag{5.3}$$

where the first equality is by linearity of expectation, and the second is by our assumption that (v_i, u) is distributed as G. The result now follows roughly from arguing the random variable $\frac{1}{k} \sum_{i=1}^{k} f(v_i)$ restricted to u_0, is non-trivially concentrated around its expectation $\mathbb{E}_{v \in L} [f(v) \mid v \sim u_0]$, which is at least 2ε-far from μ.

In particular, fix u_0 such that $\mathbb{E}_{(v_1, v_2, \ldots, v_k, u) \sim D} [f(v_i) \mid u = u_0] = \mu_{u_0} > \mu + 2\varepsilon$. Then, we bound $Z(u_0)$ by its upper tail:

$$Z(u_0) \geqslant \mathop{\mathbb{P}}_{(v_1, v_2, \ldots, v_k, u) \sim D} \left[\frac{1}{k} \sum_{i=1}^{k} f(v_i) > \mu_{u_0} - \varepsilon \ \middle| \ u = u_0 \right].$$

Denote the right-hand side by p, and as $f(v) \in [0, 1]$ we have

$$\mu_{u_0} \leqslant p + (1 - p)(\mu_{u_0} - \varepsilon).$$

Re-arranging gives

$$Z(u_0) \geqslant p \geqslant \frac{\varepsilon}{1 - \mu_{u_0} - \varepsilon} \geqslant \varepsilon.$$

The last inequality follows from the the denominator being in $(0, 1)$ ($0 < \mu_{u_0} - \varepsilon < 1$ because $\mu_{u_0} - \varepsilon \geqslant \mu + \varepsilon > 0$ and $\mu_{u_0} - \varepsilon < \mu_{u_0} \leqslant 1$).

Finally, we note the usual trick (changing f in the proof to $1 - f$) gives the bounds for

$$\mathbb{P}_{u \in R}\left[\mathbb{E}_{v \sim u, v \in L}[f(v)] < \mu - 2\varepsilon\right].$$

$\square$

5.2.1 Multiplicative Sampling

In the following sections, it will sometimes be more convenient to work with *multiplicative* sampling instead of the additive bounds of Theorem 5.1, and with the 'flipped' inclusion graph $(X(i), X(k))$. By the generic translations in Claim 4.9 and Claim 4.8 we get the following corollary for this setting.

Corollary 5.10. *Let X be a k-maximal c-nice complex. Then for every $\delta, \beta \in (0, 1)$, $(X(i), X(k))$ is a (α, β, δ)-sampler for*

$$\alpha \leqslant \frac{c'}{\beta\delta}e^{-\Omega_c(\beta^2\delta^2\frac{k}{i})}.$$

We remark that we have not attempted to optimize dependence on β and δ above. In the following sections we will only be interested in the setting where both are constant.

6 Reverse Hypercontractivity

In this section we recursively leverage optimal inclusion sampling to prove reverse hypercontractivity for high dimensional expanders.

Theorem 6.1 (Reverse Hypercontractivity)**.** *Fix $c > 0$, $\rho \in (0, 1)$, and let X be a k-uniform c-locally nice complex for k sufficiently large. Then there exist constants C, q (dependent only on c, ρ) such that*

$$\langle f, T_\rho g \rangle \geqslant C\|f\|_q\|g\|_q.$$

At its core, the proof of Theorem 6.1 really only relies on X and its links satisfying optimal i vs k sampling. With this in mind, we will instead work with the following notion of a *Link up-sampler* that captures this core property:

Definition 6.2 (Link up-sampler (LUS))**.** *Let $\tau \in (0, 1)$ and let X be a k-uniform simplicial complex. We say that X is a τ-LUS if for every $i \leqslant k - 2$, every $j \leqslant k - i$, and every $s \in X(i)$, it holds that $(X_s(j), X_s(k - i))$ is a $(\frac{1}{\tau}\exp(-\tau\frac{k-i}{j}), 0.1, 0.2)$-multiplicative sampler.*

We note the constants $0.1, 0.2$ are essentially arbitrary and fixed for convenience. The notion can also be further relaxed in a number of ways such as asking for sampling for j only up to a linear number of levels or only within links up to a certain level and the main results of this section would still hold.

We prove link-up samplers are reverse hypercontractive. Theorem 6.1 is an immediate corollary.

Theorem 6.3. *For every $\rho \in (0, 1)$ and $\tau > 0$, there is some $\ell > 1$ and $C > 0$ such that the following holds for all sufficiently large $k \in \mathbb{N}$. Let X be k-uniform and τ-LUS. Then T_ρ is $(\frac{1}{\ell}, \frac{\ell}{1-\ell}, C)$-reverse hypercontractive.*

We split the proof of Theorem 6.3 below into two parts, roughly corresponding to reverse hypercontractivity for boolean functions, and a reduction from the general to boolean case.

6.1 The Boolean Case

The main workhorse behind Theorem 6.3 is to show that any linear-step *down-up walk* on an LUS is reverse hypercontractive for sets of size up to $\exp(-\Omega(k))$.

Theorem 6.4. *Let $\tau > 0$ and $\gamma \in (0,1)$. There exist constants $c, q > 0$ such that the following holds for any sufficiently large k. Let X be a k-uniform τ-LUS. Then for any sets $A, B \subseteq X(k)$ of relative size at least $\exp(-ck)$ and $\gamma' \leqslant \gamma$ it holds that*

$$\mathbb{P}_{t,t' \sim UD_{k,\lfloor \gamma' k \rfloor}} \left[t \in A, t' \in B \right] \geqslant \mathbb{P}\left[A\right]^q \mathbb{P}\left[B\right]^q .$$

It will be convenient for us to assume at least one of the sets A and B is not too large. To this end, we first handle the case that both are large separately.

Claim 6.5. For every $p > 0.5$ there exists $q \geqslant 1$ so that the following holds. Let X be a simplicial complex and $A, B \subseteq X(k)$ subsets such that $\mathbb{P}\left[A\right], \mathbb{P}\left[B\right] \geqslant p$. Then for any $\gamma' \leqslant \gamma$:

$$\mathbb{P}_{t,t' \sim UD_{k,\lfloor \gamma' k \rfloor}} \left[t \in A, t' \in B \right] \geqslant \mathbb{P}\left[A\right]^q \mathbb{P}\left[B\right]^q .$$

The proof is largely un-enlightening computation, so we postpone it to the end of the subsection. Henceforth, assume without loss of generality that $\mathbb{P}\left[A\right] \leqslant \mathbb{P}\left[B\right]$ and has $\mathbb{P}\left[A\right] \leqslant 0.8$.

With this out of the way, we overview the main proof strategy. Assume for the moment that $\gamma' k$ is integral. As discussed in the introduction, Theorem 6.4 relies on the following elementary observation: since the down-up walk samples an edge by first sampling $s \in X(\gamma k)$ and then *independently* sampling $t, t' \in X(k)$ containing s, we can express the γ'-correlated mass of A and B as:

$$\mathbb{P}_{t,t' \sim UD_{k,\gamma' k}} \left[t \in A, t' \in B \right] = \mathbb{E}_{s \in X(\gamma' k)} \left[\mathbb{P}_t \left[A \mid t \supseteq s\right] \cdot \mathbb{P}_t \left[B \mid t \supseteq s\right] \right] .$$

This motivates us to consider the set of $s \in X(\gamma k)$ that *simultaneously* see a large enough fraction of A, B. Toward this end, for any integer $\ell_0 \leqslant k$ and real number $\delta > 0$, let

$$G^X(\ell_0, \delta) = \left\{ s \in X(\ell_0) \,\middle|\, \mathbb{P}_t \left[A \mid t \supseteq s\right] \geqslant \delta \mathbb{P}\left[A\right] \text{ and } \mathbb{P}_t \left[B \mid t \supseteq s\right] \geqslant \delta \mathbb{P}\left[B\right] \right\}$$

denote the set of elements in $X(\ell_0)$ that see a noticeable fraction of both A and B.

Recalling $\mathbb{P}\left[B\right] \geqslant \mathbb{P}\left[A\right]$, to prove Theorem 6.4 it suffices to show for any $\gamma' \leqslant \gamma$

$$\mathbb{P}\left[G^X(\gamma' k, \mathbb{P}\left[A\right]^{O(1)}) \right] \geqslant \mathbb{P}\left[A\right]^{O(1)} , \tag{6.1}$$

since (writing $G := G^X(\gamma' k, \mathbb{P}\left[A\right]^{O(1)})$) we then have

$$\mathbb{P}_{t,t' \sim UD_{k,\gamma' k}} \left[t \in A, t' \in B \right] \geqslant \mathbb{P}_{s \in X(\gamma' k)}\left[G\right] \cdot \mathbb{P}_{s \in X(\gamma' k)} \left[\mathbb{P}_t[t \in A | s \in G] \, \mathbb{P}_t[t \in B | s \in G] \right]$$

$$\geqslant \mathbb{P}[A]^{O(1)} \mathbb{P}[B] \tag{6.2}$$

As such, our main goal is to prove (6.1). A key observation toward this end is that G satisfies certain useful 'composition' behavior: we can bound $G(\gamma' k, \cdot)$ by iteratively bounding $G(i, \cdot)$ for smaller i within repeatedly deeper links of the complex. In particular, for sets A and B and any $r \in X$, define the restrictions of A and B to the link of r as

$$A_r := \{t \in X_r : r \cup t \in A\} \qquad \text{and} \qquad B_r := \{t \in X_r : r \cup t \in B\}$$

and define the *local* subset $G^{X_r}(\ell_0, \delta)$ as the set of $s \in X_r(\ell_0)$ that see a sufficiently fraction of A_r and B_r:

$$G^{X_r}(\ell_0, \delta) = \left\{ s \in X_r(\ell_0) \,\middle|\, \mathbb{P}_{t \in X_r(k-r)} \left[A_r \mid t \supseteq s\right] \geqslant \delta \mathbb{P}\left[A_r\right] \text{ and } \mathbb{P}_{t \in X_r(k-r)} \left[B_r \mid t \supseteq s\right] \geqslant \delta \mathbb{P}\left[B_r\right] \right\} .$$

G and G^{X_r} compose in the following fashion:

Observation 6.6 (Composition of G^X). Let $A, B \subset X(k)$, $\ell_1, \ell_2 \in \mathbb{N}$, and let $\delta_1, \delta_2, \eta_1, \eta_2 > 0$. If

30

1. $\mathbb{P}\left[G^X(\ell_1, \delta_1)\right] \geqslant \eta_1$.

2. $\forall r \in G^X(\ell_1, \delta_1)$: $\mathbb{P}_{X_r(\ell_2)}\left[G^{X_r}(\ell_2, \delta_2)\right] \geqslant \eta_2$

Then
$$\mathbb{P}\left[G^X(\ell_1 + \ell_2, \delta_1\delta_2)\right] \geqslant \eta_1\eta_2.$$

Proof. Sample $s \in X(\ell_1 + \ell_2)$ by sampling $r_1 \in X(\ell_1)$ and then $r_2 \in X_r(\ell_0)$ setting $s = r_1 \cup r_2$. Fix $r_1 \in G^X(\ell_1, \delta_1)$. Observe that if $r_2 \in G^{X_{r_1}}(\ell_2, \delta_2)$ then

$$\mathbb{P}_t\left[A \mid r_1 \cup r_2 \subseteq t\right] \geqslant \delta_2 \mathbb{P}_t\left[A \mid r_1 \subseteq t\right]$$
$$\geqslant \delta_1\delta_2 \mathbb{P}\left[A\right],$$

and similarly for B. Therefore in this case $s \in G^X(\ell_1 + \ell_2, \delta_1\delta_2)$. Hence

$$\mathbb{P}_s\left[G^X(\ell_1 + \ell_2, \delta_1\delta_2)\right] \geqslant \mathbb{P}_{r_1}\left[G^X(\ell_1, \delta_1)\right] \cdot \mathbb{P}_{s=r_1\cup r_2}\left[r_2 \in G^{X_{r_1}}(\ell_2, \delta_2) \;\middle|\; r_1 \in G^X(\ell_1, \delta_1)\right]$$
$$\geqslant \eta_1\eta_2.$$

$\square$

The proof strategy (modulo details), is then simple. To bound $G(\gamma'k, \mathbb{P}[A]^{q_0})$, we divide $t \in X(k)$ into pieces as $r_1 \cup r_2 \cup r_3 \ldots$, where each r_i is viewed as being drawn from the link of $r_1 \cup \ldots \cup r_{i-1}$. The local G-set for each r_i can be lower bounded by the sampling properties of $X_{\cup_{j<i}r_j}$, and by carefully choosing the size of the r_i, we can ensure that composition gives the desired bound on G.

In practice, this approach hits some complications due to the fact that the localized restrictions of A and B shrink throughout the process. To handle this, we'll instead use a 'two-phase' version of the above. First, we show that there exists some small absolute constant γ_0, potentially much smaller than the desired γ, for which the above approach does work. We call complexes satisfying such a bound lovely:

Definition 6.7 (Lovely complex). We say that X is $(c, 1 - \gamma_0, k, q)$-*lovely*, if for all $A, B \subset X(k)$ with $\mathbb{E}[B] \geqslant \mathbb{E}[A] \geqslant \exp(-ck)$
$$\mathbb{P}\left[G^X(\gamma_0 k, \mathbb{P}[A]^q)\right] \geqslant \mathbb{P}[A]^q.$$

We first argue any LUS complex is γ'-lovely for some γ' depending only on τ and γ.

Lemma 6.8. *For every $\tau > 0$ and $\gamma < 1$ there exist c, q and $\gamma_0 > 0$ so that following holds for large enough k. Let X be τ-LUS, then for every $t \in X^{\leqslant \gamma k}$, X_t is $(c, 1 - \gamma', k - |t|, q)$-lovely for any $\gamma' \leqslant \gamma_0$.*

Second, we give a similar argument showing that loveliness itself can be bootstrapped to higher levels.

Claim 6.9. Let X be a k-uniform simplicial complex. Let $q', q'' \geqslant 1$, and $c', c'', \gamma_1, \gamma_2 \geqslant 0$. Assume that:

1. X is $(c', 1 - \gamma_1, k, q')$-lovely.

2. For every $r \in X(\gamma_1 k)$, X_r is $(c'', 1 - \gamma_2, (1 - \gamma_1)k, q'')$-lovely.

Then X is $(c, (1 - \gamma_1)(1 - \gamma_2), k, q)$-lovely for $q = q''q' + q'' + q'$ and $c = \min\{\frac{c''}{(q'+1)(1-\gamma_1)}, c'\}$.

Combined with some definition chasing, Lemma 6.8 and Claim 6.9 immediately imply Theorem 6.4.

Proof of Theorem 6.4. Recall that by (6.2) it suffices to show X is $(c, 1 - \gamma, k, q)$-lovely for some constants c, q depending only on γ and τ. We first prove a slightly different statement by induction:

$$\forall \ell \in \left[\left\lceil \frac{\log(1 - \gamma)}{\log(1 - \gamma_0)} \right\rceil\right]: \quad X \text{ is } (c_\ell, (1 - \gamma_0)^\ell, k, q_\ell)\text{-lovely.}$$

for constants c_ℓ, q_ℓ dependent only on τ, γ. Note this implicitly assumes $(1 - \gamma_0)^\ell$ is integer. This is unecessary and we discuss removing this assumption at the end of the subsection. Finally, observe that for $\ell = \left\lceil \frac{\log(1 - \gamma)}{\log(1 - \gamma_0)} \right\rceil$

31

we have $(1 - \gamma_0)^\ell \leqslant 1 - \gamma$. We will argue a slight modification gives exactly $1 - \gamma''$ for any $\gamma'' \leqslant \gamma$ after showing the above.

The base case $\ell = 1$ is immediate from Lemma 6.8. Assume the statement is true for some $\ell < \left\lceil \frac{\log(1-\gamma)}{\log(1-\gamma_0)} \right\rceil$, then X is $(c(\ell), (1 - \gamma_0)^\ell, k, q(\ell))$-lovely. Moreover, for $j = 1 - (1 - \gamma_0)^\ell \leqslant \gamma k$, Lemma 6.8 implies every j-link X_t is $(c'', 1 - \gamma', k - j, q'')$-lovely for any $\gamma' \leqslant \gamma_0$. Thus by Claim 6.9, we have that X is $(c_{\ell+1}, 1 - (1 - \gamma_0)^{\ell+1}, k - |t|, q_{\ell+1})$-lovely where $c_{\ell+1}$ and $q_{\ell+1}$ are functions of γ_0, $c(\ell)$, $q(\ell)$, c'', and q'' as described in Claim 6.9. Since γ_0, c'', and q'' are themselves functions only of τ and γ, this gives the desired bound.

In order to prove X is $(c, 1 - \gamma, k, q)$-lovely, simply observe that in the final step of this induction we may choose $\gamma_0' \leqslant \gamma_0$ such that $(1 - \gamma_0)^\ell(1 - \gamma_0') = 1 - \gamma$. Exactly the same argument as above then gives X is at least $(c_\ell, 1 - \gamma, k, q_\ell)$-lovely for $\ell = \left\lceil \frac{\log(1-\gamma)}{\log(1-\gamma_0)} \right\rceil$. The same strategy can be used to achieve the same (or better) constants for any $\gamma' < \gamma$ as well. $\qquad\square$

It remains to prove Lemma 6.8 and Claim 6.9. While Lemma 6.8 is the central component, it is instructive to first prove the simpler Claim 6.9 which is in some sense a 'one-step' variant of the former.

Proof of Claim 6.9. Let A, B be sets of size at least $\exp(-ck)$. Our goal can be rephrased as showing

$$G^X\big((\gamma_1 + \gamma_2(1 - \gamma_1))k, \mathbb{P}\,[A]^{q''q' + q'' + q'}\big) \geqslant \mathbb{P}\,[A]^{q''q' + q'' + q'}$$

since $1 - (\gamma_1 + \gamma_2(1 - \gamma_1)) = (1 - \gamma_1)(1 - \gamma_2)$. We prove this via Observation 6.6. In particular since $c' < c$, $(c', 1 - \gamma_1, k, q')$-loveliness of X implies

$$G^X(\gamma_1 k, \mathbb{P}\,[A]^{q'}) \geqslant \mathbb{P}\,[A]^{q'}.$$

Moreover, for every $r \in G^X(\gamma_1 k, \mathbb{P}\,[A]^{q'})$ we have that $\mathbb{P}_{X_r}[A], \mathbb{P}_{X_r}[B] \geqslant \mathbb{P}\,[A]^{q'+1} \geqslant \exp(-c''(1 - \gamma_1)k)$. Therefore by $(c'', 1 - \gamma_2, (1 - \gamma_1)k, q'')$-loveliness of X_r it holds that

$$G^{X_{t_1}}\big(\gamma_2(1 - \gamma_1)k, \underset{X_r}{\mathbb{P}}\,[A]^{q''}\big) \geqslant \underset{X_r}{\mathbb{P}}\,[A]^{q''} \geqslant \mathbb{P}\,[A]^{(q'+1)q''}.$$

Combining the two bounds and applying Observation 6.6 we then have

$$G^X\big(\gamma_1 k + \gamma_2(1 - \gamma_1)k, \mathbb{P}\,[A]^{q''q' + q'' + q'}\big) \geqslant \mathbb{P}\,[A]^{q''q' + q'' + q'}$$

as desired. $\qquad\square$

Finally, we end with the proof of Lemma 6.8. Here, instead of decomposing $s = r \cup r'$, we will need to more carefully decompose $s = r_1 \cup r_2 \cup r_3 \ldots$ and control in every step how much A and B shrink using the link-up-sampling property.

Proof of Lemma 6.8. We start with a few simplifying assumptions. First, recall that by Claim 6.5 we may assume that $\mathbb{P}\,[A] \leqslant 0.8$. Second, note it is sufficient to prove the theorem just for X. The result then follows for all γk-links X_t since the link of a τ-LUS complex is itself τ-LUS, and the ambient dimension of each link is at least $(1 - \gamma)k$ which we take to be still sufficiently large. Finally we introduce one notational simplification. Recall in the definition of a τ-LUS that for any i-face s the graph $(X_s(j), X_s(k - i))$ is a $\big(\frac{1}{\tau}\exp(-\tau\frac{k-i}{j}), 0.1, 0.2\big)$-multiplicative sampler. In the proof of this lemma we will only consider links of faces s with size $i \leqslant \gamma_0 k$ for some sufficiently small γ_0, so these are always $\big(\frac{1}{\tau}\exp(-\tau(1 - \gamma_0)\frac{k}{j}), 0.1, 0.2\big)$-multiplicative samplers. Moreover, we will also always ensure $j \leqslant \gamma_0 k$, allowing us to subsume the constant $\frac{1}{\tau}$ into the exponent. In particular, as long as $\gamma_0 \leqslant \frac{\tau}{3\ln\frac{1}{\tau}}$, we may assume all (examined) links in the proof are $(\exp(-\tau'\frac{k}{j}), 0.1, 0.2)$-multiplicative samplers where $\tau' = \frac{\tau}{3}$.

Fix $\gamma_0 = \min\{\frac{\tau}{3\ln\frac{1}{\tau}}, \gamma\}$, let $c = \gamma_0 \frac{\tau'}{3}$, and assume $\mathbb{P}\,[B] \geqslant \mathbb{P}\,[A] \geqslant \exp(-ck)$. Fix $m = \frac{-\log\mathbb{P}[A]}{c}$, and define the 'step size' $\ell = \frac{k}{m}$. Observe that $1 \leqslant \ell \leqslant \gamma_0 k$. Assume for the moment that ℓ is also integer (see Remark 6.11). We record the following immediate observations:

Observation 6.10.

1. $\mathbb{P}[A] = \exp(-cm)$.

2. $0.8^m = \mathbb{P}[A]^{q_0}$ where $q_0 = -c^{-1}\log 0.8$ is a constant.

3. For every $i \leqslant \gamma_0 m$:
$$0.8^i \, \mathbb{P}[A] \geqslant \exp(-\tau' m) = \exp(-\tau'\frac{k}{\ell}).$$

The first two items are by definition. The third is satisfied for any γ_0, c such that $c(1 + \gamma_0 q_0) \leqslant \tau'$. One can check that taking $\gamma_0 \leqslant \min\{\gamma, \frac{\tau}{3\ln\frac{1}{\tau}}\}$ suffices. Finally, assume for simplicity k is such that $\gamma_0 k \in \mathbb{N}$.

We now move to the main argument. Let $j \in \{1, 2, \ldots, j_{fin} = \lceil \gamma_0 m \rceil\}$ and for every $j \leqslant \lfloor \gamma_0 m \rfloor$ define
$$G_j := G^X(j\ell, (0.8)^j)$$
as above. If $j_{fin} > \gamma_0 m$, additionally define
$$G_{j_{fin}} := G^X(\gamma_0 k, 0.8^{j_{fin}}) \subseteq G^X(\gamma_0 k, \mathbb{P}[A]^{\gamma_0 q_0 + 1}).$$

where in the last step we've used the assumption that $\mathbb{P}[A] \leqslant 0.8$. We will show
$$\mathbb{P}\left[G_{j_{fin}}\right] \geqslant (0.8)^{j_{fin}} \geqslant \mathbb{P}[A]^{\gamma_0 q_0 + 1}. \tag{6.3}$$

Setting $q = \gamma_0 q_0 + 1$ proves the lemma. We prove (6.3) by inductively showing that for every $j = 1, 2, \ldots, j_{fin}$,
$$\mathbb{P}[G_j] \geqslant (0.8)^j.$$

We start with the base case $j = 1$, which already contains the main idea. Given $S \subset X(k)$, define
$$T(S) := \left\{r \in X(\ell) : \mathbb{P}[S \mid r] < 0.8\,\mathbb{P}[S]\right\}$$
to be the set of 'terrible' ℓ-sets which see too little of S. Observe that by construction
$$G_1 = X(\ell) \setminus (T(A) \cup T(B)).$$

Since $\ell \leqslant \gamma_0 k$, we have $(X(k), X(\ell))$ is a $(\exp(-\tau'\frac{k}{\ell}), 0.1, 0.2)$-sampler by our earlier discussion. On the other hand, by Observation 6.10, we can bound the sizes of A and B as:
$$\mathbb{P}[B] \geqslant \mathbb{P}[A] = \exp(-cm) \geqslant \exp(-\tau'\frac{k}{\ell}).$$

Thus applying sampling we can bound $\mathbb{P}[G_1]$ by
$$\mathbb{P}[G_1] \geqslant 1 - \mathbb{P}[T(A)] - \mathbb{P}[T(B)] \geqslant 1 - 2 \cdot 0.1 = 0.8$$
as desired.

The inductive step is similar to the base case, only performed *locally* within the links of X together with Observation 6.6. By induction hypothesis $\mathbb{P}[G_i] \geqslant 0.8^i$. For every $s \in G_i$, by Observation 6.10
$$\mathbb{P}_{X_s}[B], \mathbb{P}_{X_s}[A] \geqslant (0.8)^i \, \mathbb{P}[A] \geqslant \exp(-\tau' m) \geqslant \exp(-\tau'\frac{k}{\ell}).$$

Thus by the same argument above for X_s and $A_s = \{t \setminus s \mid s \subseteq t \in A\}, B_s = \{t \setminus s \mid s \subseteq t \in B\}$, we have
$$\mathbb{P}\left[G^{X_s}(\ell, 0.8) \,\middle|\, s \in G_i\right] \geqslant 1 - \mathbb{P}[T(A)] - \mathbb{P}[T(B)] \geqslant 1 - 2 \cdot 0.1 = 0.8.$$

The inequality $\mathbb{P}[G_{i+1}] \geqslant 0.8^{i+1}$ then follows by Observation 6.6.

Finally, we note that for the last step, going between $j_{fin} - 1$ to j_{fin} (in the case that $j_{fin} > \gamma m$), we follow the same procedure except that we may need to sample less than ℓ new vertices. This only improves the sampling so the same analysis as above holds. $\qquad\square$

Remark 6.11. Throughout the proof we have 'cheated' in the typical way assuming a priori that various parameters (ℓ, γk, $\gamma_0 k$, and $(1 - \gamma_0)^i (1 - \gamma')k$) are integer. Note that we did not assume above $\gamma_0 m$ was integer because of the careful interplay between m and $\mathbb{P}[A]$. In all other cases, however, these assumptions are benign and taking the appropriate integer floor does not substantially change the proof. The only modification is that every 'step' in the process (e.g., drawing an $\lfloor \ell \rfloor$-size set) potentially moves one fewer level up the complex. In the worst case, this results in taking a constant multiplicative factor more steps than under integer assumptions, resulting in additional constant factors in c and q_0. We omit the details.

Proof of Claim 6.5. Denote by $E(X, Y)$ the set of directed edges (t, t') such that $t \in X, t' \in Y$. It is easy to see that $\mathbb{P}[E(A, B)] + \mathbb{P}[E(A, B^c)] = \mathbb{P}[A]$. Similarly we have that $\mathbb{P}[E(A, B^c)] + \mathbb{P}[E(A^c, B^c)] = \mathbb{P}[B^c]$ so $\mathbb{P}[B^c] \geqslant \mathbb{P}[E(A^c, B^c)]$. combining these together gives that $\mathbb{P}[E(A, B)] \geqslant \mathbb{P}[A] - \mathbb{P}[B^c]$. Assume that $\mathbb{P}[A] \leqslant \mathbb{P}[B]$, this implies that $\mathbb{P}[E(A, B)] \geqslant \mathbb{P}[A] - \mathbb{P}[A^c] = 2\mathbb{P}[A] - 1$. Hence it is sufficient to show that given A such that $\mathbb{P}[A] > p$, there exists a constant $q > 0$ such that $2\mathbb{P}[A] - 1 \geqslant \mathbb{P}[A]^q$ (which in turn is at least $\mathbb{P}[A]^q \mathbb{P}[B]^q$). Indeed let us consider $f(x) = 2x - 1 - x^q$ and show this function is non-negative in the range $[p, 1]$ for large enough q.

Assume without loss of generality that q is an integer. In this case $f(x) = (1 - x)(\sum_{j=1}^{q-1} x^j - 1)$. This function is greater than 0 in $[p, 1]$ if and only if $\sum_{j=1}^{q-1} x^j \geqslant 1$ in this range. In the range $[p, 0.9]$ the series $\sum_{j=1}^{q-1} x^j$ converges uniformly to $\frac{x}{1-x}$ which is always greater or equal $\frac{p}{1-p} > 1$. In the range $x \in [0.9, 1]$ we can take any $q \geqslant 3$ and one can verify that $\sum_{j=1}^{q-1} x^j > 0.9 + 0.9^2 > 1$. □

6.2 The General Case

We now prove Theorem 6.3 using our restricted reverse hypercontractive inequality for indicators. This is done via the following reduction from reverse hypercontractivity for sets to the general case.

Theorem 6.12. *For every $\ell \geqslant 1$ and $\varepsilon > 0$ there exists $\kappa' \geqslant \left(\frac{1}{5}\right)^{2\ell(1+\varepsilon)} \left(\frac{1}{18 + \frac{12}{\varepsilon}}\right)^{2(\ell-1)}$ such that the following holds. Let V be a finite probability space and D a monotone linear operator such that for every $A, B \subseteq V$,*

$$\langle 1_A, D 1_B \rangle \geqslant \kappa\, \mathbb{P}[A]^\ell\, \mathbb{P}[B]^\ell. \tag{6.4}$$

Then for every two arbitrary *functions $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ it holds that*

$$\langle f_1, D f_2 \rangle \geqslant \kappa\kappa' \|f_1\|_{\frac{1}{\ell(1+\varepsilon)}} \|f_2\|_{\frac{1}{\ell(1+\varepsilon)}}. \tag{6.5}$$

That is, D is $\left(\frac{1}{\ell(1+\varepsilon)}, \frac{1}{1 - \ell(1+\varepsilon)}, \kappa\kappa'\right)$-reverse hypercontractive.

The proof involves careful discretization and thresholding of f_1 and f_2 into level sets in a way that largely maintains the functions' moments. We defer the details to Appendix E and prove the main Theorem assuming this result.

Proof of Theorem 6.3. By Theorem 6.12 it is enough that we show that there exist $\ell > 1, \kappa > 0$ such that for every $A, B \subseteq X(k)$

$$\langle 1_A, T_\rho 1_B \rangle \geqslant \kappa\, \mathbb{P}[A]^\ell\, \mathbb{P}[B]^\ell.$$

Since $\gamma > \rho$, observe we can always take k sufficiently large so that $\mathbb{P}_{Y \sim Bin(\rho, k)}[Y \leqslant \lfloor \gamma k \rfloor] \geqslant \frac{1}{2}$. Let q, c be the constants promised by Theorem 6.4 with respect to the $UD_{\lfloor \gamma k \rfloor, k}$. By monotonicity of q and c, these apply to all walks of intersection at most $\lfloor \gamma k \rfloor$). Thus if both $\mathbb{P}[A], \mathbb{P}[B] \geqslant \exp(-ck)$ then by Theorem 6.4:

$$\mathop{\mathbb{P}}_{t_1, t_2 \sim T_\rho}[t_1 \in A, t_2 \in B] \geqslant \mathop{\mathbb{P}}_{Y \sim Bin(\rho, k)}[Y \leqslant \gamma k]\, \mathbb{P}[A]^q\, \mathbb{P}[B]^q \geqslant \frac{1}{2}\, \mathbb{P}[A]^q\, \mathbb{P}[B]^q.$$

Otherwise, it holds that $\mathbb{P}[A]\, \mathbb{P}[B] \leqslant \exp(-ck)$. On the other hand, the probability of resampling *all vertices* in the noise operation is $(1 - \rho)^k = \exp(-ck)^{\ell''}$ for some constant $\ell'' > 1$ that depends on c and on ρ. In particular,

$$\mathop{\mathbb{P}}_{t_1, t_2 \sim T_\rho}[t_1 \in A, t_2 \in B] \geqslant (1 - \rho)^k\, \mathbb{P}[A]\, \mathbb{P}[B] \geqslant \mathbb{P}[A]^{\ell''+1}\, \mathbb{P}[B]^{\ell''+1}.$$

Taking $\ell = \max\{q, \ell'' + 1\}$ completes the proof. □

6.3 Proof of Theorem 6.1

Finally, we briefly give the formal proof for reverse hypercontractivity on high dimensional expanders. This follows from the fact that any c-locally nice complex is $\tau(c)$-LUS for some appropriate setting of constants.

Corollary 6.13. *For every $c < 1$ there exists a constant $\tau > 0$ such that any c-locally nice complex is τ-LUS.*

Proof. By definition every i-link X_s is a c-nice complex, so by Corollary 5.10 every $(X_s(j), X_s(k-i))$ is a $(\frac{1}{c_1}\exp(-c_2\frac{k-i}{j}), 0.1, 0.2)$-sampler for some constants c_1, c_2 dependent only on c. Taking $\tau = \min\{c_1, c_2\}$ then suffices. $\square$

Proof of Theorem 6.1. The proof is immediate from Corollary 6.13 and Theorem 6.3. $\square$

7 Concentration for All Nice HDX

In this section we study concentration of measure for partite complexes and complexes whose down-up walk have an $\Omega(\frac{1}{d})$-spectral gap, completing the proof of Theorem 5.1. Along the way, we'll show that complexes near the TD barrier satisfy exponential concentration for Lipschitz functions.

7.1 Chernoff-Hoeffding for Partite HDX

We'll start with the setting of partite HDX. As in the two-sided case, we first study the special case of $(X(k), X(1))$. In the partite case different components may behave differently, making the splittability argument for two-sided HDX fail. Instead, we will rely on a *localization* approach, leveraging the local sampling properties between bipartite components of the complex. The resulting quantitative bounds via this technique are largely incomparable with our prior two-sided analysis. While they are weaker for large λ, they approach the error of true independent sampling for small enough λ and remove the restrictive condition of two-sided expansion. The latter in particular actually leads to sparser complexes with subgaussian concentration due to better known degree bounds for partite constructions.

Before stating the result formally, we briefly introduce some relevant notation. Given a complex X, let X^{ind} denote the complete complex whose vertices are $X(1)$, that is:

$$X^{ind}(k) = \{\{v_1, v_2, \ldots, v_k\} : v_i \in X(1)\}$$

where each v_i is drawn independently from $X(1)$ *with repetition*.[8] Similarly for X partite, let $X^{ind,p}$ be the complete k-partite complex whose sides are $X[1], X[2], \ldots, X[k]$:

$$X^{ind,p}(k) = \{\{v_1, v_2, \ldots, v_k\} : v_1 \in X[1], v_2 \in X[2], \ldots, v_k \in X[k]\}$$

where each v_i is drawn independently from the marginal over $X[i]$.

Finally, let $\sigma(X, k, \varepsilon)$ denote the worst case 'k vs 1' ε-sampling error of X.

$$\sigma(X, k, \varepsilon) = \max\left\{\Pr_{s \in X(k)}\left[\left|\Pr_{v \in X(1)}[A \mid v \in s] - \Pr_{v \in X(1)}[A]\right| > \varepsilon\right] \;\middle|\; A \subseteq X(1)\right\}$$

We prove that sampling in two-sided and one-sided partite HDX match the bounds in their respective independent complex up to $\text{poly}(k)\lambda$-error.

Theorem 7.1. *Let $\eta > 0, \varepsilon \in (0, \frac{1}{2})$, and $\lambda \leqslant \frac{\eta^{2.5}}{16k^4}$. Then*

1. *If X is k-uniform λ-two-sided local-spectral expander:*

$$\sigma(X, k, \varepsilon) \leqslant \sigma(X^{ind}, k, \varepsilon) + \eta$$

---

[8] We remark we have abused notation somewhat as this is not formally a simplicial complex and the faces are *multi*-sets, but this is irrelevant for our purposes.

2. If X is a k-partite λ-one-sided local-spectral expander:

$$\sigma(X, k, \varepsilon) \leqslant \sigma(X^{ind,p}, k, \varepsilon) + \eta$$

In particular, in both cases X is then a (ε, β)-sampler for $\beta = \eta + \exp(-\Omega(\varepsilon^2 k))$ by Chernoff-Hoeffding.

In the two-sided case, note that Theorem 5.2 is tighter when $\lambda > \exp(-\Omega(k))$, while for small enough λ Theorem 5.2 wins out achieving parameters arbitrarily close to independent sampling. We note that here we only give a bound on sampling sets, but this can be converted to concentration for lifts of bounded functions with essentially no loss by Claim 4.11.

The core of Theorem 7.1 is a lower bound on the k-wise correlation across coordinates of a λ-product.

Proposition 7.2. *Let X be a k-uniform λ-product. For every $\eta < 1$ and functions $f_i : X[i] \to [0,1]$:*

$$\mathop{\mathbb{E}}_{\{v_1, v_2, \ldots, v_k\} \in X(k)} \left[\prod_{i=1}^{k} f_i(v_i) \right] \geqslant (1-\eta)^k \prod_{i=1}^{k} \mathbb{E}[f_i]$$

whenever $\lambda \leqslant \frac{\eta^{1.5} \min_i \{\mathbb{E}[f_i]^2\}}{\sqrt{8k^3}}$.

Proof. The proof relies on the following adaptation of an elementary connection between bipartite expansion and sampling observed in [DK17]:

Claim 7.3. Let $G = (L, R, E)$ be a λ-bipartite expander and let A be its bipartite adjacency operator. Let $f : L \to [0,1]$ be a function with $\mathbb{E}[f] \geqslant \mu$ and let $0 < \varepsilon < \mu$. Then

$$\mathop{\mathbb{P}}_{u \in R}[Af(u) < \mu - \varepsilon] < \frac{\lambda^2 \mu}{\varepsilon^2}.$$

We prove this in Appendix C for completeness. Using this fact, we prove the following more general claim. Write $\mu_i = \mathbb{E}[f_i]$ and fix $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k > 0$ such that $\mu_i - (i-1)\varepsilon_i > 0$. For $\varepsilon = \min_i\{\varepsilon_i\}$ we show

$$\mathop{\mathbb{E}}_{\{v_1, v_2, \ldots, v_k\} \in X(k)} \left[\prod_{i=1}^{k} f_i(v_i) \right] \geqslant \prod_{i=1}^{k} \left(\mu_i - (i-1)\varepsilon_i - \lambda^2 \frac{k-i}{\varepsilon^2 \cdot \min_{j>i}(\mu_j - (i-1)\varepsilon_j)} \right). \tag{7.1}$$

Setting $\varepsilon_i = \frac{\mu_i \eta}{2i}$ and applying our assumed bound on λ then gives the desired result.

We prove (7.1) by induction on k, the uniformity of the complex. For $k = 1$ the bound is trivial since $\mathbb{E}[f_1] = \mu_1$. Assume the statement holds for any $(k-1)$-uniform λ-product, and for each $i = 2, \ldots, k$ let T_i denote the vertices in $X[1]$ that under-sample f_i:

$$T_i = \left\{ v_1 \in X[1] \ \middle| \ \mathop{\mathbb{E}}_{v_i \in X_{v_1}[i]}[f(v_i)] < \mu_i - \varepsilon_i \right\}.$$

Since each $(X[1], X[i])$ is a λ-bipartite expander, Claim 7.3 implies

$$\mathbb{P}\left[\bigcup_{i=2}^{k} T_i \right] \leqslant \sum_{i=2}^{k} \frac{\lambda^2}{\mu_i \varepsilon_i^2} \leqslant \lambda^2 \frac{k-1}{\varepsilon^2 \cdot \min_{j>1}\{\mu_j\}}.$$

We de-correlate our expected product along the first coordinate by conditioning on $G = X[1] \setminus \{\bigcup_{i=2}^{k} T_i\}$:

$$\mathbb{E}\left[\prod_{i=1}^{k} f_i(v_i) \right] \geqslant \mathbb{E}\left[\mathbf{1}_G(v_0) \prod_{i=1}^{k} f_i(v_i) \right]$$

$$\geqslant \mathbb{E}[\mathbf{1}_G(v_0) f_0(v_0)] \cdot \min_{v_1 \in G} \left(\mathop{\mathbb{E}}_{\{v_2, v_3, \ldots, v_k\} \in X_{v_1}} \left[\prod_{i=2}^{k} f_i(v_i) \right] \right).$$

36

We bound the two terms separately. For the first:

$$\mathbb{E}[\mathbf{1}_G(v_0)f_0(v_0)] \geqslant \mathbb{E}[f_0(v_0)] - \mathbb{P}\left[\bigcup_{i=2}^{k} T_i\right] \tag{7.2}$$

$$\geqslant \mu_0 - \lambda^2 \frac{k-1}{\varepsilon^2 \cdot \min_{j>1}\{\mu_j\}}. \tag{7.3}$$

To bound the minimum term, we apply the inductive hypothesis on every link X_{v_1}. Toward this end, for notational simplicity re-index the sides of X_{v_1} to be from 1 to $k-1$ and set:

1. $f_i' = f_{i+1}$,

2. $\varepsilon_i' = \varepsilon_{i+1}$,

3. $\mu_i' = \mu_{i+1} - \varepsilon_{i+1}$ and $\mu' = \min_{j>1} \mu_i'$.

Since $\mathbb{E}_{X_{v_1}[i]}[f_i'] \geqslant \mu_i'$ by assumption on the links, applying the inductive hypothesis to X_{v_1} and $\{f'\}$ bounds the minimum by

$$\prod_{i=1}^{k-1}\left(\mu_i' - (i-1)\varepsilon_i' - \lambda^2 \frac{k-i}{\varepsilon^2 \cdot \min_{j>i}(\mu_j' - (i-1)\varepsilon_j')}\right)$$

Combining this with the prior term (corresponding to $i = 0$) gives the desired bound

$$\prod_{i=1}^{k}\left(\mu_i - (i-1)\varepsilon_i - \lambda^2 \frac{k-i}{\varepsilon^2 \cdot \min_{j>i}(\mu_j - (i-1)\varepsilon_j)}\right).$$

$\square$

We note a more involved variant of the above applying different bounds in each level of induction implies the following finer-grained bound:

Claim 7.4. Let X be a k-partite λ-product. Let $f_i : X[i] \to [0,1]$ be such that $\mathbb{E}_{v_i \in X[i]}[f(v_i)] \geqslant \mu_i$. Let $\{\varepsilon_{i,\ell} > 0\}_{1 \leqslant i < \ell \leqslant k}$ and let $s_{i,\ell}^{(m)} = \sum_{j=\ell}^{m} \varepsilon_{j,i}$ (and for brevity $s_{m+1,\ell}^{(m)} = 0$). Then

$$\mathbb{E}_{\{v_0, v_1, \dots, v_k\} \in X(k)}\left[\prod_{i=1}^{k} f_i(v_i)\right] \geqslant \prod_{i=1}^{k}\left(\mu_i - s_{i,i}^{(k-1)} - \lambda^2 \sum_{j=0}^{i-1} \frac{\mu_j - s_{i,j}^{(k-1)}}{\varepsilon_{i,j}^2}\right). \tag{7.4}$$

We omit the proof which is an unenlightening technical extension of the above.

Using our lower bound on negative correlation, we show that for any function tuple $f = (f_1, f_2, \dots, f_k)$, the random variables $f(X)$ and $f(X^{ind})$ are actually *distributionally* close. This will allow us to immediately recover any concentration bound on X^{ind} up to the distributional error. More formally, given a k-partite simplicial complex X, let D_{ind} denote the distribution supported by of $X^{ind,p}$, i.e. $\mathbb{P}_{D_{ind}}[\{v_1, v_2, \dots, v_k\}] = \prod_{i=1}^{k} \mathbb{P}_{X[i]}[v_i]$. Moreover, for any distribution D over $\Omega = X[1] \times X[2] \times \cdots \times X[k]$, and a function $\bar{f} : \Omega \to \mathbb{R}^k$ define the *push forward* distribution $\bar{f}_* D : \mathbb{R}^k \to [0,1]$ as

$$\bar{f}_* D(\bar{x}) = \mathbb{P}_{\omega \sim D}[\bar{f}(\omega) = \bar{x}].$$

Let $D_X = \pi_k$ denote the distribution over k-sets of X. We show the pushforward distributions $\bar{f}_* D_X$ and $\bar{f}_* D_{ind}$ are close in TV-distance.

Lemma 7.5. *Let $\varepsilon, \eta > 0$ and let X be a k-partite λ-product for $\lambda \leqslant \frac{\eta^{2.5}}{8k^4}$. Then for any $f_i : X[i] \to \{0,1\}$ and $\bar{f} = (f_1, f_2, \dots, f_k)$:*

$$d_{TV}(\bar{f}_* D_X, \bar{f}_* D_{ind}) \leqslant \eta.$$

Proof. Note that both distributions are supported in $\{0, 1\}^k$ by our assumption on the domain of $\bar{f}$. Thus we need to show for every $B \subseteq \{0, 1\}^k$:

$$\left| \mathop{\mathbb{P}}_{\bar{f}_* D_X} [B] - \mathop{\mathbb{P}}_{\bar{f}_* D_{ind}} [B] \right| \leqslant \eta.$$

Indeed since $\mathbb{P}_{\bar{f}_* D_X} [B] - \mathbb{P}_{\bar{f}_* D_{ind}} [B] = \mathbb{P}_{\bar{f}_* D_{ind}} [\bar{B}] - \mathbb{P}_{\bar{f}_* D_X} [\bar{B}]$ for $\bar{B} = \{0, 1\}^k \setminus B$, it is enough to just show the one-sided bound

$$\mathop{\mathbb{P}}_{\bar{f}_* D_X} [B] \geqslant \mathop{\mathbb{P}}_{\bar{f}_* D_{ind}} [B] - \eta.$$

Fix any event $B \subseteq \{0, 1\}^k$. We'll decompose B into two parts. First, we'll look at the subset B_1 on which f takes many 'unlikely' values, and argue the difference is small just because D_X and D_{ind} have the same marginals over each coordinate of $\{0, 1\}^k$. In particular, let $I = \left\{ i \in [k] \,\middle|\, \mathbb{E}[f_i] \leqslant \frac{\eta}{2k} \right\}$ and $J = \left\{ i \in [k] \,\middle|\, \mathbb{E}[f_i] \geqslant 1 - \frac{\eta}{2k} \right\}$ be the sets of coordinates with extreme expectations and let

$$B_1 = B \cap \left\{ x \,\middle|\, \left(\bigvee_{i \in I} \{f_i(x) = 1\} \right) \vee \left(\bigvee_{j \in J} \{f_j(x) = 0\} \right) \right\}.$$

Set $B_2 = B \setminus B_1$.

We first show that $\mathbb{P}_{\bar{f}_* D_X} [B_1] - \mathbb{P}_{\bar{f}_* D_{ind}} [B_1] \leqslant \frac{\eta}{2}$. Since both distributions have the same marginals, we can write for either D_X or D_{ind} by a union bound:

$$\mathbb{P}[B_1] \leqslant \sum_{i \in I} \mathbb{P}[f_i = 1] + \sum_{j \in J} \mathbb{P}[f_j = 0] \leqslant (|I| + |J|) \frac{\eta}{2k} \leqslant \frac{\eta}{2}.$$

In particular the difference is no more than $\frac{\eta}{2}$.

Moving on to B_2, we decompose the probability as a sum over k-wise products:

$$\mathop{\mathbb{P}}_{\bar{f}_* D_X} [B_2] = \sum_{\bar{b} \in B_2} \mathop{\mathbb{P}}_{D_X} \left[\bigwedge_{i=1}^{k} f_i = b_i \right] = \sum_{\bar{b} \in B_2} \mathop{\mathbb{E}}_{\{v_1, v_2, \ldots, v_k\} \sim D_X} \left[\prod_{i=1}^{k} |b_i - f_i(v_i)| \right].$$

We note that the functions $|b_i - f_i(v_i)|$ have values in $[0, 1]$ and by assumption have expectation $\geqslant \frac{\eta}{2}$ (since otherwise $\bar{b} \in B_1$). Then by [Proposition 7.2](#) setting $\varepsilon = \frac{\eta}{2}$ we have

$$\mathop{\mathbb{P}}_{D_X} [B_2] \geqslant (1 - \frac{\eta}{2k})^k \sum_{\bar{b} \in B_2} \prod_{i=1}^{k} \mathop{\mathbb{E}}_{v_i \in X[i]} [|b_i - f_i(v_i)|]$$

$$= (1 - \frac{\eta}{2k})^k \sum_{\bar{b} \in B_2} \prod_{i=1}^{k} \mathop{\mathbb{P}}_{D_{ind}} \left[\mathbf{1}_{f_i = b_i} \right]$$

$$= (1 - \frac{\eta}{2k})^k \sum_{\bar{b} \in B_2} \prod_{i=1}^{k} \mathop{\mathbb{P}}_{D_{ind}} \left[\mathbf{1}_{f_i = b_i} \right]$$

$$\geqslant \mathop{\mathbb{P}}_{D_{ind}} [B_2] - \frac{\eta}{2}$$

where we've used the marginal equivalence of D_X and D_{ind} and independence of D_{ind} in the second and third steps respectively. Combining our bounds on B_1 and B_2 gives the result. $\qquad \square$

We can now easily prove [Theorem 7.1](#).

Proof of [Theorem 7.1](#). Let $f : X(1) \to [0, 1]$. For the partite case, set $f_i = f|_{X[i]}$. Consider the set of 'bad' k-tuples that mis-sample f:

$$B = \left\{ \left| \sum_{j=1}^{k} f(v_j) - \mu \right| > \varepsilon \right\}.$$

By Lemma 7.5, the TV-distance between the distribution over values of $\bar{f}(x) = f(x_1, x_2, \ldots, x_k)$ over X and $X^{ind,p}$ is at most η so we have:

$$
\begin{aligned}
\sigma(X, k, \varepsilon) &= \max_{\bar{f}} \; \mathbb{P}_{\bar{f}_* D_X} [B] \\
&\leqslant \max_{\bar{f}} \; \mathbb{P}_{\bar{f}_* D_{ind}} [B] + \eta \\
&\leqslant S(X^{ind,p}, k, \varepsilon) + \eta
\end{aligned}
$$

as desired. The two-sided case follows the same argument after taking the partitification $P(X)$ of X and setting $f_i(v, i) = f(v)$. $\qquad\square$

7.2 Lipschitz Concentration for 'Weak' HDX

In this section we study concentration of measure under weaker quantitative assumptions on the underlying complex, requiring only that the down-up walk of X has a good spectral gap. Our results in this special case are incomparable to our prior strategies. While we cannot show Chernoff (subgaussian) strength bounds directly in this setting, we are able to give exponential concentration for the substantially broader class of *Lipschitz* functions. To achieve the subgaussian bounds needed for Theorem 5.1, we pass to lower level skeletons. Before stating the formal results we briefly recall some notation from Section 4.6.

Given a function $f : X(k) \to \mathbb{R}$, let $Z = f(x_1, \ldots, x_k)$ and $Z'_{(i)} = f(x_1, \ldots z_i, \ldots, x_k)$ where z_i is sampled conditional on x_{-i}. We call a f ν-Lipschitz if with probability 1 over Z, Z'

$$
\sum_{i=1}^{k} (Z - Z'_{(i)})_+^2 \leqslant \nu
$$

and say it has ν-bounded difference if for all adjacent (s, s') in the down-up walk:

$$
(f(s) - f(s'))^2 \leqslant \frac{\nu}{k}
$$

We prove that ν-Lipschitz functions on HDX satisfy exponential concentration of measure.

Theorem 7.6. *Let X be a d-uniform simplicial complex such that $U_{d-1}D_d$ has spectral gap at least $\frac{1}{Cd}$ for some $C > 0$. Then for any $0 \leqslant k \leqslant d$ and ν-Lipschitz function $f : X(k) \to \mathbb{R}$:*

1. *Upper Tail:* $\mathbb{P}[f - \mathbb{E}[f] \geqslant t] \leqslant 2e^{-t/\sqrt{(C+1)\nu}}$

2. *Lower Tail:* $\mathbb{P}[f - \mathbb{E}[f] \leqslant -t] \leqslant 2e^{-t/\sqrt{(C+1)\nu}}$

As (almost) immediate corollaries, we get exponential concentration for TD and SI complexes.

Corollary 7.7 (Concentration up to the TD Barrier). *Let X be a d-uniform λ-TD complex for $\lambda < 1$. Then for any $0 \leqslant k \leqslant d$ and ν-Lipschitz function $f : X(k) \to \mathbb{R}$*

1. **Upper tail:** $\mathbb{P}[f - \mathbb{E}[f] \geqslant t] \leqslant 2e^{-\frac{t}{\sqrt{c_\lambda \nu}}}$

2. **Lower tail:** $\mathbb{P}[f - \mathbb{E}[f] \leqslant -t] \leqslant 2e^{-\frac{t}{\sqrt{c_\lambda \nu}}}$

where $c_\lambda \leqslant 1 + e^{\frac{\lambda}{(1-\lambda)}}$.

Corollary 7.8 (Concentration under Spectral Independence). *Fix $\eta > 0$ and let X be a d-uniform η-spectrally independent complex. Then for any $0 \leqslant k \leqslant d$, $g : X(1) \to [0, 1]$, and $f = U_{1,k}g$:*

1. **Upper tail:** $\mathbb{P}[f - \mathbb{E}[f] \geqslant \varepsilon] \leqslant 4e^{-c_\eta \varepsilon \sqrt{k}}$

2. **Lower tail:** $\mathbb{P}[f - \mathbb{E}[f] \leqslant -\varepsilon] \leqslant 4e^{c_\eta \varepsilon \sqrt{k}}$

where $c_\eta \leqslant \frac{1}{12 \max\{1, \sqrt{\eta}\}}$.

We note it is possible to extend the latter to lifted Lipschitz functions from an appropriate skeleton.

The proof of Theorem 7.6 is via a variant of the *Herbst Argument*, a classical strategy for proving concentration of random variables based on applying functional inequalities like MLSI and LSI to the moment generating function. While sparse simplicial complexes cannot have bounded (modified) log-sobolev constants, a somewhat lesser known variant of the Herbst argument using spectral gap was developed by Aida and Stroock [AS94], Bobkov and Ledoux [BL97], and Boucheron, Lugosi, and Massart [BLM13]. We will give an elementary adaption of their method to HDX achieving the above bounds.

The key to Theorem 7.6 is really the following approximate variant of the Efron-Stein inequality, sometimes called 'approximate tensorization of variance'. We remark that similar inequalities for local-spectral expanders appear in the literature [KM20; Che+21].

Lemma 7.9 (Approximate Efron-Stein Inequality). *Let X be a d-uniform simplicial complex such that $\lambda_2(U_{d-1}D_d) \leqslant 1 - \frac{1}{Cd}$ for some $C > 0$. Then for any $0 \leqslant k \leqslant d$, $f : X(k) \to \mathbb{R}$, and $Z = f(X)$ we have:*

$$Var(f) \leqslant (C + 1) \sum_{i=1}^{k} \mathbb{E}\left[(Z - Z'_{(i)})^2_+\right]$$

Proof. It is sufficient to instead prove that any k-uniform complex whose down-up walk $U_{k-1}D_k$ has gap $\frac{1}{C'k}$ satisfies

$$Var(f) \leqslant C' \sum_{i=1}^{k} \mathbb{E}\left[(Z - Z'_{(i)})^2_+\right].$$

The result then follows from [Lee23, Theorem 3.5], who shows that $\lambda_2(U_{d-1}D_d) \leqslant 1 - \frac{1}{Cd}$ implies $\lambda_2(U_{k-1}D_k) \leqslant 1 - \frac{1}{(C+1)k}$ for any $0 \leqslant k \leqslant d$.

With this in mind, let E_i be the ith averaging operator:

$$E_i f(s) = \underset{s_i \sim X_{s_{-i}}}{\mathbb{E}} [f(s_{-i} \cup s_i)],$$

where we recall s_{-i} is s without the ith vertex in its ordering. Define the ith Laplacian $L_i = I - E_i$, and the total laplacian operator as $L = I - UD = \frac{1}{k+1} \sum L_i$. By assumption on the spectral gap of UD, we have:

$$Var(f) \leqslant Ck\langle f, Lf \rangle$$

On the other hand expanding out the Laplacian gives:

$$Ck\langle f, Lf \rangle = C \sum_{i=1}^{k} \langle f, L_i f \rangle$$

$$= C \sum_{i=1}^{k} \langle L_i f, L_i f \rangle$$

$$= C \sum_{i=1}^{k} \underset{s \sim X(k)}{\mathbb{E}} [Var_s^{(i)}(f)]$$

where $Var_s^{(i)}(f)$ is the variance of f over the link $X_{s_{-i}}$. Finally examining this expected variance, we have:

$$\underset{s \sim X(k)}{\mathbb{E}} [Var_s^{(i)}(f)] = \frac{1}{2}\mathbb{E}[(Z - Z'_{(i)})^2] = \mathbb{E}[(Z - Z'_{(i)})^2_+]$$

where both equalities follow from the fact that Z and $Z'_{(i)}$ are i.i.d conditioned on s_{-i}. Combining with the above completes the proof. $\qquad\square$

We can now prove Theorem 7.6 closely following the approach of [BLM13].

Proof of Theorem 7.6. We prove only the upper tail. The lower tail follows similarly. It is sufficient to prove the following exponential integrability of the moment generating function:

$$\mathbb{E}\left[e^{\frac{1}{\sqrt{C\nu}}(f-\mathbb{E}[f])}\right] \leqslant 2, \tag{7.5}$$

since we then have:

$$\mathbb{P}[f - \mathbb{E}[f] > t] = \mathbb{P}\left[\exp\left(\frac{1}{\sqrt{C\nu}}(f-\mathbb{E}[f])\right) > \exp\left(\frac{1}{\sqrt{C\nu}}t\right)\right]$$
$$\leqslant 2\exp\left(-\frac{1}{\sqrt{C\nu}}t\right)$$

by Markov.

For notational simplicity, let $g = f - \mathbb{E}[f]$ and let $Z = g(X)$. Following [BLM13], we use Efron-Stein to set up a recurrence relating $e^{\lambda Z}$ to $e^{\lambda Z/2}$ for small enough $\lambda > 0$. Applying Efron-Stein to the latter gives:

$$\mathbb{E}[e^{\lambda Z}] - \mathbb{E}[e^{\lambda Z/2}]^2 \leqslant C\sum_{i=1}^{k}\mathbb{E}\left[(e^{\lambda Z/2} - e^{\lambda Z'_{(i)}/2})_+^2\right]$$
$$\leqslant C\sum_{i=1}^{k}\mathbb{E}\left[e^{-\lambda Z}(1 - e^{\lambda(Z'_{(i)}-Z)/2})_+^2\right]$$
$$\leqslant \frac{C\lambda^2}{4}\mathbb{E}\left[e^{-\lambda Z}\sum_{i=1}^{k}(Z - Z'_{(i)})^2\right]$$
$$\leqslant \frac{C\lambda^2\nu}{4}\mathbb{E}\left[e^{-\lambda Z}\right]$$

since f (and therefore g) are ν-bounded. Re-arranging we have the recurrence:

$$\mathbb{E}[e^{\lambda Z}] \leqslant \frac{1}{1 - \frac{\lambda^2 C\nu}{4}}\mathbb{E}[e^{\lambda Z/2}]^2.$$

It is then an elementary exercise to solve when $\lambda \leqslant \frac{1}{\sqrt{C\nu}}$ and derive (7.5) (see [BLM13, Section 3.6] for the full derivation). $\qquad\square$

We now prove the corollaries for SI/TD-complexes

Proof of Corollary 7.7. By Theorem 4.23, X is a $\frac{\lambda}{(d-1)(1-\lambda)}$-one-sided local-spectral expander, so Theorem 4.19 gives $\lambda_2(U_{d-1}D_d) \leqslant 1 - \frac{1}{d}e^{\frac{\lambda}{1-\lambda}}$. Plugging this into Theorem 7.6 gives the result. $\qquad\square$

In the spectral independence regime, one must be slightly more careful since an η-spectrally independent system has $C \approx k^\eta$. We recover concentration by looking only at lifted functions and analyzing concentration on some lower skeleton.

Proof of Corollary 7.8. We prove the upper tail. The lower tail follows similarly. We break the proof into two cases. First, assume that $\eta \leqslant 2$ and let $k' = \lfloor\frac{d}{3}\rfloor$. The k'-skeleton of X is a one-sided $\frac{1}{k'}$-local-spectral expander. Thus by Theorem 4.19 $\lambda(U_{k'-1}D_{k'}) \leqslant 1 - \frac{1}{ek'}$ and Theorem 7.6 implies every $k \leqslant k'$ has concentration:

$$\mathbb{P}[f - \mathbb{E}[f] > \varepsilon] \leqslant 2e^{-\frac{1}{\sqrt{e+1}}\varepsilon\sqrt{k}}.$$

Further, by Lemma 4.15 every $k > k'$ has concentration:

$$\mathbb{P}[f - \mathbb{E}[f] > \varepsilon] \leqslant 2e^{-\frac{1}{2}\frac{1}{\sqrt{e+1}}\varepsilon\sqrt{k'}}(1 - e^{-\frac{\varepsilon^2}{12}k})$$

41

$$\leqslant 4e^{-\frac{1}{2\sqrt{6(e+1)}}\varepsilon\sqrt{d}}$$

$$\leqslant 4e^{-\frac{1}{12}\varepsilon\sqrt{d}}$$

assuming $k \geqslant 6$ and $\varepsilon^2 \geqslant \frac{12}{k}$ (else the bound is trivial).

Otherwise, assume $\eta > 2$ and let $k' = \lfloor \frac{d}{\eta} \rfloor$. The k'-skeleton of X is a $\frac{\eta}{(1-\frac{1}{\eta})d}$-one-sided local-spectral expander. Then by Theorem 4.19 $\lambda(U_{k'-1}D_{k'}) \leqslant 1 - \frac{1}{e^2 k'}$ and Theorem 7.6 implies for every $k \leqslant k'$:

$$\mathbb{P}[f - \mathbb{E}[f] > \varepsilon] \leqslant 2e^{-\frac{1}{\sqrt{e^2+1}}\varepsilon\sqrt{k}}.$$

Further, by Lemma 4.15 every $k > k'$ has concentration:

$$\mathbb{P}[f - \mathbb{E}[f] > \varepsilon] \leqslant 2e^{-\frac{1}{2\sqrt{e^2+1}}\varepsilon\sqrt{k'}}(1 - e^{-\frac{\varepsilon^2}{12}k})$$

$$\leqslant 4e^{-\frac{1}{2\sqrt{4\eta(e^2+1)}}\varepsilon\sqrt{d}}$$

$$\leqslant 4e^{-\frac{1}{12\sqrt{\eta}}\varepsilon\sqrt{d}}$$

assuming $\eta \leqslant \frac{d}{2}$ and $\varepsilon^2 \geqslant \frac{12}{k+1}$ (again, otherwise the stated bound is trivial). $\qquad\square$

7.2.1 Subgaussian Concentration on Skeletons

We now show that it is possible to recover subgaussian concentration from the above analysis by passing to skeletons of size roughly $\Theta(\sqrt{d})$.

Corollary 7.10 (Subgaussian concentration for skeletons)**.** *Let X be a d-uniform simplicial complex with $\lambda_2(U_{d-1}D_d) \leqslant 1 - \frac{1}{Cd}$. Then for any $k \leqslant d$, $\nu > 0$, and function $f : X(k) \to \mathbb{R}$ with ν-bounded difference:*

1. **Upper Tail:** $\mathbb{P}[f - \mathbb{E}[f] \geqslant t] \leqslant 2e^{-\frac{t}{\sqrt{4(C+1)\nu}} \cdot \sqrt{\frac{d}{k}}} + e^{-\frac{t^2}{4\nu}}$

2. **Lower Tail:** $\mathbb{P}[f - \mathbb{E}[f] \leqslant -t] \leqslant 2e^{-\frac{t}{\sqrt{4(C+1)\nu}} \cdot \sqrt{\frac{d}{k}}} + e^{-\frac{t^2}{4\nu}}$

Proof. The proof is immediate from Theorem 7.6 and Lemma 4.16, setting $f_{up}(t,\nu) = f_{low}(t,\nu) = 2e^{\frac{-t}{\sqrt{(C+1)\nu}}}$. $\qquad\square$

Corollary 7.10 does not imply full subgaussian concentration for all functions with bounded difference, but does for certain critical sub-classes such as bounded 1-lifts:

Corollary 7.11. *Let X be a d-uniform simplicial complex with $\lambda_2(U_{d-1}D_d) \leqslant 1 - \frac{1}{Cd}$. Then for any $k \leqslant d$ and $f : X(1) \to [0,1]$ of expectation μ:*

1. **Upper Tail:** $\displaystyle\mathop{\mathbb{P}}_{s \sim X(k)}\left[U_{1,k}f(s) - \mu \geqslant \varepsilon\right] \leqslant 3e^{-\frac{1}{4\sqrt{C+1}}\varepsilon^2 k}$

2. **Lower Tail:** $\displaystyle\mathop{\mathbb{P}}_{s \sim X(k)}\left[U_{1,k}f(s) - \mu \leqslant -\varepsilon\right] \leqslant 3e^{-\frac{1}{4\sqrt{C+1}}\varepsilon^2 k}$

Proof. The proof is immediate from Lemma 4.16 and the observation that $U_{1,k}f$ has $\frac{1}{k}$-bounded difference. $\quad\square$

Note that by similar arguments as above, this implies subgaussian concentration for the skeletons of λ-TD and η-SI complexes. This is already implicit in the statement of Theorem 5.1, since TD/SI-complexes are c-nice since an appropriate skeleton of these complexes are nice (Claim 4.28). Finally, we remark it is likely possible to generalize the above strategy from bounded difference to Lipschitz functions on sufficiently strong local-spectral expanders by more carefully analyzing the eigenspaces of the Laplacian and higher order random walks associated with the d-lift (e.g. as in [KO20; Dik+18; Gai+22]).

7.3 Inclusion Sampling and the Proof of Theorem 5.1

We are finally ready to prove Theorem 5.1 in full. We again restate the theorem in the final two cases for convenience.

Theorem 7.12 (Sampling on HDX). *Let X be a k-uniform complex which is either*

1. *The skeleton of a d-uniform partite 2^{-cd}-one-sided local-spectral expander Y for $d \geq 3k \log k$*

2. *The skeleton of a d-uniform complex Y with $\lambda(U_{d-1}D_d) \leq 1 - \frac{1}{cd}$ for $d \geq 2k^2$*

Then for any $i < k$ the containment graph $(X(k), X(i))$ is a (ε, β) sampler for

$$\beta = \frac{64}{\varepsilon} \exp(-\Omega_c(\varepsilon^2 \frac{k}{i})).$$

Moreover, if X is a k-skeleton of a d-uniform complex Y, then for every $k < k' \leq d$, the graph $(Y(k'), Y(i))$ is also $(\varepsilon, O(\beta))$-function samplers.

Proof of Theorem 7.12 (Theorem 5.1, partite and spectral gap case). In both cases note the 'moreover' part follows directly from Lemma 4.15.

Spectral Gap. Let Y be the promised d-uniform complex such that X is an (at most) $\sqrt{\frac{d}{2}}$-uniform skeleton of Y such that Y's down-up walk has spectral gap $\lambda(U_{d-1}D_d) \leq 1 - \frac{1}{cd}$. For any $g : X(i) \to [0,1]$ Theorem 7.6 implies $U_{i,d}g$ has the following exponential tail:

$$\mathbb{P}[U_{i,d}g(s) - \mathbb{E}[g] \geq \varepsilon] \leq 2e^{-\Omega(\varepsilon \frac{\sqrt{d}}{i})} \leq 2e^{-\Omega(\varepsilon \frac{k}{i})}$$

since $U_{i,d}g$ has $\frac{i^2}{d}$-bounded difference. By Lemma 4.16, $U_{i,k}g$ therefore has tail:

$$\mathbb{P}[U_{i,k}g(s) - \mathbb{E}[g] \geq \varepsilon] \leq 2e^{-\Omega(\varepsilon \frac{k}{i})} + \frac{2}{\varepsilon}e^{-\Omega(\varepsilon^2 \frac{k}{i})},$$

where the latter term is from the following concentration bound on the complete complex:

Claim 7.13. For any $i \leq k \leq n$ and $g : \Delta_n(i) \to [0,1]$:

$$\mathbb{P}_{\Delta_n(k)}[U_{i,k}g - \mathbb{E}[g] > \varepsilon] \leq \frac{2}{\varepsilon}e^{-\Omega(\varepsilon^2 \frac{k}{i})}$$

and likewise for the lower tail.

We prove this claim in Appendix A via a simple application of Lemma 5.9. The lower tail follows similarly.

Partite HDX. The partite case of Theorem 5.1 introduces a new challenge. We cannot directly reduce to the complete complex since we are not on a sufficiently low dimensional skeleton, and the faces complex of X isn't a (sufficiently strong) HDX to blackbox apply Theorem 5.2. Nevertheless, lacking expansion does not necessarily mean the faces complex is a poor sampler. We will argue F_X has nearly subgaussian concentration by arguing its sampling can be split into two components 1) sampling a good partition of *colors* 2) sampling a good subset conditioned on this partition. We show the first step may be reduced to concentration on the faces complex of the complete complex, while the second reduces to concentration for partite HDX (Theorem 7.1).

Formally, we will start by showing that the kth level of a d-partite $2^{-\Omega(d)}$-HDX has nearly optimal concentration for lifted functions when $k = \lfloor \frac{d}{3} \rfloor$. That is for any $f : X(\ell) \to [0,1]$:

$$\mathbb{P}_{X(k)}[|U_{\ell,k}f - \mathbb{E}[f]| > \varepsilon] \leq 5 \exp\left(\frac{k}{\ell} \cdot \frac{1}{8 \log(ek)}\right). \tag{7.6}$$

The stated result is then an immediate consequence of Lemma 4.16.

The key to Equation (7.6) is the concentration properties of the faces complex of the complete complex, an object we'll call the *swap complex*.

Definition 7.14 (The Swap Complex). Let $n \geqslant k \geqslant \ell \in \mathbb{N}$. The (ℓ, k, n)-swap complex is the $\lfloor \frac{k}{\ell} \rfloor$-uniform simplicial complex $C = C_{\ell,k,n}$ whose vertices are all size-ℓ subsets $[n]$:

$$C(1) = \binom{[n]}{\ell},$$

and whose top-level faces are all possible pair-wise disjoint ℓ-sets:

$$C\left(\left\lfloor \frac{k}{\ell} \right\rfloor\right) = \left\{ \{s_1, s_2, \ldots, s_{\lfloor \frac{k}{\ell} \rfloor}\} : \bigcup_{i \neq j} s_i \cap s_j = \emptyset \right\}$$

endowed with the uniform distribution.

In other words, $C_{\ell,k,n}$ is exactly $F^\ell_{\Delta_n(k)}$, the faces complex of of the k-uniform complete complex on n vertices. We prove that the swap complex satisfies (nearly) subgaussian concentration of measure.

Theorem 7.15. *Let $C = C_{\ell,k,n}$ for $n \geqslant 3k$ and $f : C\left(\left\lfloor \frac{k}{\ell} \right\rfloor\right) \to \mathbb{R}$ be any ν-Lipschitz function. Then:*

1. **Upper tail:** $\mathbb{P}[f - \mathbb{E}[f] > t] \leqslant \exp(\frac{t^2}{\nu} \cdot \frac{1}{2\log(ek)})$

2. **Lower tail:** $\mathbb{P}[f - \mathbb{E}[f] < -t] \leqslant \exp(\frac{t^2}{\nu} \cdot \frac{1}{2\log(ek)})$

The proof of Theorem 7.15 follows via bounding the log-sobolev constant of the down-up walk on C and applying the entropic Herbst argument. The proof is given in Appendix B. We remark that the extra log factor in the exponent above is the only barrier to proving optimal concentration for the top level of partite HDX. This is discussed further in Appendix B, where we observe the result would be implied by showing optimal entropy contraction for the well-studied Kneser graphs.

We are now ready to handle the partite case of Theorem 7.1. Let X be a k-uniform skeleton of a d-partite $2^{-\Omega(d)}$-HDX for $3k \leqslant d$. Recall by Proposition 5.8 it is enough to prove the $F^{(\ell)}_X$ is an $(\varepsilon, 8e^{-\Omega(\varepsilon^2 \frac{k}{\ell})})$ sampler for any $\varepsilon > 0$.

Fix $f : X(\ell) \to [0, 1]$ of expectation μ viewed as a function on the vertices of the faces complex, and let $m = \lfloor \frac{k}{\ell} \rfloor$. Consider the random variable $\mathcal{J} := (col(s_1), col(s_2), \ldots, col(s_m))$ corresponding to the partition generated by drawing $(s_1, \ldots, s_m)$ from the faces complex of X. For a fixed $\mathcal{J} = \{I^1, I^2, \ldots, I^m\}$, denote the conditional expectation of f under $\mathcal{J}$ as

$$\mu_{\mathcal{J}} := \mathop{\mathbb{E}}_{s \sim X(\ell)}[f \mid col(s) \in \mathcal{J}].$$

Further let $C_{\mathcal{J}}$ denote the complex generated by restricting to top-level faces of C with partition $\mathcal{J}$. We define two 'bad' events based on $\mathcal{J}$ outside which our sampling is ε-accurate:

1. **(Bad Colors):** E_1 the event that $\mu_{\mathcal{J}}$ deviates significantly from μ
$$|\mu_{\mathcal{J}} - \mu| \geqslant \varepsilon/2.$$

2. **(Bad Sampling):** E_2 the event that $(s_1, \ldots, s_m) \sim C_{\mathcal{J}}$ samples $\mu_{\mathcal{J}}$ poorly
$$\left| \frac{1}{m} \sum_{i \in [m]} f(s_i) - \mu_{\mathcal{J}} \right| \geqslant \varepsilon/2.$$

It is then sufficient to prove the following three claims:

1. $\mathbb{P}_{(s_1, s_2, \ldots, s_m) \sim C}\left[\left| \frac{1}{m} \sum_{i \in [m]} f(s_i) - \mu \right| > \varepsilon \right] \leqslant \mathbb{P}[E_1] + \mathbb{P}[E_2 \wedge \neg E_1].$

2. $\mathbb{P}[E_1] \leqslant 2\exp\left(-\frac{\varepsilon^2 m}{8\log(ek)}\right).$

3. $\mathbb{P}[E_2 \wedge \neg E_1] \leqslant 4\exp(-\varepsilon^2 m).$

The first claim is immediate from definition, as the event $|\sum f - \mu| > \varepsilon$ only occurs if $E_1 \vee E_2$ holds, the probability of which is bounded by $\mathbb{P}[E_1] + \mathbb{P}[E_2 \wedge \neg E_1]$ as desired.

Bounding $\mathbb{P}[E_1]$. Observe $\mathcal{J}$ is equidistributed with the complex $C_{\ell,k,d}$ and define $g : C(1) \to [0,1]$ by $g(I^j) = \mathbb{E}_C[f \mid I^j \in col(s)]$. For $\mathcal{J} = \{I^1, \ldots, I^m\}$ we then have:

1. $\mathbb{E}[g] = \underset{I \sim C(1)}{\mathbb{E}} [\mathbb{E}[f \mid col(s) \in I]] = \mu$

2. $\frac{1}{m} \sum_{j=1}^{m} g(I^j) = \mu_{\mathcal{J}}.$

Combining these observations with Theorem 7.15 we can write

$$\mathbb{P}[E_1] = \underset{\mathcal{J}}{\mathbb{P}} \left[|\mu_{\mathcal{J}} - \mu| > \frac{\varepsilon}{2} \right]$$

$$= \underset{\{I^1,\ldots,I^m\} \sim C(m)}{\mathbb{P}} \left[\left| \frac{1}{m} \sum_{j=1}^{m} g(I^j) - \mathbb{E}[g] \right| > \frac{\varepsilon}{2} \right]$$

$$\leqslant 2 \exp \left(- \frac{\varepsilon^2 m}{8 \log(ek)} \right).$$

since $\frac{1}{m} \sum_i g(s_i)$ is $\frac{1}{m}$-bounded.

Bounding $\mathbb{P}[E_2 \wedge \neg E_1]$. It is left to argue that for any fixed $\mathcal{J} = \{I^1, I^2, \ldots, I^m\}$, $C_{\mathcal{J}}$ is a partite $2^{-\Omega(k)}$-HDX. The result is then immediate from applying Theorem 7.1 with $\eta = \exp(-\frac{1}{4}\varepsilon^2 m)$:

$$\mathbb{P}[E_2 \wedge \neg E_1] \leqslant \underset{\{s_1,\ldots,s_m\} \sim D}{\mathbb{P}} [E_2 \mid \mathcal{J} \in \neg E_1]$$

$$= \underset{\mathcal{J} \in \neg E_1}{\mathbb{E}} \left[\underset{s \sim C_{\mathcal{J}}}{\mathbb{P}} \left[\left| \frac{1}{m} \sum_{i \in [m]} f(s_i) - \mu_{\mathcal{J}} \right| \geqslant \frac{\varepsilon}{2} \right] \right]$$

$$\leqslant 3 \exp \left(- \frac{1}{4} \varepsilon^2 m \right).$$

By construction it is easily checked $C_{\mathcal{J}}$ is a partite complex whose links are isomorphic to swap walks S_{I_i,I_j} within a link of the original complex X. By Theorem 4.21, $C_{\mathcal{J}}$ is therefore a partite λ-one-sided HDX for $\lambda \leqslant 2^{-\Omega(d)}$ as well. $\qquad \square$

8 Agreement Testing

In this section we use our reverse hypercontractive inequality to prove several new agreement testing theorems on HDX. In the first subsection, we give new agreement tests in the 99%-regime between any two levels of an HDX. In the second subsection, we explore the role of reverse hypercontractivity in the 1%-regime, and prove that HDX admit agreemment tests with optimal *local* soundness (a weaker notion of soundness only requiring globalness on links), and show under the stronger assumption of ℓ_∞-expansion that this guarantee can be propagated to a true tester with optimal soundness.

8.1 Background

Suppose U is some finite set of vertices and $S \subseteq \mathbb{P}(U)$ are subsets of U. An *agreement test* is a procedure that inputs a family of assignments to the subsets $\{f_s : s \to \{0,1\}\}_{s \in S}$, and aims to check whether the assignments are consistent with a *global* assignment to the ground set $G : U \to \{0,1\}$. This is a classical setup in hardness of approximation (in particular in the construction of PCPs), where one repeats a problem in parallel and needs to check that answers across repetitions are consistent.

Let us give the simplest setup first, slightly generalizing it later on. Let (U, S, Σ, D) be such that

1. U is a finite set ("Universe").

2. $S \subseteq \mathbb{P}(U)$ ("Sets").

3. Σ is another finite set ("Alphabet").

4. A distribution $s_1, s_2 \sim D$ such that $s_1, s_2 \in S$ ("Distribution").

An *ensemble of functions* in this context is a family $\mathcal{F} = \{f_s : s \to \Sigma\}_{s \in S}$. An ensemble is called *global* if there exists a function $G : U \to \Sigma$ such that for every $s \in S$ it holds that $f_s = G|_s$.

We wish to relate the following two quantities. The first quantity is the "agreement" of an ensemble. For $\eta \in [0, 1)$ we denote by

$$Agree_\eta(\mathcal{F}) := \mathop{\mathbb{P}}_{s_1, s_2, t \sim D} \left[f_{s_1}|_{s_1 \cap s_2} \overset{\eta}{\approx} f_{s_2}|_{s_1 \cap s_2} \right], \tag{8.1}$$

where the notation inside the probability means that there is a set $t' \subseteq s_1 \cap s_2$ of size $|t'| \geqslant (1 - \eta)|s_1 \cap s_2|$ such that $f_{s_1}|_{t'} = f_{s_1}|_{t'}$. We will also use $\overset{\eta}{\not\approx}$ to denote the negation of this.

The second quantity is the distance from global functions. For $\eta, \varepsilon > 0$ we say that $\mathcal{F}$ is (η, ε)-close to a global function if the exists some $G : U \to \Sigma$ such that

$$\mathop{\mathbb{P}}_s \left[f_s \overset{\eta}{\approx} G|_s \right] \geqslant 1 - \varepsilon.$$

The probability distribution in which we take $s \in S$ above is the marginal distribution of D.[9]

We denote by

$$\mathrm{dist}_\eta(\mathcal{F}, Glob) = \min \left\{ \varepsilon \geqslant 0 \,|\, \text{ there exists } G : U \to \Sigma \text{ such that } \mathcal{F} \text{ is } (\eta, \varepsilon)\text{-close to } G \right\}.$$

A *high acceptance agreement theorem* (also known as a "99% agreement theorem") is a theorem that states that if a set of local functions pass the test with almost perfect probability ("99%"), then the functions are close to being global. In other words, this is a theorem that states for a specific (U, S, Σ, D) that for every ensemble $\mathcal{F}$,

$$Agree_\eta(\mathcal{F}) = 1 - \varepsilon \implies \mathrm{dist}_{\eta'}(\mathcal{F}, Glob) \leqslant \varepsilon'.$$

where $\varepsilon, \varepsilon', \eta, \eta' > 0$. Ideally, one wants such a statement for every ε, η small enough, with $\varepsilon' = O(\varepsilon)$ and $\eta' = O(\eta)$. But other parameter regimes are interesting in applications as well.

Agreement tests originated in low-degree testing results such as [RS96] and were later abstracted by [GS97] to roughly the setup above. They are an important component of Dinur's proof of the PCP Theorem [Din07]. An optimal high acceptance agreement theorem was proven by [DS14] on the complete complex. This was extended by [DFH19] to another setup related to the complete complex we discuss below. Agreement theorems on sparse complexes such as high dimensional expanders are also known, pioneered by [DK17].

A *low acceptance agreement theorem* (also known as a "1% agreement theorem") deals with the regime where it is only assumed that the functions pass the agreement test with small (sometimes even sub-constant) probability ("1%"). This is a theorem of the form

$$Agree_\eta(\mathcal{F}) = \delta \geqslant \varepsilon \implies \mathrm{dist}_{\eta'}(\mathcal{F}, Glob) \leqslant 1 - \mathrm{poly}(\delta).$$

for some $(\varepsilon, \eta, \eta')$ (and some explicit polynomial in δ). We note here that in this test even when $\eta = 0$, an $\eta' > 0$ is sometimes unavoidable (see discussions in [DG08b; DL17]).

Agreement theorems in the low acceptance regime are typically harder to prove than in the high acceptance regime, since the promised structure on the family of functions is mild. However, in applications such that a parallel repetition theorem [IKW12] and low-degree testing [AS97], they are often crucial. Dinur and Goldenberg [DG08b] proved a low acceptance agreement theorem on the complete complex. Their proof was simplified by [IKW12], who also derived a more sophisticated Z-test (see below) and gave the first sparse structure supporting such agreement tests using subspaces. Work by [DL17] gave an improved analysis of [IKW12], asymptotically matching the lower bound of a random set of local functions. Their work heavily relied on reverse hypercontractivity, a strategy we follow below. Recently, [DD23b] and [BM23b] gave characterizations of high dimensional expanders that support agreement theorems, and [DDL24; BLM24] constructed such complexes.

---

[9]Formally, draw $(s, s') \sim D$, then s or s' with probability $1/2$ independently.

There are also other kinds of theorems about agreement tests that don't strictly fall into any of these categories. One example such as [GK22] list agreement theorem, that inspired the former line [DD23b; BM23b; DDL24; BLM24].

8.1.1 Varying the Test

We present two simple extensions to the agreement test above. The first comes from sampling more than two sets. For example, in the Z-test by [IKW12], three sets s_1, s_2, s_3 are sampled, and the test passes if $f_{s_1} = f_{s_2}$ and $f_{s_2} = f_{s_3}$ (equality is with respect to the respective intersections). See the formal definition of this test below.

Another extension to the definition of a test, mainly for technical reasons, is a test that checks $f_{s_1}|_t = f_{s_2}|_t$ only on some $t \subset s_1 \cap s_2$ which is also sampled by the tester. This is sometimes more convenient to analyze than the test checking agreement on the whole intersection. To summarize this discussion more formally, the distribution D in our setup samples $(\{s_i\}_{i=1}^k, \{t_{i,j}\}_{1 \leqslant i < j \leqslant k})$ such that $\{s_i\}_{i=1}^k \subseteq S$ are sets, and for every $i < j$, $t_{i,j} \subseteq s_i \cap s_j$. The agreement test passes if for all $i < j$, $f_{s_i}|_{t_{i,j}} \overset{\eta}{\approx} f_{s_j}|_{t_{i,j}}$.

8.2 Agreement Testing for Subsets

We now move to the first setting of interest, agreement tests on simplicial complexes in the 99% regime. In particular we study the following setup due to [DFH19]. Let $j < k < d$ and let X be a d-uniform simplicial complex. Let us denote by $SS(X, \Sigma, j, k, d) := (U, S, \Sigma, D)$ where:

1. $U = X(j)$.

2. $S = \left\{ \binom{s}{j} \ \middle| \ s \in X(d) \right\}$.

3. The distribution $D = D_{d,k}$ samples s_1, s_2 according to the down-up walk. The 'test' set $t \subseteq s_1 \cap s_2$ is a randomly chosen k-face.

For simplicity of notation we sometimes refer to the sets as s instead of $\binom{s}{j}$, and also index functions by f_s for $s \in X(d)$ (and not $f_{\binom{s}{j}}$). However, in this context when writing $v \in s$, we mean a j-set instead of a vertex.

In other words, we study the following variant of the classical 'V-test' in this setting:

The V-test on $X(k)$:

1. Draw $t \in X(k)$, and $s_1, s_2 \in X_t(d)$ independently

2. Accept if:
$$f_{t \cup s_1}|_t \overset{\eta}{\approx} f_{t \cup s_2}|_t$$

When $j = 1$, this reduces to the heavily studied setting of (de-randomized) 'direct product testing'.

We are now ready state the main result of the subsection, our generalized 99%-tester:

Theorem 8.1. *Let j, d be integers, let $\gamma, \eta, c, \nu < 1$. Denote by $k = \gamma d$. Suppose that $j < k$, $\eta \lfloor \frac{k}{j} \rfloor \geqslant 192 \log \frac{16}{\eta}$ and $\nu < 1$. Let X be a d-uniform c-locally nice complex such that for every link of a set $v \in X(j)$, $\lambda(U D_{d-j,k-j}(X_v)) \leqslant \nu$. Then the following holds for the set system $SS(X, \Sigma, j, k, d)$; for every alphabet Σ, family $\mathcal{F}$, and $\varepsilon > 0$*

$$Agree_\eta(\mathcal{F}) \geqslant 1 - \varepsilon \quad \implies \quad \mathrm{dist}_{8\frac{\eta+\varepsilon}{1-\nu}}(\mathcal{F}, g_{maj}) \leqslant \varepsilon^{O_{\gamma,c}(1)} + 2^{-\Omega_{\gamma,c}(d)} \tag{8.2}$$

where $g_{maj} := \underset{s \in X(d): s \supset v}{plurality} \{f_s(v)\}$ *is the plurality decoding.*[10]

---
[10]Note that ties may be broken arbitrarily and the weighting is according to the (normalized) complex weights π_d.

We note that every c-nice complex has $\lambda(UD_{d-j,k-j}) \leqslant \frac{k-j}{d-j}(1 + \frac{1}{c'(d-j)})^{(1-\gamma)(d-j)}$, so this is not an extra requirement in most cases (see [Lee23, Theorem 3.5]).

We remark that for the setup of $j = 1$ there exist agreement theorems that achieve (exact) closeness $1 - O(\varepsilon)$ for high dimensional expanders [DK17; DD19; KM20]. Theorem 8.1 is incomparable to these results. On the one hand, in Theorem 8.1 the resulting 'globalness' is weaker than [DK17; DD19; KM20] where one gets a guarantee of $\text{dist}_0(\mathcal{F}, g_{maj}) \leqslant O(\varepsilon)$. On the other hand, the initial assumption on the sets of local functions in the aforementioned results is $Agree_0(\mathcal{F}) \geqslant 1 - \varepsilon$, wheras this result assumes $Agree_\eta(\mathcal{F}) \geqslant 1 - \varepsilon$ for $\eta > 0$. This relaxed assumption is crucial in the low-soundness regime theorem we give in the next section.

For $j > 1$, to our knowledge Theorem 8.1 is the first agreement theorem beyond the complete complex, where Dinur, Filmus, and Harsha [DFH19] showed

$$Agree_0(\mathcal{F}) \geqslant 1 - \varepsilon \Rightarrow \text{dist}_0(\mathcal{F}, Glob) \geqslant 1 - O_j(\varepsilon). \tag{8.3}$$

It is plausible that, with more effort, (8.3) could be proven directly on HDX. The advantage of our method in this sense lies in its generality and (relative) simplicity.

Proof of Theorem 8.1. Given a family $\mathcal{F} = \{f_s\}$ with high agreement, we show it can be (approximately) decoded to the majority function $G = g_{\text{maj}}$. We argue this can be inferred directly from reverse hypercontractivity and the following claim bounding the *average* probability of disagreement with G for a random pair $v \subseteq s$ drawn from $(X(j), X(d))$:

Claim 8.2. $\quad \mathbb{P}_{v \subset s}[f_s(v) \neq G(v)] \leqslant \frac{\varepsilon+\eta}{1-\nu}$.

We first show this implies the result. Define the set of 'bad' d-faces which noticeably disagree with G as

$$B := \left\{ s \in S \,\middle|\, \mathbb{P}_{v \in s}[f_s(v) \neq G(v)] > \frac{8(\varepsilon+\eta)}{1-\nu} \right\}.$$

It is enough to show $\mathbb{P}[B] \leqslant \max\{\varepsilon^{O_{\gamma,c}(1)}, \exp(-\Omega_{\gamma,c}(d))\}\}$.

Toward bounding B, define an intermediary set A of 'good' faces:

$$A := \left\{ s \in S \,\middle|\, \mathbb{P}_{v \in s}[f_s(v) \neq G(v)] \leqslant \frac{4(\varepsilon+\eta)}{1-\nu} \right\}$$

and note that $\mathbb{P}[A] \geqslant \frac{1}{2}$ by Claim 8.2 and Markov's Inequality. The idea is to argue that, on the one hand, reverse hypercontractivity implies the number of edges (s_1, s_2) between A and B is at least some power of $\mathbb{P}[B]$, while, on the other hand, agreement implies the number of such edges is at most $O(\varepsilon)$ because they (typically) contain many $v \in s_1 \cap s_2$ that disagree and therefore fail the test.

More formally, denote the event that $(s_1, s_2) \in (A \times B)$ as $E_{A,B}$. Assuming $\mathbb{P}[B] \geqslant \exp(-c'd)$ (for c' depending on c, γ as in Theorem 6.4), reverse hypercontractivity lower bounds the measure of this event by

$$\mathbb{P}[E_{A,B}] \geqslant \left(\frac{\mathbb{P}[B]}{2}\right)^q$$

for some constant q depending only on c, γ.

On the other hand, since the test rejects with probability ε, we can upper bound the measure of $E_{A,B}$ by $O(\mathbb{P}[\text{Test rejects}]) = O(\varepsilon)$ we will obtain a bound of the form $\mathbb{P}[B] \leqslant O(\varepsilon^{1/q})$. To this end, note

$$\mathbb{P}[E_{A,B}] \leqslant \frac{\varepsilon}{\mathbb{P}[\text{Test rejects}|\ E_{A,B}]}.$$

It is therefore enough to argue the conditional rejection probability is at least some constant (say $\frac{1}{2}$). Toward this end, for a face $s \in X(d)$, define

$$T_s := \{v \in s : f_s(v) \neq G(v)\}$$

to be the j-faces on which f_s disagrees with majority, and recall that an edge in our test is sampled by first picking $s_1 \in X(d)$, then $t \subset s_1$ in $X(\gamma d)$, then $s_2 \supset t$ conditionally. Given t, denote by $t(j) = \binom{t}{j}$ the set of sub j-faces in t, for shortness. We define two bad events outside of which the test rejects and prove they occur with vanishing probability.

48

1. **Mis-sampling A:** E_1 the event that t sees 'too many' $v \subset s_1$ that disagree with majority:

$$E_1 := \left\{ \frac{|t(j) \cap T_{s_1}|}{\binom{k}{j}} \geqslant 5 \frac{\varepsilon + \eta}{1 + \gamma} \right\}$$

2. **Mis-sampling B:** E_2 the event that t sees 'too few' $v \subset s_2$ that disagree with majority:

$$E_2 := \left\{ \frac{|t(j) \cap T_{s_2}|}{\binom{k}{j}} \leqslant 7 \frac{\varepsilon + \eta}{1 + \gamma} \right\}$$

Conditioning on $\neg E_1$ and $\neg E_2$ the test rejects since s_1 and s_2 disagree on a $2\frac{\varepsilon+\eta}{1-\nu}$-fraction of t. Thus by a union bound:

$$\mathbb{P}[\text{Test Rejects}|\ E_{A,B}] \geqslant 1 - \mathbb{P}[E_1|\ E_{A,B}] - \mathbb{P}[E_2|\ E_{A,B}]$$

and it is enough to show that each $\mathbb{P}[E_i|\ E_{A,B}] \leqslant \frac{1}{4}$

Bounding E_1 and E_2. Events E_1 and E_2 occur with vanishing probability due to the sampling properties of the complete complex Δ (Claim A.1), namely that for any $\alpha, \delta > 0$ and $j \leqslant k$, the inclusion graph $(\Delta(k), \Delta(j))$ is an $(\alpha, \frac{4}{\alpha\delta} \exp(\frac{\delta^2}{12}\alpha \lfloor \frac{k}{j} \rfloor), \delta)$-sampler.

With this in mind, recall $t \in X(\gamma d)$ is drawn uniformly at random from s_1, so we may equivalently view t as being drawn from a γd-uniform complete complex Δ on the d vertices of s_1. By assumption, $s_1 \in A$, so $T_{s_1} \subset \Delta(j)$ is of measure at most $4\frac{\varepsilon+\eta}{\gamma}$, and sampling then implies

$$\mathbb{P}\left[\frac{|t(j) \cap T_{s_1}|}{\binom{k}{j}} > 5\frac{\varepsilon+\eta}{1+\gamma} \right] \leqslant \frac{4(1+\gamma)}{\varepsilon+\eta} \exp\left(-\frac{\eta}{96} \left\lfloor \frac{k}{j} \right\rfloor \right) \leqslant \frac{1}{4}.$$

Since we can equivalently sample an edge (s_1, s_2) by first picking s_2, then $t \subset s_2$, then $s_1 \supset t$, E_2 can be bounded using the lower tail of Δ in exactly the same fashion to get

$$\mathbb{P}\left[\frac{|t(j) \cap T_{s_2}|}{\binom{k}{j}} < 7\frac{\varepsilon+\eta}{1+\gamma} \right] \leqslant \frac{4(1+\gamma)}{\varepsilon+\eta} \exp\left(-\frac{\eta}{192} \left\lfloor \frac{k}{j} \right\rfloor \right) \leqslant \frac{1}{4}.$$

$\square$

Proof of Claim 8.2. Fix $v \in X(j)$ and let $D_{d,k}^{(v)}$ denote the conditional distribution over $(s_1, t, s_2) \sim D_{d,k}$ such that $v \subset t$. Note that by construction, $D_{d,k}^{(v)}$ (after removing v from each face) is distributed exactly as the down-up walk $D_{d-j,k-j}$ within the link of v. Partition the vertices of this graph into sets $S_\sigma^{(v)} := \{s|\ f_s(v) = \sigma\}_{\sigma \in \Sigma}$ by the alphabet value they assign v, and observe that since the marginals of $D_{d,k}^{(v)}$ are distributed as $X_v(d-j)$ we have

$$1 - \mathbb{P}_{v \subset s}[f_s(v) \neq G(v)] = \mathbb{P}_{v \subset s}[f_s(v) = G(v)] = \mathbb{E}_{v \sim X(j)}\left[\max_{\sigma \in \Sigma} \left\{ \mathbb{P}\left[S_\sigma^{(v)} \right] \right\} \right],$$

where the inner probability is over the v-conditioned marginal.

The idea, which is fairly standard (see e.g. [DK17, Claim 5.2]), is now to relate this maximum partition size to the *local* disagreement of $\mathcal{F}$ using expansion. In particular, for each $v \in X(j)$ define the v-local disagreement of $\mathcal{F}$ as:

$$\varepsilon_v := \mathbb{P}_{(s_1, s_2) \sim D_{d,k}^{(v)}}[f_{s_1}(v) \neq f_{s_2}(v)].$$

It is easy to see that $\mathbb{E}_v[\varepsilon_v] \leqslant \varepsilon + \eta$. This follows from observing that sampling a random $v \in X(j)$, then $(s_1, t, s_2) \sim D_{d,k}^{(v)}$ is equivalent to *first* sampling $(s_1, t, s_2) \sim D_{d,k}$, then sampling v uniformly from t. The

bound is clear from the latter interpretation since the tester rejects with probability at most ε, and otherwise agrees on all but an η fraction of the j-sets in t. It is therefore enough to prove

$$\max_{\sigma \in \Sigma} \left\{ \mathbb{P}\left[S_\sigma^{(v)}\right] \right\} \geqslant 1 - \frac{\varepsilon_v}{1-\nu},$$

as taking expectation on both sides gives the result.

Toward this end, define $E = E^{(v)}$ to be the set of edges $(s_1, s_2) \sim D_{d,k}^{(v)}$ that cross the alphabet partition. Since the marginal is distributed as the down-up walk in v's link, this graph is a ν-spectral expander by assumption. Writing $S_\sigma = S_\sigma^{(v)}$ for simplicity of notation, we have

$$\begin{aligned}
\varepsilon_v = \mathbb{P}\left[E\right] &= \sum_{\sigma \in \Sigma} \mathbb{P}_{(s_1, s_2)} \left[s_1 \in S_\sigma, s_2 \notin S_\sigma\right] \\
&= \sum_{\sigma \in \Sigma} \langle \mathbf{1}_{S_\sigma}, D_{d,k}^{(v)}(\mathbf{1} - \mathbf{1}_{S_\sigma}) \rangle \\
&= 1 - \sum_{\sigma \in \Sigma} \langle \mathbf{1}_{S_\sigma}, D_{d,k}^{(v)} \mathbf{1}_{S_\sigma} \rangle
\end{aligned} \tag{8.4}$$

where $\mathbf{1}_{S_\sigma}$ is the indicator of S_σ and $\mathbf{1}$ is the all-ones function and the final equality is from by stochasticity of $D_{d,k}^{(v)}$. Let $f_\sigma = \mathbf{1}_{S_\sigma} - \mathbb{P}\left[S_\sigma\right] \mathbf{1}$ be the projection of $\mathbf{1}_{S_\sigma}$ onto $\mathbf{1}^\perp$. By orthogonality it holds that $\|f_\sigma\|^2 = \|\mathbf{1}_{S_\sigma}\|^2 - \mathbb{P}\left[S_\sigma\right]^2 \|\mathbf{1}\|^2 = \mathbb{P}\left[S_\sigma\right] - \mathbb{P}\left[S_\sigma\right]^2$. Therefore exploiting the expansion of $D_{d,k}^{(v)}$ we have

$$\begin{aligned}
\langle \mathbf{1}_{S_\sigma}, D_{d,k}^{(v)} \mathbf{1}_{S_\sigma} \rangle &= \mathbb{P}\left[S_\sigma\right]^2 \langle \mathbf{1}, \mathbf{1} \rangle + \langle f_\sigma, D_{d,k}^{(v)} f_\sigma \rangle \\
&\leqslant \mathbb{P}\left[S_\sigma\right]^2 + \nu \|f_\sigma\|^2 \\
&= \nu \mathbb{P}\left[S_\sigma\right] + (1-\nu) \mathbb{P}\left[S_\sigma\right]^2.
\end{aligned}$$

Inserting this back to (8.4) we have that

$$\begin{aligned}
\varepsilon_v &\geqslant 1 - \sum_{\sigma \in \Sigma} \left(\nu \mathbb{P}\left[S_\sigma\right] + (1-\nu) \mathbb{P}\left[S_\sigma\right]^2 \right) \\
&= (1-\nu) \left(1 - \sum_{\sigma \in \Sigma} \mathbb{P}\left[S_\sigma\right]^2 \right) \\
&\geqslant (1-\nu) \left(1 - \max_{\sigma \in \Sigma} \left\{ \mathbb{P}[S_\sigma] \right\} \right).
\end{aligned}$$

Re-arranging gives the desired result. $\qquad \square$

8.2.1 Testing all the intersection

We note that one could also consider the standard V-test where s and s' are still drawn from the down-up walk, but we test agreement over $t = s_1 \cap s_2$ (whereas in the definition of the test s_1 and s_2 may have some elements in their intersection not in t). The success in these tests are related by the following claim.

Claim 8.3. Let $\eta > 0$. Let D_1 be the test above. Let D_2 be the test with a similar distribution to D_1, with the distinction that $t = s_1 \cap s_2$. Then for any $\mathcal{F}$, $Agree_{\eta, D_2}(\mathcal{F}) \geqslant 1 - \varepsilon$ implies that $Agree_{2\eta, D_1}(\mathcal{F}) \geqslant 1 - \varepsilon - 2^{-\Omega_\gamma(\eta^2 k/j)}$.

Proof of Claim 8.3. Let s_1, s_2 be a pair supported by D_2 such that $\mathcal{F}$ passes the D_2-test on the pair. Let A be the set of j-faces contained in $s_1 \cap s_2$ on which f_{s_1}, f_{s_2} disagree on. By assumption $|A| \leqslant \eta|s_1 \cap s_2|$. By the sampling properties of the complete complex Claim A.1 the fraction of $t \subseteq s_1 \cap s_2$ such that $|A \cap t| \geqslant 2\eta|t|$ is $2^{-\Omega_\gamma(\eta^2 k/j)}$. Thus the probability of *failing* the D_1 test is at most the probability of failing the D_2 test plus $2^{-\Omega_\gamma(\eta^2 k/j)}$. $\qquad \square$

Remark 8.4. When $\eta \lesssim \varepsilon$, the 'additive' error $\exp(-\Omega(d))$ can be removed from Theorem 8.1 by interpolating the result with a standard averaging argument over Claim 8.2. This technique stops working as soon as η is much larger than ε. While the additional $\exp(-\Omega(d))$ error is not harmful in most applications, it is nevertheless interesting to ask whether the term can be removed for general η.

8.2.2 Agreement Testing for the non-Lazy Down-Up Walk

Using a similar technique, we can also prove an agreement test for the following distribution based on the swap walks. Define the test distribution $(s_1, s_2, t) \sim A_{d,k}$ by the following procedure:

1. Sample $t \in X(k)$.

2. Sample $r_1, r_2 \in X(d-k)$ in the swap-walk inside X_t.

3. Output $s_i = t \cup r_i$ and t.

The resulting edge distribution over (s_1, s_2) is sometimes called the 'partial' swap walk [AJT19]. Note this walk has the property that the intersection is fixed. The following claim shows that when the local swap walks expand, an agreement theorem for this walk follows as well.

Claim 8.5. Let j, d be integers, let $\gamma, \eta, c, \lambda, \nu < 1$. Let $k = \gamma d$ and suppose $j < k$, $\eta \lfloor \frac{k}{j} \rfloor \geqslant 192 \log \frac{16}{\eta}$ and $\nu < 1$. Let X be a $(2d-k)$-uniform c-locally nice complex such that for every $v \in X(j)$, $\lambda\big(A_{d-j,k-j}(X_v)\big) \leqslant \nu$. Suppose additionally that for every $t \in X(k)$, $S_{d-k,d-k}(X_t) \leqslant \lambda$. Then the following holds for the set system $SS(X, \Sigma, j, k, d)$; for every alphabet Σ, family $\mathcal{F}$, and $\varepsilon > 0$

$$Agree_\eta(\mathcal{F}) \geqslant 1 - \varepsilon \quad \implies \quad \mathrm{dist}_{8\frac{\eta+\varepsilon}{1-\nu}}(\mathcal{F}, g_{\mathrm{maj}}) \leqslant \varepsilon^{O_{\gamma,C}(1)} + \lambda^{O_{\gamma,C}(1)}. \tag{8.5}$$

where $g_{\mathrm{maj}} := \underset{s \in X(d):s \supset v}{\mathrm{plurality}} \{f_s(v)\}$ is the plurality decoding.

Proof of Claim 8.5. As in the proof of Theorem 8.1, we define

$$B := \left\{ s \in S \ \middle| \ \underset{v \in s}{\mathbb{P}}\left[f_s(v) \neq G(v)\right] > \frac{8(\varepsilon+\eta)}{1-\nu} \right\}.$$

and

$$A := \left\{ s \in S \ \middle| \ \underset{v \in s}{\mathbb{P}}\left[f_s(v) \neq G(v)\right] \leqslant \frac{4(\varepsilon+\eta)}{1-\nu} \right\}.$$

As in the proof of Theorem 8.1, $\mathbb{P}[A] \geqslant \frac{1}{2}$ by an argument similar to Claim 8.2 and Markov's inequality. More precisely, we need to repeat the proof of Claim 8.2, only instead of using the down-up walk conditioned on v, we use $A_{d,k}$ conditioned on $v \in t$. This is just $A_{d-j,k-j}$ inside X_v.

In addition, as in the proof of Theorem 8.1, it also holds that the relative fraction of edges crossing between A and B is small, i.e. $\mathbb{P}[E_{A,B}] \leqslant 2\varepsilon$.

Our goal is to show that $\mathbb{P}[B] \leqslant \max\{\varepsilon^{\Omega_{\gamma,c}(1)}, \exp(-\Omega_{\gamma,c}(d)), \lambda^{\Omega_{\gamma,c}(1)}\}$.

Assume otherwise, and that in particular reverse hypercontractivity for indicators applies for sets of size $\mathbb{P}[B]$. By Theorem 6.4

$$\underset{t \in X(k)}{\mathbb{P}}\left[\underset{s \supseteq t}{\mathbb{P}}[A] \cdot \underset{s \supseteq t}{\mathbb{P}}[B] \geqslant (\mathbb{P}[A]\,\mathbb{P}[B])^{O_{\gamma,c}(1)} \right] \geqslant \mathbb{P}[A]\,\mathbb{P}[B]^{O_{\gamma,c}(1)}.$$

Let us denote by G the set of t satisfying the above condition. Suppose we sampled $t \in G$ in the first step of our agreement distribution. Recall that now we sample and edge by going in the swap walk in the link of t. The probability we sampled an edge between A and B is therefore

$$\underset{r_1, r_2 \sim S(X_t)}{\mathbb{P}}[t \cup r_1 \in A, t \cup r_2 \in B] \geqslant (\mathbb{P}[A]\,\mathbb{P}[B])^{O_{\gamma,c}(1)} - \lambda = p \cdot \mathbb{P}[B]^{O_{\gamma,c}(1)} - \lambda$$

by the expansion of the swap walk and the expander mixing lemma (for some constant p). Thus if $p\,\mathbb{P}[B]^{O_{\gamma,c}(1)} \geqslant 2\lambda$ then we have that

$$\Omega(\mathbb{P}[B]^{O_{c,\gamma}(1)}) \leqslant \mathbb{P}[G]\,\mathbb{P}[s_1 \in A, s_2 \in B \mid G] \leqslant \mathbb{P}[E_{A,B}] \leqslant 4\varepsilon$$

and the claim follows. $\qquad\square$

8.3 Low Soundness and the Z-Test

In this section we will apply Theorem 8.1 and reverse hypercontractivity to prove soundness of [IKW12]'s Z-test in the 1%-regime under certain stronger assumptions on the complex. Before moving on, we briefly comment on the notation in this subsection. Unlike before, we will use the convention that s, t are $\frac{k}{2}$-sets (and not k-sets as in the previous subsections). We will typically denote k-sets by a capital letter instead. We first recall their test:

The Z-test on $X(k)$:

1. Draw $t \in X\left(\frac{k}{2}\right)$, and $s, s' \in X_t\left(\frac{k}{2}\right)$ independently

2. Draw $s'' \sim X_{s'}\left(\frac{k}{2}\right)$

3. Accept if:
$$f_{t \cup s} = f_{t \cup s'} \quad \text{and} \quad f_{s' \cup t} = f_{s' \cup s''}$$

We prove soundness for complexes which are 'λ-global', a slight strengthening of ℓ_∞-independence [KKS21]:

Definition 8.6 (Global Complex). A k-uniform complex X is called λ-global if $\forall t \in X\left(\frac{k}{2}\right)$:

$$\|S_t - \pi_{\frac{k}{2}}\|_{TV} \leqslant \lambda$$

where S_t is the distribution over neighbors of t in $S_{\frac{k}{2}, \frac{k}{2}}$ and $\pi_{\frac{k}{2}}$ is the induced distribution on $X\left(\frac{k}{2}\right)$.

Globality of X essentially promises that every $\frac{k}{2}$-set in X 'sees' most other $\frac{k}{2}$-sets in X, and is equivalent to bounding the matrix infinity norm of the swap walk from its stationary operator $\frac{1}{2}\|S_{\frac{k}{2}, \frac{k}{2}} - \Pi_{\frac{k}{2}}\|_\infty \leqslant \lambda$.

Theorem 8.7. $\forall \lambda, \eta > 0$ and large enough k, let X be a k-uniform, λ-global, c-locally nice complex. Then for any $\delta \in \left(8\lambda + e^{-\Omega_c(\eta k)}, \frac{1}{8}\right)$ if $Agree_0^Z(\mathcal{F}) \geqslant \delta$:

$$\exists g : X(0) \to \Sigma, \quad \mathop{\mathbb{P}}_{s \in X(k)}[f_s \overset{\eta}{\approx} g(s)] \geqslant \delta/8.$$

A couple of remarks are in order. First, we mention that it is well known that no tester can do better than $\exp(-\Omega(k))$ soundness [DG08b; DL17], so our bound is essentially optimal for small λ. Second, we remark that many (dense) complexes of interest are indeed λ-global HDX, including basic examples such as the complete complex and random complexes, and more involved examples such as skeletons of various Ising models, independent sets, list-colorings, or more generally essentially any of the myriad complexes studied in the approximate sampling literature (see e.g. [ALG21; Liu21; CLV21; Fen+22; Bla+22] among many others). Theorem 8.7 shows the Z-test is sound on these families in a blackbox fashion, even up to $\exp(-\eta k)$ for the appropriate parameters. We elaborate on these examples further in the end of this section.

The proof of Theorem 8.7 largely follows the strategy of [IKW12; DL17]. The main difference lies in replacing certain ad-hoc arguments for the complete complex with more general methods based on reverse hypercontractivity and globality.

8.3.1 A Local Agreement Theorem

The core of Theorem 8.7 is a *local* structure theorem for the V-test on high dimensional expanders. Stating this requires a few definitions. We remark that these are standard notions in the literature, and we follow the notation of [DL17]. We first define *restrictions*.

Definition 8.8 (Restriction). A *restriction* is a pair $\tau = (t, \sigma)$ such that $t \in X\left(\frac{k}{2}\right)$ and $\sigma : t \to \Sigma$ is an assignment to t.

Given a family of functions as above, there is a natural distribution over restrictions that comes from sampling a random function and a random restriction of it. More formally:

1. Sample $t \in X\left(\frac{k}{2}\right)$ and $s \in X_t\left(\frac{k}{2}\right)$.

2. Output $\tau = (t, f_{s \cup t}|_t)$.

Whenever we mention a distribution over restrictions τ we always mean $\tau = (t, f_{s \cup t}|_t)$ sampled as above. Given such a restriction τ, it will also be useful to have notation for the faces in X that are *consistent* with τ.

Definition 8.9 (Consistent Strings). Given a restriction $\tau = (t, \sigma)$, denote by $\mathcal{V}_\tau$ the set of faces $s \in X_t\left(\frac{k}{2}\right)$ consistent with τ:

$$\mathcal{V}_\tau := \left\{ s \in X_t\left(\frac{k}{2}\right) : f_{s \cup t}|_t = \sigma \right\}$$

A restriction is called *good* if it has many consistent faces.

Definition 8.10 (Good Restriction). A restriction $\tau = (t, \sigma)$ is called δ-good if $\mathbb{P}_{s \in X_t}[\mathcal{V}_\tau] \geqslant \frac{\delta}{2}$.

When δ is clear from context, we just call such restrictions good. Finally, a restriction is called 'DP' if its consistent strings agree with a global function ('DP' is for 'direct product').

Definition 8.11 (DP-Restriction). A restriction $\tau = (t, \sigma)$ is called (η, δ)-DP if there exists $g_\tau : X_t(1) \to \Sigma$ such that:

$$\mathbb{P}_{s \sim X_t}\left[f_{t \cup s} \overset{\eta/4}{\not\approx} g_\tau \; \middle| \; s \in \mathcal{V}_\tau \right] \leqslant \delta^2$$

As before, when η, δ are clear from context we just write 'DP'. We can now state the local agreement theorem for high dimensional expanders.

Theorem 8.12 (The Local Agreement Theorem). *For any $\eta > 0$ and k sufficiently large, let X be a k-uniform c-locally nice complex. Then for all $\delta > e^{-\Omega(\eta k)}$, if $\mathrm{Agree}_0^V(\mathcal{F}) \geqslant \delta$:*

1. *Many restrictions are good:*

$$\mathbb{P}_\tau[\tau \text{ is good}] \geqslant \frac{\delta}{2}$$

2. *Almost all 'good' restrictions are DP:*

$$\mathbb{P}_\tau[\tau \text{ is DP} \mid \tau \text{ is good}] \geqslant 1 - \delta^2$$

We remark that while Theorem 8.12 does not give a true agreement tester in its own right, it is a powerful tool independent of Theorem 8.7. It roughly states that good agreement implies a non-negligible fraction of good restrictions, *and* that if f_s agrees with a restriction of a face t it is almost certainly a restriction of a global assignment on the link X_t. This tool is a critical part of many of the previous works on agreement testing [DG08b; DS14; IKW12; DL17].

We prove Theorem 8.12 in the next section. We first show it implies the main theorem under the additional assumption of λ-globality. The argument is a simple adaptation of [IKW12].

Proof of Theorem 8.7. As in [IKW12], we replace the standard Z-test in Figure 8.3 with a related proxy test where we first sample (t, s) as before, but then draw (s', s'') *independent of* (t, s). If $s' \in X_t$ we continue as in the Z-test, and if $s' \notin X_t$ we accept. The probability of passing the new test is only greater or equal to the probability of passing the original test, but by λ-globalness, the swap walk starting at t is λ-close to the stationary distribution hence the probability that s' misses the link of t is at most λ.

With this in mind, let τ denote the restriction $(t, f_{s \cup t}|_t)$ and observe that if τ is not good, the new Z-test accepts with probability at most $\delta/2 + \lambda \leqslant \delta$. Thus it must be the case that conditioned on τ being good, the test still accepts with probability at least δ. Condition on this event, and observe that by Theorem 8.12, τ is also a DP-restriction except with probability at most $\delta^2 \leqslant \delta/8$ (recall that $\delta < \frac{1}{8}$).

Fix such a DP restriction τ, and denote by B the set of faces in $\mathcal{V}_\tau$ on which $\mathcal{F}$ is $\eta/4$-close to the DP function g_τ promised by Theorem 8.12 (if there is more than one, take any such function). Assume toward contradiction that $\mathcal{F}$ is not $(\eta, \delta/8)$-close to g_τ. Let $S = s' \cup s''$ and let H denote the event that $f_S \overset{\eta}{\approx} g_\tau$.[11] By assumption $\mathbb{P}\left[H\right] \leqslant \frac{\delta}{8}$, because S is drawn as in the distribution of $X(k)$. We define four events, outside which the test rejects:

1. E_1: The event that $s' \notin X_t$,

2. E_2: The event that $s' \in \mathcal{V}_\tau \setminus B$,

3. E_3: The event that $S \in H$,

4. E_4: The event that $S \notin H$, and $f_S|_{s'} \overset{\eta/2}{\approx} g_\tau|_{s'}$.

We first argue the test rejects if none of the E_i occurs. In particular, note by $\overline{E_1} \cap \overline{E_2}$, s' must either be in B, or in $X_t \setminus \mathcal{V}_\tau$. In the latter case, the first query of the proxy Z-test rejects since $s' \in X_t$ but $f_{t \cup s'} \neq f_{t \cup s}$. On the other hand, in the former case $s' \in B$ and therefore $f_{t \cup s'}|_{s'} \overset{\eta/4}{\approx} g_\tau|_{s'}$. However, by $\overline{E_3} \cap \overline{E_4}$, we have $S \notin H$ and, moreover, that the restriction to s' has disagreement at least $f_S|_{s'} \overset{\frac{\eta}{2}}{\not\approx} g_\tau|_{s'}$. Thus there must be a vertex on which $f_{t \cup s'}$ and f_S disagree and the second query rejects.

Finally by a union bound it is enough to prove $\sum \mathbb{P}[E_i] \leqslant \delta/2$. The first event occurs with probability at most $\delta/8$ by λ-globality where $\lambda < \frac{\delta}{8}$. The second event occurs with probability at most $\delta^2 \leqslant \delta/8$ by Theorem 8.12. The third event occurs with probability at most $\delta/8$ by assumption. The final event occurs with probability at most $e^{-\Omega(\eta k)} \leqslant \delta/8$ by Chernoff, since f_S η-disagrees with g_τ and s' is a random subset, which completes the proof. $\qquad \square$

8.3.2 Proof of the Local Agreement Theorem

Let us begin by proving the first property in Theorem 8.12.

Claim 8.13. Under the assumptions of Theorem 8.12, $\mathbb{P}_\tau\left[\tau \text{ is good}\right] \geqslant \frac{\delta}{2}$.

Proof. Let $p = \mathbb{P}_\tau\left[\tau \text{ is good}\right]$. Since $Agree_0^V(\mathcal{F}) \geqslant \delta$ and we have

$$
\begin{aligned}
\delta &\leqslant Agree_0^V(\mathcal{F}) \\
&= p \mathop{\mathbb{P}}_{s,t,s'}\left[f_{t \cup s} = f_{t \cup s'} \mid (t, f|_{s \cup t}) \text{ is good}\right] + (1-p) \mathop{\mathbb{P}}_{s,t,s'}\left[f_{t \cup s} = f_{t \cup s'} \mid (t, f|_{s \cup t}) \text{ is not good}\right].
\end{aligned}
$$

By definition $\mathbb{P}_{s,t,s'}\left[f_{t \cup s} = f_{t \cup s'} \mid (t, f|_{s \cup t}) \text{ is not good}\right] \leqslant \frac{\delta}{2}$ so $\delta \leqslant p + \frac{\delta}{2}$ which implies the claim. $\qquad \square$

The rest of this section is dedicated to proving the second property in Theorem 8.12, which is more complicated. The proof has two main steps. First, we will need a variant of the notion of an *excellent restriction*, that appeared in [DG08b; IKW12; DL17]. Roughly, an excellent restriction is a restriction (t, σ) so that if $s_1, s_2 \in \mathcal{V}_t$, the probability that $f_{t \cup s_1}|_{s_1 \cap s_2} \approx f_{t \cup s_2}|_{s_1 \cap s_2}$ is $1 - \text{poly}(\delta)$. The probability in which we choose $s_1, s_2 \in X_t$ is stated below. We show most good restrictions are in fact excellent.

Second, we show that excellent restrictions are DP. Here we use a variant of the "smoothing" technique introduced by [DL17] in order to get an ensemble of local functions $\left\{\tilde{f}_s : s \to \Sigma \;\middle|\; s \in X_t\left(\frac{k}{2}\right)\right\}$ that agree with high probability $1 - \text{poly}(\delta)$ and use our 99%-tester (Theorem 8.1) to show the restriction is DP.

Formally we require some additional notation to properly define these notions. First, denote the $\frac{k}{4}$-step down-up walk in the link of t as N_t and denote by $(s, t', s') \sim N_t$ the distribution where s, s' are chosen according to N_t and t' is the intermediate set chosen in the first down step (we note that $t' \subseteq s \cap s'$ but equality doesn't necessarily hold). Second, let N_t^3 denote three steps of N_t. Let $(s_1, t_1, s_2, t_2, s_3, t_3, s_4) \sim N_t^3$ be the full walk sampled. Finally, let $(s_1, t_2, t', s_4) \sim N_t^3$ be such that s_1, t_2, s_4 are the marginals as before, and $t' = t_1 \cap t_2 \cap t_3$ the vertices that are in all sets in the sampled walk.

Fix some sufficiently small constant $c_1 > 0$. We can now define excellence.

---

[11]Formally here we extend g_τ to a function on $X(1)$ by arbitrarily assigning values to any vertex in $X(1) \setminus X_t(1)$.

Definition 8.14 (Excellent Restriction). A good restriction $\tau = (t, \sigma)$ is called excellent if:

1. $\mathbb{P}_{(s,t',s')\sim N_t}\left[(s, s' \in \mathcal{V}_\tau) \wedge \left(f_s|_{t'} \overset{\eta/1500}{\not\approx} f_{s'}|_{t'}\right)\right] \leqslant e^{-c_1\eta k}$ and,

2. $\mathbb{P}_{(s_1,t_2,t',s_4)\sim N_t^3}\left[(s, s' \in \mathcal{V}_\tau) \wedge \left(f_{s_1}|_{t'} \overset{\eta/1500}{\not\approx} f_{s_4}|_{t'}\right)\right] \leqslant e^{-c_1\eta k}$.

where we've abused notation and written f_s as shorthand for $f_{s\cup t}$.

We note that this definition is slightly cumbersome due to the fact that we don't measure the agreement of $f_s, f_{s'}$ (respectively f_{s_1}, f_{s_4}) on the intersection of $s \cap s'$ (respectively $s_1 \cap s_4$), but instead on a fixed subset. This is for technical reasons and we encourage the readers to think of the case of N_t where $t' = s \cap s'$ (and analogously for N_t^3).

For notational convenience, let $\mu := e^{-c_1\eta k}$ and $\eta' = \eta/1500$. The following lemma (or variants thereof) is fairly standard [Imp+08; IKW12; DL17]:

Lemma 8.15 (Good restrictions are excellent). *There exists $c_2 > 0$ such that*
$$\mathbb{P}_{\tau=(t,f_{s\cup t}|_t)}\left[\tau \text{ is good but not excellent}\right] \leqslant e^{-c_2\eta k}.$$

Proof. We only prove that N_t^3 case, which is somewhat less standard. The N_t proof is analogous. It is sufficient to show that over a random restriction $\tau = (t, f_{s_0\cup t}|_t)$:

$$\mathbb{P}_{\tau=(t,f_{s_0\cup t}|_t)}\left[\mathbb{P}_{(s_1,t_2,t',s_4)\sim N_t^3}\left[s_1, s_4 \in \mathcal{V}_t \wedge f_{s_1}|_{t'} \overset{\eta'}{\not\approx} f_{s_4}|_{t'}\right] > e^{-c\eta k}\right] \leqslant e^{-\Omega(\eta k)} \quad (8.6)$$

where s_1, t_2, t', s_4 is chosen independently of s_0. In particular if this holds then, by claim Claim 8.13, conditioning on being good can increase the probability of this event to hold to at most $\frac{2e^{-\Omega(\eta k)}}{\delta} \leqslant e^{-c_2\eta k}$ for the right choice of constants.

The key to (8.6) is to observe that this process can be sampled by the following equivalent method: sample $T \in X\left(\frac{k}{2} + \frac{k}{4}\right)$, and randomly partition it into $T = t \cup t_2$ so that $|t| = \frac{k}{2}$ and $|t_2| = \frac{k}{4}$. The faces s_1, t_2, t_3, and s_4 are now drawn independently from the "up-down-up" walk $U_{\frac{k}{4},\frac{k}{2}}D_{\frac{k}{2},\frac{k}{4}}U_{\frac{k}{4},\frac{k}{2}}$ within X_t starting from t_2.

Observe that $|t'| \geqslant \frac{k}{32}$ with probability $e^{-\Omega(k)}$ by a standard Chernoff bound. Conditioning on this event, observe that if $s_1, s_4 \in \mathcal{V}_\tau$ and $f_{s_1}|_{t'} \overset{\eta'}{\not\approx} f_{s_4}|_{t'}$, then it must be the case that

1. t_2 has at least $\frac{\eta'k}{32}$ elements on which $f_{s_1} \neq f_{s_4}$

2. t has no elements on which $f_{s_1}|_t \neq f_{s_4}|_t$.

However, t and t_2 are a uniformly random partition of T, so such a split occurs with probability at most $e^{-\Omega(\eta k)}$ by Chernoff as desired. $\square$

To proving excellent restrictions are DP, we'll use a variant of Dinur and Livni-Navon's 'smoothing' operation that spreads the consistent strings $\mathcal{V}_\tau$ over the entire link. Instead of using the noise operator as in their work, we use the down-up walk which avoids a number of technical complications.

Definition 8.16 (Smoothed Assignment). For every $\tau = (t, f_{t\cup s_0}|_t)$, define the smoothed assignment $\widetilde{\mathcal{F}}_\tau = \{\tilde{f}_s\}_{s\in X_t}$ as
$$\tilde{f}_s(v) = \underset{(t',s')\sim N_t(s):s'\in\mathcal{V}_\tau,t'\ni v}{\text{Plurality}} \{f_{s'}(v)\}.$$

We break ties arbitrarily. If the plurality is not well defined (i.e. there are no such $s' \in \mathcal{V}_\tau$ that contain v) we define it to be $\tilde{f}_s(v) = \perp$.

The main idea is to show that on excellent restrictions, $\widetilde{\mathcal{F}}_\tau$ is 1) highly consistent with the original family $\mathcal{F}$ on $\mathcal{V}_\tau$, and 2) has very high agreement on the *entire* link. This reduces the problem to an (approximate) 99%-regime test within the link of t, which we can solve via our tester from the previous section. More formally, the following lemmata suffice to prove Theorem 8.12.

Lemma 8.17. *Every excellent $\tau = (t, \sigma)$ satisfies:*

$$\mathop{\mathbb{P}}_{s \sim X_t}\left[f_s \overset{10\eta'}{\not\approx} \tilde{f}_s \ \middle|\ s \in \mathcal{V}_\tau \right] \leqslant \frac{\delta^2}{2},$$

that is, $\mathcal{F}$ and $\widetilde{\mathcal{F}}$ are close on $\mathcal{V}_\tau$ (note $f_s \overset{10\eta'}{\not\approx} \tilde{f}_s$ is over just the vertices of s, not $s \cup t$)

Lemma 8.18. *There exists $c_3 > 0$ such that every excellent $\tau = (t, \sigma)$ satisfies*

$$\mathop{\mathbb{P}}_{(s, t', s') \sim N_t}\left[\tilde{f}_s|_{t'} \overset{10\eta'}{\not\approx} \tilde{f}_{s'}|_{t'} \right] \leqslant e^{-c_3 \eta k},$$

that is, $\widetilde{\mathcal{F}}$ is highly agreeing on X_t.

Proof of Theorem 8.12. The first property is proved in Claim 8.13. Toward the second, we will prove that every excellent τ is a DP-restriction. Then by Lemma 8.15 we have that

$$\mathop{\mathbb{P}}_\tau\left[\tau \text{ is DP} \mid \tau \text{ is good} \right] \geqslant \mathop{\mathbb{P}}_\tau\left[\tau \text{ is excellent} \mid \tau \text{ is good} \right] \geqslant 1 - \delta^2$$

where we've used the assumption that $\delta \geqslant e^{-\Omega(\eta k)}$. Fix any excellent $\tau = (t, \sigma)$. One can check these satisfy the conditions of Theorem 8.1 with $\nu = 1/2$. Together with Lemma 8.18 and Claim 8.3 that translates between the agreement guarantee in Lemma 8.18 to that of Theorem 8.1, we have that there exists a global function $g : X_t(0) \to \Sigma$ such that

$$\mathop{\mathbb{P}}_{s \sim X_t}[\tilde{f}_s \overset{320\eta'}{\not\approx} g|_s] \leqslant e^{-c_4 \eta k}$$

where c_4 is some constant depending on c_3, η and the c-local niceness of X. We note that after conditioning over $s \in \mathcal{V}_\tau$, we still have that this is at most $\frac{2e^{-c_4 \eta k}}{\delta} \leqslant \frac{\delta^2}{2}$, where this holds by assumption that δ is large enough.

Combining this with Lemma 8.17, we then have

$$\mathop{\mathbb{P}}_{s \sim X_t}\left[f_s \overset{\eta/4}{\not\approx} g(s) \ \middle|\ s \in \mathcal{V}_t \right] \leqslant \mathop{\mathbb{P}}_{s \sim X_t}\left[f_s \overset{10\eta'}{\not\approx} \tilde{f}_s \ \middle|\ s \in \mathcal{V}_\tau \right] + \mathop{\mathbb{P}}_{s \sim X_t}\left[\tilde{f}_s \overset{320\eta'}{\not\approx} g(s) \ \middle|\ s \in \mathcal{V}_t \right]$$
$$\leqslant \delta^2$$

as desired. □

It is left to prove the key lemmata. This is the main place we use reverse hypercontractivity (other than the application of Theorem 8.1 in the proof of Theorem 8.12). The proofs for both properties follow the strategy of [DL17] adapted to our setting.

Proof of Lemma 8.17. We define two families of 'bad' sets outside of which f and $\tilde{f}$ approximately agree. First, we look at the set of 'lonely' strings that don't sufficiently see $\mathcal{V}_\tau$.

$$L := \left\{ s \in X_t : \mathop{\mathbb{P}}_{s' \sim N_t(s)}[s' \in \mathcal{V}_\tau] \leqslant \delta' \right\}$$

for $\delta' = e^{-c_5 \eta k}$ for some $c_5 > 0$ sufficiently small. We will assume below that $\delta' \leqslant \delta^4$ (which is possible due to the assumptions on δ). We record the following properties of L that we prove at the end of the subsection.

Claim 8.19. There exists a sufficiently small constant $c_5 > 0$ such that the following holds for $\delta' = e^{-c_5 \eta k}$.

1. $\mathbb{P}_{s \in X_t}[L] \leqslant \frac{\delta'^2}{2}$.

2. For any $s \notin L$, the fraction of vertices $v \in s$ s.t. $\mathbb{P}_{(t',s') \sim N_t(s)}[s' \in \mathcal{V}_\tau \mid t' \ni v] \leqslant \delta'/3$ is at most η'.

Second, we look at the family of sets within $\mathcal{V}_\tau$ with strong disagreement:

$$B := \left\{ s \in \mathcal{V}_\tau : \mathbb{P}_{(t',s') \sim N_t(s)}\left[s' \in \mathcal{V}_\tau \wedge f_s|_{t'} \overset{\eta'}{\not\approx} f_{s'}|_{t'} \right] \geqslant \frac{\delta'}{30} \right\}.$$

We claim that both L and B have measure at most $\delta^2/2$ within $\mathcal{V}_\tau$. The former follows from the first item of Claim 8.19 since $\mathbb{P}[L|\mathcal{V}_\tau] \leqslant \frac{\delta'^2}{\delta} \leqslant \frac{\delta^2}{2}$. For the latter we have $\mathbb{P}[B] \leqslant \frac{30\mu}{\delta'} \leqslant \frac{\delta^2}{2}$ from Markov's inequality and excellence of τ (here we also need that the c_5 in the definition of δ' is sufficiently smaller than the c_1 in the definition of μ).

It is left to observe that $f_s \overset{10\eta'}{\approx} \tilde{f}_s$ for any face $s \in \mathcal{V}_\tau \setminus (L \cup B)$. Fix such an s, and define D_0 to be the set of vertices on which f and $\tilde{f}$ disagree:

$$D_0 := \left\{ v \in s : f_s(v) \neq \tilde{f}_s(v) \right\}.$$

We need to argue $|D_0| \leqslant 10\eta'|s|$. Assume toward contradiction otherwise. Let D' denote the vertices in s such that $\mathbb{P}_{(t',s') \sim N_t(s)}[s' \in \mathcal{V}_\tau \mid t' \ni v] \leqslant \delta'/3$ and write $D = D_0 \setminus D'$. By Claim 8.19, we have $|D| \geqslant 9\eta'|s|$.

Consider now the following probabilistic experiment. Sample $v \in D$ uniformly, then some $t', s' \sim N_t(s)$ conditioned on $v \in t'$ and $s' \in \mathcal{V}_\tau$. We denote this distribution by $\mathcal{D}$. Note that for every fixed $v \in D$, this is exactly the distribution over which the plurality in the definition of $\tilde{f}_s(v)$ was defined.

With this in mind consider the event $P = \{f_s(v) \neq f_{s'}(v)\}$ and observe that $\mathbb{P}_{\mathcal{D}}[P] \geqslant \frac{1}{2}$ because for every fixed $v \in D$, by construction $f_s(v)$ disagrees with the plurality vote on v. On the other hand, we show toward contradiction that P is contained in events whose total probability is strictly less than $\frac{1}{2}$. Namely:

1. E_1: the event $\{f_s|_{t'} \overset{\eta'}{\not\approx} f_{s'}|_{t'}\}$,

2. E_2: the event $|t' \cap D| < 6\eta'|t'|$,

3. E_3: the event $P \setminus (E_1 \cup E_2)$.

Obviously $P \subseteq E_1 \cup E_2 \cup E_3$ so it is enough to bound the probability of the three events, starting with E_3. In this case we have that $f_s, f_{s'}$ disagree on at most an η'-fraction of t', but $|t' \cap D| \geqslant 6\eta'|t'|$. Conditioned on t', v is drawn uniformly from $|D \cap t'|$, so the probability that $f_s(v) \neq f_{s'}(v)$ is at most $\frac{1}{6}$.

Moving on to E_2, notice that if D is an $9\eta'$-fraction of s, a random $t' \subset s$ contains *less than a* $6\eta'$-fraction of D with probability $\exp(-\Omega(\eta'^2 k))$ by Chernoff. While the distribution $\mathcal{D}$ doesn't sample t' uniformly (since it conditions on $v \in t'$ and $s' \in \mathcal{V}_\tau$), we do know for any fixed $v \in D$, $v \in t'$ with probability at least $\frac{1}{2}$, and because $s \notin L$ and $v \notin D'$ the probability that $s' \in \mathcal{V}_\tau$ is at least $\frac{\delta'}{3}$ even after conditioning on v. Hence, even after conditioning on these events the probability that $|t' \cap D| < 6\eta'|t'|$-fraction of D is at most $2\delta^{-1}\exp(-\Omega(\eta'^2 k)) < \frac{1}{10}$ for large enough k.

Finally, we bound the probability of E_1 by similar reasoning.

$$\mathbb{P}_{\mathcal{D}}[E_1] = \mathbb{P}_{v \sim D, (t',s') \sim N_t(s)}\left[f_s|_{t'} \overset{\eta'}{\not\approx} f_{s'}|_{t'} \;\middle|\; s' \in \mathcal{V}_\tau, v \in t' \right]$$

$$\leqslant \frac{3}{\delta'} \mathbb{P}_{v \sim D, s' \sim N_t(s)}\left[s' \in \mathcal{V}_\tau \wedge f_s|_{t'} \overset{\eta'}{\not\approx} f_{s'}|_{t'} \;\middle|\; v \in t' \right]$$

$$\overset{s \notin B}{\leqslant} \frac{3}{\delta'} \cdot \frac{\delta'}{30} \cdot \mathbb{P}[v \in t']^{-1} \leqslant \frac{1}{5}.$$

Thus $\frac{1}{2} \leqslant \mathbb{P}[P] \leqslant \mathbb{P}[E_1] + \mathbb{P}[E_2] + \mathbb{P}[E_3] < \frac{1}{2}$ and a contradiction is reached. $\qquad\square$

Proof of Lemma 8.18. The proof is similar to the first property. Let L be as in Lemma 8.17. This time we define a set of bad *triples* as follows. For a set $(s, t', s') \sim N_t$ we consider the following quantity

$$F(s, t', s') := \mathop{\mathbb{P}}_{(s_1, t_1, s_2, t_2, s_3, t_3, s_4) \sim N_t^3} \left[s_1, s_4 \in \mathcal{V}_\tau \wedge f_{s_1}|_{t''} \overset{\eta'}{\not\approx} f_{s_4}|_{t''} \;\middle|\; s_2 = s, t_2 = t', s_3 = s' \right]$$

where $t'' = t_2 \cap t_3 \cap t_4$. We define B to be the set of triples with large F-value:

$$B := \left\{ (s, t', s') : F(s, t', s') \geqslant \frac{\delta'^2}{80} \right\},$$

and keep L as in the proof of Lemma 8.17. By excellence of τ and Markov's inequality $\mathbb{P}_{(s,t',s') \sim N_t}[B] \leqslant \frac{80\mu}{\delta'^2} \leqslant \exp(-\Omega(\eta k))$. Since s and s' are both marginally distributed as $X_t\left(\frac{k}{2}\right)$, we also have by Claim 8.19 and a union bound that either one of $s \in L$ or $s' \in L$ with probability at most $\delta'^2 = e^{-2c_5 \eta k}$. It is therefore enough to show $\tilde{f}_s|_{t'} \overset{10\eta'}{\approx} \tilde{f}_{s'}|_{t'}$ under the assumption that $(s, t', s') \notin B$ and $s, s' \notin L$.

Toward this end, fix (s, t', s') and consider the set of disagreeing vertices on the smoothings at s and s':

$$D_0 := \left\{ v \in t' : \tilde{f}_s(v) \neq \tilde{f}_{s'}(v) \right\}.$$

Assume for the sake of contradiction that $|D_0| \geqslant 10\eta'|t'|$. Similar to before, we actually consider $D = D_0 \setminus D'$ where D' is the set of vertices in t' such that $\mathbb{P}_{(t_3, s_4) \sim N_t(s')}[s_4 \in \mathcal{V}_\tau | t_3 \ni v]$ (likewise for $s_1, t_1 \sim N_t(s)$). By Claim 8.19, both s and s' have at most $\eta'|s|$ such vertices so $|D| \geqslant 6\eta'|t'|$.

We now consider the following probabilistic experiment. Sample $v \in D$ uniformly and $(s_1, t_1, s_2, t_2, s_3, t_3, s_4)$ conditioned on $s_2 = s, t_2 = t', s_3 = s'$. We then condition this experiment on:

1. $v \in t'' = t_1 \cap t_2 \cap t_3$.

2. $s_1, s_4 \in \mathcal{V}_\tau$.

Note that for a fixed $v \in D$, the marginals $(s_1, t_1, s_2 = s)$ and $(s_3 = s', t_3, s_4)$ are precisely the distributions that were used in the definition of $\tilde{f}_s(v), \tilde{f}_{s'}(v)$ respectively. As before, define the event $P = \{f_{s_1}(v) \neq f_{s_4}(v)\}$ and observe that $\mathbb{P}_\mathcal{D}[P] \geqslant \frac{1}{2}$ because for every $v \in D$ the plurality labeling for v over N_t under our conditioned process is different at s and s'.

We now define events whose union contains P which occur with total probability less than $\frac{1}{2}$. These are

1. E_1: the event $\{f_{s_1}|_{t''} \overset{\eta'}{\not\approx} f_{s_4}|_{t''}\}$

2. E_2: the event $|t'' \cap D| \leqslant 5\eta'|t''|$.

3. E_3: the event $P \setminus (E_1 \cup E_2)$.

By the same analysis as Lemma 8.17 we have $\mathbb{P}[E_3] < \frac{1}{5}$ and $\mathbb{P}[E_2] \leqslant \frac{1}{10}$ (we omit the repeated details). As for E_1 we have

$$\mathbb{P}_\mathcal{D}[E_1] = \mathop{\mathbb{P}}_{v \in D, s_1, t_1, t_3, s_4} \left[f_{s_1}|_{t''} \overset{\eta'}{\not\approx} f_{s_4}|_{t''} \;\middle|\; s_1, s_4 \in \mathcal{V}_\tau, v \in t'' \right]$$

$$\leqslant \frac{9}{\delta'^2} \mathop{\mathbb{P}}_{v \sim D, s' \sim N_t(s)} \left[s_1, s_4 \in \mathcal{V}_\tau \wedge f_{s_1}|_{t''} \overset{\eta'}{\not\approx} f_{s_4}|_{t''} \;\middle|\; v \in t'' \right]$$

$$\overset{(s,t',s') \notin B}{\leqslant} \frac{9}{\delta'^2} \cdot \frac{\delta'^2}{80} \cdot \mathbb{P}\left[v \in t'\right]^{-1} \leqslant \frac{1}{5}.$$

We reach a similar contradiction as before. $\qquad\square$

Proof of Claim 8.19. Let us begin with the first item. Recall under the assumptions of Theorem 8.7 N_t is reverse hypercontractive for indicators. Namely let c_6 be the constant promised by Theorem 6.4, and observe

that by assumption $\mathbb{P}_{X_t}[\mathcal{V}_\tau] \geqslant \delta/2 \geqslant e^{-c_6\eta k}$. If $\mathbb{P}_{X_t}[L] \leqslant 2e^{-2c_5\eta k}$ we are done, so assume otherwise. Then by definition of L and Theorem 6.4 we have the chain of inequalities:

$$\mathbb{P}[L]\delta' \geqslant \mathop{\mathbb{P}}_{(s,s')\sim N_t}[s \in L, s' \in \mathcal{V}_\tau] \geqslant 2^{-q-1}\,\mathbb{P}[L]^q\delta^q.$$

where q is the constant given in Theorem 6.4. Re-arranging gives $\mathbb{P}[L] \leqslant (2^{q+1}\delta'\delta^{-q})^{\frac{1}{q-1}} \leqslant \exp(-c_7\eta k) \leqslant \frac{\delta^2}{2}$ for the appropriate choice of constant $c_7 > 0$. Thus if c_5 is sufficiently small we get the first item.

For the second item, let D' denote the set of $v \in s$ such that $\mathbb{P}_{(t',s')\sim N_t(s)}[s' \in \mathcal{V}_\tau \mid t' \ni v] \leqslant \frac{\delta'}{3}$. Define $g : \binom{s}{|t'|} \to [0,1]$ to be $g(t_0) = \mathbb{P}_{(t',s')\sim N_t(s)}[s' \in \mathcal{V}_\tau \mid t' = t_0]$. Observe that $\mathbb{E}[g] \geqslant \delta'$ since by assumption $s \notin L$. Moreover by construction we have $v \in D'$ if and only if $\mathbb{E}_{t_0 \ni v}[g(t_0)] \leqslant \frac{1}{3}\mathbb{E}[g(t_0)]$.

The problem is now reduced to sampling on the complete complex. In other words let $L = \binom{s}{|t'|}$, $R = s$ and $v \sim t$ if $v \in t$. By Chernoff for every α, this graph is an $(\alpha, \beta, \frac{1}{2})$-multiplicative sampler for $\beta = \exp(-\Omega(\alpha k))$. By Claim 4.8, this implies that the opposite graph is a $(\frac{6\beta}{\alpha}, 2\alpha, \frac{2}{3})$-multiplicative sampler. We take $\alpha = \eta'/2$, and the constant $c_5 > 0$ defined above small enough so that $\frac{6\beta}{\alpha} \leqslant \delta'$. As such we have for every function $g : L \to [0,1]$ of expectation at least δ' the fraction of vertices $v \in s$ such that $\mathbb{E}_{t'\subseteq s, v\in t'}[g] < \frac{1}{3}\mathbb{E}_{t'\subseteq s}[g(t')]$ is at most η' as desired. $\square$

8.4 Examples of Global Complexes

We conclude the section with several examples of 'global HDX' to which Theorem 8.7 applies in a blackbox manner. We start with a few basic direct examples and then argue a vast array of complexes studied in the approximate sampling literature give rise to global HDX. These are:

1. The complete complex.

2. Erdos-Renyi Hypergraphs.

3. Skeletons of the full linear matroid over $\mathbb{F}_q^d$.

4. Skeletons of ℓ_∞-independent complexes, including many classic spin systems.

5. Faces complexes of any of the above.

The first and simplest example of global complexes (other than the complete complex) come from the classical random model of Linial and Meshulam [LM06].

Definition 8.20 (Erdos-Renyi Hypergraphs). The Erdos-Renyi hypergraph $X \sim G_k(n,p)$ is a random k-uniform simplicial complex whose $(k-1)$-skeleton is complete, and such that every k-face is sampled into X with probability p independently.

The second direct example we'll give is based on the full linear matroid over $\mathbb{F}_q^d$, whose faces consist of linearly independent vectors over $\mathbb{F}_q^d$:

$$X(i) = \{\{v_1, v_2, \ldots, v_i\} \mid v_1, v_2, \ldots, v_i \text{ are independent}\},$$

endowed with the uniform distribution.

Claim 8.21 (Global complexes). The following families are global and HDX:

1. The k-uniform complete complex on n vertices for large enough n is:

    – (one-sided) 0-local-spectral HDX
    – $n^{-\Omega(k)}$-global

2. The Erdos-Renyi hypergraph $G_k(n,p)$ for large enough n is w.h.p:

    – $o_n(1)$-local-spectral HDX

– $(1 - p + o_n(1))$-global

3. The k-skeleton of the full linear matroid over $\mathbb{F}_q^d$ for large enough q, d is :

   – (one-sided) 0-local-spectral HDX
   – $q^{-\Omega(d)}$-global

Proof. We cover each case individually:

Complete Complex. All links of the complete complex are complete and therefore 0-one-sided local spectral expanders. For globalness, observe that the only $\frac{k}{2}$-sets which do not appear in the link of some t are those that intersect it, which are an $1 - \frac{\binom{n-k}{k}}{\binom{n}{k}} \leqslant n^{-\Omega(k)}$ fraction.

Random Hypergraphs. It is well known for fixed p that this complex is an $o_n(1)$-local-spectral expander with high probability so long as $p \gg \frac{k \log(n)}{n}$ (the top links are simply random $G(n, p)$ graphs). Since all other links are at least connected with high probability, Theorem 4.23 implies the local-spectral bound. Toward globality, fix any $\frac{k}{2}$-set $t \subset \binom{[n]}{k/2}$. By Chernoff the probability that more than $(1 - p + o_n(1))$-fraction of $\frac{k}{2}$-sets disjoint from t fail to appear in X is much less than $n^{-\Omega(k)}$ for large enough n. Thus union bounding over t, the worst-case total variation of S_t from uniform over $\binom{n}{\frac{k}{2}}$ is at most $1 - p + o_n(1)$ as desired.[12]

Linear Matroid. All matroids (and therefore their skeletons) are 0-local spectral expanders [Ana+19] (for this matroid specifically, it is folklore). Toward globality, fix $\frac{k}{2}$ linearly independent vectors $B_1 = \{v_1, \ldots, v_{k/2}\}$. The set $B_2 = \{w_1, \ldots, w_{k/2}\}$ only fails to appear in the link of B_1 if B_1 and B_2 are linearly dependent, which happens with probability roughly $\frac{q^{k+(k-1)d}}{q^{dk}} \leqslant q^{-\Omega(d)}$ (ignoring low-order terms). Since $\Pi_{k/2}$ is uniform over linearly independent sets this gives the result. $\qquad\square$

The sampling literature is rife with examples of "ℓ_∞-independent" complexes (a term formally coined in [KKS21]), which are closely related to global complexes. We'll consider the variant implicit in [CLV21]. Let μ be a distribution over $[q]^d$, and observe μ induces a natural d-partite simplicial complex X_μ with

$$X_\mu(d) := \{(1, a_1), \ldots, (d, a_d) : (a_1, \ldots, a_d) \in Supp(\mu)\}$$

and $\pi_X((1, a_1), \ldots, (d, a_d)) = \mu(a_1, \ldots, a_d)$. Let $S \subset [d]$ and $z \in [q]^S$ be such that $\mathbb{P}_{x \sim \mu}[x_S = z] > 0$ (we call such configurations 'feasible'). The (S, z)-*influence matrix* $\Psi_{z \to S}$ has entries

$$\Psi_{z \to S}((u, i), (v, j)) = \begin{cases} \mathbb{P}_{x \sim \mu}[x_i = u | x_j = v, x_S = z] - \mathbb{P}_{x \sim \mu}[x_i = u | x_S = z] & \text{if } j \neq i \\ 0 & \text{if } j = i. \end{cases}$$

A complex is called ℓ_∞-independent if $\|\Psi\|_\infty$ is bounded for all feasible configurations.

Definition 8.22 (ℓ_∞-independence)**.** Fix $q, d \in \mathbb{N}$ and let μ be a distribution over $[q]^d$. μ is called D-ℓ_∞-independent if for all feasible $S \subset [d]$ and $z \in \{0, 1\}^S$:

$$\|\Psi_{S \to z}\|_\infty \leqslant D.$$

We remark that in this regime, one typically thinks of q as fixed or small and the uniformity d as going to infinity. We show any sufficiently low-dimensional skeleton of an ℓ_∞-independent complex is a global HDX:

Proposition 8.23. *Let μ be a D-ℓ_∞-independent distribution over $[q]^d$ and $3 \leqslant k \leqslant d$. Then the k-skeleton $X_\mu^{\leqslant k}$ satisfies:*

1. *$X_\mu^{\leqslant k}$ is a $\frac{D+q}{d+3-k}$-two-sided local-spectral expander*

---

[12]Formally, distance here should be measured from $\pi_{\frac{k}{2}}$, but this is within $o_n(1)$ of uniform in TV with high probability.

2. $X_\mu^{\leqslant k}$ is $\frac{k^2}{8} \cdot \frac{D+q}{d+3-k}$-global.

We remark that (tighter) spectral variants of both items above are well known (see e.g. [ALG21; AP23]). Proposition 8.23 implies essentially every spin-system studied in the recent breakthrough line of work on approximate sampling through spectral independence admits optimal 1% agreement testers (taking $k \leqslant \sqrt{\log(d)}$ skeletons).[13] See [KKS21, Section 3] for an overview including distributions with the stochastic covering property, independent sets, various Ising/Potts models, list colorings, and more.

The proof of Proposition 8.23 relies on an intermediate notion of [Hop24] called local-ℓ_∞-expansion.

Definition 8.24 ((λ, ∞)-expansion)**.** A simplicial complex X is called a (λ, ∞)-local expander if for every $\tau \in X$ with $|\tau| \leqslant d - 2$:
$$\|A_\tau - \Pi_\tau\|_\infty \leqslant \lambda$$
where A_τ is the weighted adjacency matrix of X_τ's 1-skeleton, and Π_τ is its corresponding stationary operator.

We require two lemmas regarding this notion. First, we note that up to dependence on the alphabet q (typically thought of as constant in this regime) ℓ_∞-expansion is essentially equivalent to ℓ_∞-independence.

Lemma 8.25. *Let μ be a D-ℓ_∞-independent distribution over $[q]^d$. Then for every feasible $S \subset [d]$ and $x_S \in \{0,1\}^S$:*
$$\|A_{x_S} - \Pi_{x_S}\|_\infty \leqslant \frac{D+q}{d - |S|}.$$

Proof. We prove the case where $S = \emptyset$. The general statement follows from applying this argument to the links of X_μ. With this in mind, observe that by definition:

$$
\begin{aligned}
\|A_\emptyset - \Pi\|_\infty &= \max_{(u,i)} \sum_{(v,j) \in X_{(u,i)}} \left| \frac{1}{d-1} \mathbb{P}[x_i = u | x_v = j] - \frac{1}{d} \mathbb{P}[x_i = u] \right| \\
&\leqslant \frac{1}{d} \max_{(u,i)} \left(\sum_{(v,j) \in X_{(u,i)}} \left| \mathbb{P}[x_i = u | x_j = v] - \mathbb{P}[x_i = u] \right| + \frac{1}{d-1} \mathbb{P}[x_i = u | x_j = v] \right) \\
&\leqslant \frac{q}{d} + \frac{1}{d} \max_{(u,i)} \sum_{(v,j) \in X_{(u,i)}} \left| \mathbb{P}[x_i = u | x_j = v] - \mathbb{P}[x_i = u] \right| \\
&= \frac{q}{d} + \frac{1}{d} \max_{(u,i)} \sum_{(v,j) \in X_{(u,i)}} \left| \mathbb{P}[x_i = u | x_j = v] - \mathbb{P}[x_i = u] \right| \\
&\leqslant \frac{D+q}{d}
\end{aligned}
$$

$\square$

Second, we'll need the slightly more involved fact from [Hop24] that any ℓ_∞-local expander is global. We reproduce the proof here for completeness.

Lemma 8.26. *Let X be a d-uniform λ-ℓ_∞-local expander. Then X is $\frac{d^2}{8}\lambda$-global.*

Proof. Recall that it is sufficient to bound the infinity norm:
$$\|S_{\frac{d}{2}, \frac{d}{2}} - \Pi_{\frac{d}{2}, \frac{d}{2}}\|_\infty \leqslant \frac{d^2}{4}\lambda.$$

We prove a slightly stronger statement by induction. For any $\tau \in X$ and feasible i, j both the ∞-norm and 1-norm are bounded by:
$$\|S_{i,j}^\tau - \Pi_{i,j}^\tau\|_1, \|S_{i,j} - \Pi_{i,j}\|_\infty \leqslant ij\lambda.$$

---

[13] Note while spectral independence is weaker than what we require, almost all known methods for spin systems actually bound ℓ_∞-independence.

We induct on $i+j$. For the base case $i=j=1$, first observe the ∞-norm is exactly the λ-ℓ_∞-expansion of the 1-skeleton. A classical consequence of Hölder duality states that for any Hölder conjugates (p,q) and operator M

$$\|M\|_p = \|M^*\|_q,$$

where M^* is M's adjoint. Since $S_{1,1} - \Pi_{1,1}$ is self-adjoint we therefore have

$$\|S_{0,0}^\tau - \Pi_{0,0}^\tau\|_1 = \|S_{0,0}^\tau - \Pi_{0,0}^\tau\|_\infty \leqslant \lambda$$

as desired.

Assume now by induction that, for some fixed $i+j > 0$, all $i'+j' < i+j$ and $\tau \in X$ satisfy

$$\|S_{i',j'}^\tau - \Pi_{i',j'}^\tau\|_\infty \leqslant i'j'\lambda.$$

Let $p \in \{1,\infty\}$ and let $\bar{p}$ denote the complement of p. We first argue we may assume without loss of generality that $i > 1$. This is again by Hölder duality since

$$\|S_{i,j} - \Pi_{i,j}\|_p = \|S_{j,i} - \Pi_{j,i}\|_{\bar{p}}.$$

Thus we are done if we show the result for both $p \in \{1,\infty\}$ just for the case $i \geqslant j$ (and thus $i > 1$).

Toward this end, fix any $f : X(j) \to \mathbb{R}$. As is typically the case, the idea is to localize the input by first drawing an $(i-1)$-face, then a vertex from its link:

$$\|S_{i,j}f - \Pi_{i,j}f\|_p = \left\| \|(S_{i,j}f)|_s - (\Pi_{i,j}f)|_s\|_{p,v\in X_s(1)} \right\|_{p,s\in X(i-1)}$$

where for any $g \in X(i) \to \mathbb{R}$, $g|_s(v) = g(s \cup v)$ denotes the localization of g to the link of s. For concreteness, we write out the above explicitly for the $p=1$ case in slightly less compact notation. The $p=\infty$ case is the same replacing expectations with absolute maxima:

$$\|S_{i,j}f - \Pi_{i,j}f\|_p = \mathop{\mathbb{E}}_{t\in X(i)}[|S_{i,j}f(t) - \Pi_{i,j}f(t)|]$$

$$= \mathop{\mathbb{E}}_{s\in X(i-1)}\left[\mathop{\mathbb{E}}_{v\in X_s(1)}[|S_{i,j}f(s \cup v) - \Pi_{i,j}f(s \cup v)|] \right]$$

$$= \mathop{\mathbb{E}}_{s\in X(i-1)}\left[\mathop{\mathbb{E}}_{v\in X_s(1)}[|S_{i,j}f|_s(v) - \Pi_{i,j}f|_s(v)|] \right]$$

$$= \mathbb{E}_{s\in X(i-1)}\left[\mathbb{E}_{v\in X_s(1)}[|S_{i,j}f|_s(v) - \Pi_{i,j}f|_s(v)|] \right]$$

$$= \mathbb{E}_{s\in X(i-1)}\left[\left| \mathbb{E}_{v\in X_s(1)}[|S_{i,j}f|_s(v) - \Pi_{i,j}f|_s(v)|] \right| \right]$$

$$= \left\| \|(S_{i,j}f)|_s - (\Pi_{i,j}f)|_s\|_{p,v\in X_s(1)} \right\|_{p,s\in X(i-1)}.$$

We stick to the more compact norm notation for the remainder of the proof. The trick is now to observe that as a function of $X_s(1)$, $S_{i,j}f|_s$ is exactly $S_{1,j}^s f^s$, where $f^s : X_s(j) \to \mathbb{R}$ is the restriction $f^s(\tau) = f(\tau)$. Then by adding and subtracting the corresponding local stationary operator we have:

$$\|S_{i,j}f - \Pi_{i,j}f\|_p = \left\| \|S_{1,j}^s f^s - \Pi_{1,j}^s f^s + \Pi_{1,j}^s f^s - (\Pi_{i,j}f)|_s\|_{p,v\in X_s(1)} \right\|_{p,s\in X(i-1)}$$

$$\leqslant \left\| \|S_{1,j}^s f^s - \Pi_{1,j}^s f^s\|_{p,v\in X_s(1)} \right\|_{p,s\in X(i-1)} + \left\| \|\Pi_{1,j}^s f^s - (\Pi_{i,j}f)|_s\|_{p,v\in X_s(1)} \right\|_{p,s\in X(i-1)}$$

62

by the triangle inequality. The first term is now bounded by the inductive hypothesis applied in the link of s:

$$\left\| \|S_{1,j}^s f^s - \Pi_{1,j}^s f^s\|_{p,v \in X_s(1)} \right\|_{p,s \in X(i-1)} \leqslant \left\| j\lambda \|f^s\|_{p,X_s(j)} \right\|_{p,s \in X(i-1)} = j\|f\|_{p,X(j)}$$

Toward analyzing the second term, observe that $\Pi_{1,j}^s f^s = S_{i-1,j} f(s)$ and $(\Pi_{i,j} f)|_s(v) = \Pi_{i-1,j} f(s)$ so:

$$\left\| \|\Pi_{1,j}^s f^s - (\Pi_{i,j} f)|_s\|_{p,v \in X_s(0)} \right\|_{p,s \in X(i-1)} = \|S_{i-1,j} f - \Pi_{i-1,j} f\|_{p,s \in X(i-1)} \leqslant (i-1)j\lambda \|f\|_{p,X(j)}$$

by the inductive hypothesis. Altogether this gives

$$\|S_{i,j} f - \Pi_{i,j} f\|_p \leqslant \lambda j + \lambda(i-1)j = ij\lambda$$

as desired. $\qquad\square$

The proof of Proposition 8.23 is now essentially immediate.

Proof of Proposition 8.23. The first fact is immediate from Lemma 8.25 and the fact that $\|A_\tau - \Pi_\tau\|_2 \leqslant \|A_\tau - \Pi_\tau\|_\infty$. The second fact is an immediate consequence of combining Lemma 8.25 and Lemma 8.26. $\quad\square$

We conclude the subsection by the observation that if a complex X is global, then so is its faces complex (Definition 5.7). Since the faces complex also inherits the two-sided expansion of the original complex up to factors in uniformity (c.f. the proof of Theorem 5.1), this shows that one can also take as examples faces complexes of any of the examples above.

Claim 8.27. Let X be a k-uniform and λ-global complex. Let ℓ be such that ℓ divides k, then the ℓ-faces complex F_X^ℓ is λ-global.

Proof. Let $m = \frac{k}{\ell}$ for notational convenience. Fix $t = \{s_1, s_2, \ldots, s_{\frac{m}{2}}\} \in F_X^\ell(\frac{m}{2})$. Observe that the set of neighbors of t with respect to the swap walk in $F^\ell X$ are exactly those t' such that $\cup t$ and $\cup t'$ are neighbors in the swap walk on $X(\frac{k}{2})$ (here $\cup t = s_1 \cup s_2 \cup \cdots \cup s_{\frac{m}{2}}$ and similarly for t'). Moreover, one samples a neighbor of t by sampling a random neighbor τ of $\cup t$, and then sampling a uniform at random partition of τ to $t' = \{s_1', s_2', \ldots, s_{\frac{m}{2}}\}$ (the partition is chosen independent of t).

Thus one observes that the TV-distance between S_t and Π can be written as

$$\frac{1}{2} \sum_{t' \in F^\ell X(\frac{m}{2})} |\underset{S_t(F^\ell X)}{\mathbb{P}}[t'] - \underset{\Pi}{\mathbb{P}}[t']| = \frac{1}{2} \sum_{\tau \in X(\frac{k}{2})} \sum_{t': \cup t' = \tau} |\underset{S_t(F^\ell X)}{\mathbb{P}}[t'] - \underset{\Pi}{\mathbb{P}}[t']|$$

$$= \frac{1}{2} \sum_{\tau \in X(\frac{k}{2})} \sum_{t': \cup t' = \tau} p|\underset{S_{\cup t}(X)}{\mathbb{P}}[\tau] - \underset{\Pi(X)}{\mathbb{P}}[\tau]|$$

where p is one over the number of partitions of τ to t'. Obviously this is equal to

$$\frac{1}{2} \sum_{\tau \in X(\frac{k}{2})} \underset{S_{\cup t}(X)}{\mathbb{P}}[\tau] - \underset{\Pi(X)}{\mathbb{P}}[\tau] = d_{TV}(S_{\cup t}(X), \Pi(X)) \leqslant \lambda$$

as desired. $\qquad\square$

9   Analytic, Geometric, and Combinatorial Applications

In this section, we cover several brief applications of sampling and reverse hypercontractivity in well-studied combinatorial and analytic settings.

9.1 New Double Samplers

Double samplers are a strengthened notion of sampling introduced in [DK17; Din+18a] that consist of two "stacked" samplers with additional local sampling properties under 'closure' of the top layer. These interesting objects have powerful applications in agreement testing [DK17] and list-decoding [Din+18a], and have even seen algorithmic use in the construction of space-efficient data structures for the heavy hitter problem [DW22].

Definition 9.1 (Double sampler, [Din+18a]). A double sampler consists of a triple (V_2, V_1, V_0), where V_0 is the ground set, V_1 is a collection of i-subsets of V_0 and V_2 is a collection of k-subsets of V_0, where $k > i \in \mathbb{N}$. We say that (V_2, V_1, V_0) is an $(\varepsilon, \beta, \varepsilon_0, \beta_0)$-*double sampler* if

- The inclusion graphs on (V_2, V_1) and (V_2, V_0) are (ε, β)-additive samplers and the inclusion graph on (V_1, V_0) is an $(\varepsilon + \varepsilon_0, \beta + \beta_0)$-additive sampler[14] (recall the inclusion graph is defined by connecting two subsets by an edge if one contains the other).

- For every $T \in V_2$, let $V_1(T) = \{S \in V_1 \ : \ S \subset T\}$ be the sets in V_1 that are contained in T. Let $G_{|T}$ be the bipartite inclusion graph connecting elements in T (viewed as elements in the ground set V_0) to subsets in $V_1(T)$. We require that for every $T \in V_2$, the graph $G_{|T}$ is an (β_0, ε_0)-additive sampler.

Double samplers are only known to arise from high dimensional expanders. In their original work, [DK17; Din+18a] use the Ramanujan complexes [LSV05a] to construct explicit double samplers for all $(\varepsilon, \beta, \varepsilon_0, \beta_0)$:

Theorem 9.2 ([Din+18a, Theorem 2.11]). *For every $(\varepsilon, \beta, \varepsilon_0, \beta_0) > 0$, there exists an explicit infinite family of double samplers $\{(V_0^{(n)}, V_1^{(n)}, V_2^{(n)})\}$ such that for all n:*

$$\frac{|V_1^{(n)}|}{|V_0^{(n)}|}, \frac{|V_2^{(n)}|}{|V_0^{(n)}|} = \exp\left(\text{poly}\left(\frac{1}{\beta}, \frac{1}{\beta_0}, \frac{1}{\varepsilon}, \frac{1}{\varepsilon_0}\right)\right).$$

One of the main open questions posed in [Din+18a] is to settle the overhead of double samplers. In particular, while it is known that exponential dependence on ε_i is necessary (even for standard samplers), it is plausible the dependence on the 'failure probability' β_i could be improved, perhaps even to match the corresponding polynomial dependence of optimal standard samplers. Since typical related applications (e.g. for low soundness agreement tests and PCPs [IKW12; DL17]) take ε to be constant (or at most some polylog in β), this latter dependence is where the crux of the problem lies.

To the best of our knowledge, there have been no lower bounds or improvements over Theorem 9.2 since [Din+18a]'s original work. Leveraging our concentration bounds and their corresponding optimality (see Section 11.1) we take a significant step toward resolving this problem: quasi-polynomial size double samplers exist and, under reasonable assumptions on the underlying complex, are the best possible.

Theorem 9.3 (Quasi-Polynomial Double Samplers). *for every $(\varepsilon, \beta, \varepsilon_0, \beta_0) > 0$, there exists an explicit family of $(\beta, \varepsilon), (\beta_0, \varepsilon_0)$ double samplers $\{(V_2^{(n)}, V_1^{(n)}, V_0^{(n)})\}$ of size at most:*

$$\frac{|V_1^{(n)}|}{|V_0^{(n)}|}, \frac{|V_2^{(n)}|}{|V_0^{(n)}|} \leqslant \exp\left(\tilde{O}\left(\frac{\log^3 \frac{1}{\beta} \log^3 \frac{1}{\beta_0}}{\varepsilon^6 \varepsilon_0^6}\right)\right).$$

Moreover, for $\varepsilon \in (0, 0.01)$, if the underlying complex family $\{X_n\}$ is $\frac{\varepsilon}{3}$-hitting, vertex-uniform, and non-contracting,[15] this is optimal up to polynomial factors in the exponent:

$$\frac{|V_2^{(n)}|}{|V_0^{(n)}|} \geqslant \exp\left(\Omega\left(\frac{\log(\frac{1}{\beta}) \log(\frac{1}{\beta_0})}{\varepsilon^2 \varepsilon_0^2}\right)\right)$$

---

[14]The weights over edges in (V_i, V_j) are the marginals of the probabilistic experiment where we first choose $T \in V_2$ (according to some given weight distribution Π), and then choose $v \in V_0$ and $S \in V_1$ such that $v \in S \subseteq T$, uniformly at random over all such pairs S, v.

[15]Here we mean 1) $\pi_0(v) = |X(1)|^{-1}$ for all $v \in X(1)$, and 2) $|X(i)| \geqslant |X(i-1)|$ for all $2 \leqslant i \leqslant d$.

Note the above does not fully rule out the existence of simplicial poly-size double samplers. While the assumption on regularity in the lower bound can be easily relaxed, it is unclear to what extent the assumption on hitting-set is an artifact of our proof technique (see Section 11.1 for details), though it is certainly a natural property for such objects to satisfy. Moreover, our bound strongly relies on the assumption that $|V_2| \geqslant \exp(d)|V_0|$ (implied by X being non-contracting), which is not generically true. Nevertheless, high dimensional expanders exhibit both good hitting set behavior (see Proposition 10.12) and at least exponential degree (see Proposition 9.6), so it is at least accurate to say double samplers arising from HDX are at best quasi-polynomial in β^{-1}.

The proof of Theorem 9.3 relies on two further results. The first, toward our upper bound, is Lubotzky, Samuels, and Vishne's [LSV05b] classical construction of Ramanujan complexes.

Theorem 9.4 (LSV-Complexes [LSV05b])**.** *For any $\lambda > 0$ and $d \in \mathbb{N}$, there exists an explicit infinite family of d-partite one-sided λ-local-spectral expanders $\{X^{(n)}, \Pi^{(n)}\}$ with degree $deg(X) \leqslant \lambda^{-O(d^2)}$. Moreover, for every $d \geqslant k > \ell \geqslant 0$, $\max\{\pi_k\} \leqslant o_n(1)$.*

The second, toward the lower bound, is the following optimality result we prove in Section 11.1.

Corollary 9.5 (Corollary of Theorem 11.1)**.** *For any $\varepsilon \in (0, 0.01)$, let X be a d-uniform, $\frac{\varepsilon}{3}$-hitting vertex-uniform complex and $i \leqslant k \leqslant d$. Then there exists a set $A \subset X(i)$ such that:*

$$\mathbb{P}_{s \in X(k)}[|U_{i,k}1_A(s) - \mu_A| > \varepsilon] \geqslant \exp\left(O\left(\varepsilon^2 \frac{k}{i}\right)\right).$$

We are now ready to prove Theorem 9.3.

Proof of Theorem 9.3. We split into the upper and lower bounds.

Upper Bound. Let $k = O\left(\frac{\log(1/\beta)\log(1/\beta_0)}{\varepsilon^2 \varepsilon_0^2}\right)$, $d = k\log(k)$, and $\{X_n\}$ be an explicit infinite family of d-uniform LSV-complexes with local-spectral expansion $\lambda = 2^{-\Omega(d)}$, and take $V_2 = X(k)$, $V_1 = X(i)$, and $V_0 = X(1)$ for $i = \Theta(\frac{\log(1/\beta_0)}{\varepsilon_0^2})$. By Theorem 5.1, (V_2, V_1) is an (ε, β)-sampler for the right choice of constants. Moreover, every closure graph is simply the $(\frac{\log(1/\beta_0)}{\varepsilon_0^2})$-uniform complex on k vertices, which is negatively correlated and therefore also a (ε_0, β_0)-sampler by standard Chernoff-Hoeffding for negatively correlated variables. The overhead bound is then immediate from the degree of LSV complexes in Theorem 9.4.

Lower Bound. Recall that by definition $V_0^{(n)} = X_n(1) = [n]$, $V_1^{(n)} = X_n(i)$, and $V_2^{(n)} = X_n(k)$ for some family of simplicial complexes $\{X_n\}$ and set sizes $1 < i < k$. By Corollary 9.5, for any $\varepsilon \in (0, 0.01)$, $(X(k), X(i))$ is at best an $(\varepsilon, e^{O(\varepsilon^2 \frac{k}{i})})$-sampler, therefore any such double sampler requires:

$$\frac{k}{i} \geqslant \Omega\left(\frac{\log(1/\beta)}{\varepsilon^2}\right).$$

On the other hand, it is also required for every $T \in X(k)$ that the induced graphs $G|_T = (V_2|_T, V_1|_T)$ are (ε_0, β_0)-samplers. By standard sampling lower bounds [CEG95], this graph is at best a $(\varepsilon_0, e^{O(\varepsilon_0^2 i)})$-sampler, forcing

$$i \geqslant \Omega\left(\frac{\log(1/\beta_0)}{\varepsilon_0^2}\right).$$

Combining these bounds we have that for any valid choice of i the top uniformity k of an $(\varepsilon, \beta, \varepsilon_0, \beta_0)$-double sampler is at least

$$k \geqslant \Omega\left(\frac{\log(\frac{1}{\beta})\log(\frac{1}{\beta_0})}{\varepsilon^2 \varepsilon_0^2}\right).$$

It is therefore enough to argue that the overhead is at least exponential in the uniformity k:

$$\frac{|V_2|}{|V_0|} \geqslant 2^{\Omega(k)}.$$

This is true for any non-contracting complex. In particular, since $|X(k)| \geqslant |X(k/2)|$, the average degree of every $(k/2)$-face (and hence every vertex) is at least $2^{\Omega(k)}$. □

We remark that it may also be possible to construct double samplers with smaller overhead by removing the *simplicial* requirement from the definition (see e.g. the relaxed notion in [DW22]). It is likely, for instance, that 'Grassmannian HDX' [KT21; Gol23; DLZ23] based on subspace structure lead to double-samplers with different parameters-overhead trade-offs.

9.2   A Degree Lower Bound for HDX

One of the most classical results in the study of expander graphs is the Alon-Boppana Theorem [Nil91], roughly stating that any family of bounded-degree λ-expanders must have degree at least $\frac{2}{\lambda^2}$. In higher dimensions, the situation is less understood. Not only should degree scale in some way with the local expansion λ, but dependence on *dimension* becomes a critical parameter in application. In this section, we give (to our knowledge) the first systematic study of the degree of high dimensional expanders. In particular, we prove that various families of hyper-regular HDX require super-exponential degree and prove a threshold-phenomenon exists at the TD-barrier: there exist (reasonably balanced) constructions of 1-TD complexes achieving exponential degree at every level, while any (reasonably balanced) λ-TD complex must have degree $2^{-\Omega(k^2)}$ for any level $k \leqslant \sqrt{d}$.

Before applying our stronger concentration bounds, we first look at what one can infer about the degree of HDX directly from spectral methods. The following proposition, though not to our knowledge appearing anywhere in the literature, is the result of applying elementary spectral methods to the problem:

Proposition 9.6 (Spectral Lower Bound)**.** *Let X be a d-uniform $\{\lambda_i\}$-two-sided local-spectral expander. Then for every $k < d$, the degree of X's k-skeleton is at least:*

$$\deg^{(k)}(X) \geqslant \frac{1}{(k-1)!} \prod_{i=1}^{k-1} \frac{1}{\lambda_i}$$

Proof. Observe the k-th degree of a vertex $v \in X(1)$ can be bounded by

$$\frac{1}{(k-1)!} \prod_{i=1}^{k-1} \min_{t \in X(i)} \{|X_t(1)|\}$$

simply by counting the number of ways a face $\{v_2, \ldots, v_{k-1}\} \in X_{v_1}(k)$ can be sampled by choosing $v_2 \sim X_{v_1}(1)$, $v_2 \in X_{\{v_1,v_2\}}(1)$ and correcting for the $\frac{1}{(k-1)!}$ possible orderings of the face. The result now follows from the elementary fact (see e.g. [Gai+22, Lemma A.7]) that any two-sided λ-expander has at least $\frac{1}{\lambda}$ vertices. □

We remark that one might reasonably think to use Alon-Boppana [Nil91] (the classic 'optimal' trade-off between expansion and degree) to prove a version of the above for one-sided expansion. This does not work for high dimensional expanders, since their links may have low diameter.

Proposition 9.6 implies that any (hyper-regular) $o(\frac{1}{d})$-two-sided HDX must have super-exponential degree but is essentially trivial for complexes near the TD-barrier, under spectral independence, or even for arbitrarily strong partite one-sided HDX. In these cases, even if one is willing to look e.g. at polynomial size cutoffs (as is quite common in the literature) Proposition 9.6 only gives a slightly super-exponential lower bound:

$$\deg^{(k)}(X) \geqslant \exp(\Omega(k \log(k))).$$

Using concentration, we improve this bound to strongly super-exponential. We start with the partite case.

Theorem 9.7. *Fix $c \in [0, 1]$ and let X be a d-uniform hyper-regular partite $2^{-\Omega(d^c)}$-HDX. Then either:*

1. *X's 1-skeleton is dense:*[16]

$$\deg^1(X) \geqslant \frac{|X(1)|}{2}$$

---

[16]While this condition may seem strange, note it is necessary to handle dense settings. Take, e.g., $\{0, 1\}^n$, which is a hyper-regular 0-local-spectral expander whose $\sqrt{n}$-skeleton has degree $\exp(\sqrt{n}\log(n))$.

2. X has super-exponential degree:

$$\deg^{(d)}(X) \geqslant \exp(\Omega(\max\{d^{2c}, d\}))$$

Second we prove an analogous bound for skeletons of TD and SI-complexes.

Theorem 9.8. *Let $\lambda, c \in [0, 1)$, $\eta > 0$, and $d \in \mathbb{N}$. If X is a d-uniform hyper-regular λ-TD or η-SI complex then for any $k \leqslant d^c$:*

1. X's 1-skeleton is dense:

$$\deg^1(X) \geqslant \frac{|X(1)|}{2}$$

2. X's k-skeleton has super-exponential degree:

$$\deg^{(k)}(X) \geqslant \exp(\Omega_{\lambda,\eta}(\min\{k^2, k^{1/c}\}))$$

The latter setting above is tight in the sense that when $\lambda = 1$, that is *at* the TD-barrier, there exist (reasonably balanced) complexes where $\deg^{(k)}(X) = \exp(O(k))$ for all $k \leqslant d$. On the other hand, beyond the TD barrier the best known degree upper bound for a hyper-regular (or even reasonably balanced) HDX is $\exp(\exp(d))$ [FI20], leaving a substantial gap with the above. Nevertheless, Theorem 9.8 is the first to exhibit quantitative threshold behavior for degree at the TD-barrier.

Theorem 9.7 and Theorem 9.8 leave open two critical questions:

1. Can we prove super-exponential bounds for the *top* level of a λ-TD/η-SI complex?

2. Can the assumption of hyper-regularity be dropped?

As we will see in the proof of Theorem 9.8, the first question would be resolved if one could give *any* asymptotic improvement over exponential concentration for weak HDX. In other words, any bound of the form:

$$\mathbb{P}_{s \in X(d)} \left[\frac{1}{d}|s \cap A| - \mu > \frac{1}{2} \right] \leqslant \exp(\omega(\sqrt{d}))$$

for X and its links would imply a super-exponential degree lower bound. If, for instance, we have a bound of $\exp(-\Omega(d))$ holds, this would imply a degree lower bound of $\exp(\Omega(d^2))$ on the top level of X.

Regarding the second question, our proof technique extends to any complex whose links are reasonably 'balanced' in the sense that there should not be a $2^{-\Omega(\sqrt{d})}$ fraction of vertices making up an $\Omega(1)$ fraction of the mass. Unfortunately, many classical HDX constructions, e.g. the Ramanujan complexes of [LSV05b], actually are highly unbalanced in this sense. The links of these constructions essentially approximate an 'unbalanced product': they are partite complexes which are regular on each individual part, but many of the parts are vanishingly small and therefore make up a significant portion of the mass. While known constructions of this type all have minimum part size at least exponential in d (and therefore result in degree $\exp(d^2)$), it is unclear how to rule out a construction whose links look, e.g., like $\{0, 1\}^d \times [2^d]$. Such a complex could potentially satisfy optimal inclusion sampling while maintaining exponential degree.

We leave the possible extension of our lower bounds to the top level of unbalanced complexes (or conversely the construction of highly locally unbalanced exponential-size HDX) as one of the main open problems suggested by this work:

Question 9.9 (Lower Bounds for Locally Unbalanced HDX)**.** Let $\{X_n\}$ be an infinite family of bounded-degree HDX that are either λ-TD or η-SI. Is the degree of every (sufficiently large) X_n super-exponential:

$$\deg(X_n) \geqslant \exp(\omega(d))?$$

Question 9.9 is critical from the standpoint of applying HDX in other areas of theoretical computer science where degree dependence corresponds to the 'overhead' incurred by using the complex as a gadget. For instance, the lower bound in Theorem 9.3 can be achieved by HDX essentially if and only if Question 9.9 is false. Similarly, a variant of Question 9.9 where d may depend on n is critical to the study of agreement

testers whose soundness scale with the *number of vertices* of the system. Constructing polynomial size testers with inverse polynomial soundness is the combinatorial core of the so-called *sliding scale conjecture* in PCP theory. If the above lower bound holds, HDX cannot give such objects: d must be taken to be $o(\log(n))$ to maintain polynomial size, and the resulting soundness of any HDX-based poly-size agreement tester would be at best $2^{-o(\log(n))}$, missing the inverse polynomial mark.

We now move to the proofs of Theorem 9.7 and Theorem 9.8 which are inspired by our *localization* technique for proving concentration that samples $s = \{v_1, v_1, \ldots, v_d\}$ by iteratively sampling $v_0 \in X$, then $v_2 \in X_{v_1}$, $v_3 \in X_{\{v_1,v_2\}}$... and analyzing the local concentration at each step. Upon further inspection, this technique yields the fact that the 1-skeleton of a complex with good sampling in 0 and 1-links must itself be an excellent sampler graph.

Lemma 9.10. *Fix $\varepsilon, \delta_1, \delta_2 > 0$. Let X be a d-uniform simplicial complex such that:*

1. *$(X(d), X(1))$ is a $\left(\frac{\varepsilon}{2}, \delta_1\right)$-additive sampler.*

2. *For every $v \in X(1)$, $(X_v(d), X_v(1))$ is an $\left(\frac{\varepsilon}{2}, \delta_2\right)$-additive sampler.*

Then the partite double cover of X's 1-skeleton is a $\left(\varepsilon, \frac{2\delta_1}{1-\delta_2}\right)$-sampler.

Proof. Fix $A \subseteq X(1)$. At a high level, the proof follows from observing that if the base graph is a poor sampler, then after drawing the first vertex v_1 of our d-set $t = \{v_1, \ldots, v_d\}$, there is a non-trivial chance that the neighborhood of v_1 (i.e. $X_{v_1}(1)$) 'sees' A in the wrong proportion. However, since the remainder of the face $\{v_2, \ldots, v_d\} \in X_{v_1}(d)$, if $(X_{v_1}(d), X_{v_1}(1))$ is indeed a good sampler it will maintain this incorrect proportion from the first step and violate sampling of A in $(X(d), X(1))$.

We now formalize this argument. Given our assumptions we'd like to show:

$$\mathbb{P}_{v \in X(1)}\left[\mathbb{P}_{u \in X_v(1)}[A] > \mathbb{P}[A] + \varepsilon\right] \leqslant \frac{\delta_1}{1-\delta_2}.$$

The lower tail follows from applying this on the complement set. Toward this end, denote by B the set of 'bad' vertices in A that over-sample the base graph:

$$B = \left\{v \in X(1) \;\middle|\; \mathbb{P}_{u \in X_v(1)}[A] > \mathbb{P}[A] + \varepsilon\right\}.$$

We need to show $\mathbb{P}[B] \leqslant \frac{\delta_1}{1-\delta_2}$. We bound $\mathbb{P}[B]$ by relating it to the set of bad d-faces:

$$C = \left\{t \in X(d) \;\middle|\; \left|\mathbb{P}_{u \in t}[A] - \mathbb{P}[A]\right| > \frac{\varepsilon}{2}\right\},$$

which has measure at most $\mathbb{P}[C] \leqslant \delta_1$ by Condition (2). The key to relate B and C is to observe we can sample a random d-face t by first sampling a vertex $v \in X(1)$, then sampling t from its link:

$$\mathbb{P}[C] = \mathbb{E}_{t \in X(d)}[\mathbf{1}_C(t)] = \mathbb{E}_{v \in X(1)}\left[\mathbb{E}_{t \in X_v(d)}[\mathbf{1}_C(t)]\right] \geqslant \mathbb{P}[B]\,\mathbb{E}_{v \in X(1)}\left[\mathbb{E}_{t \in X_v(d)}[\mathbf{1}_C(t)] \;\middle|\; v \in B\right].$$

When $v \in B$, we have $\mathbb{P}_{u \in X_v(1)}[A] \geqslant \mathbb{P}[A] + \varepsilon$ by assumption, but then by Condition (1)

$$\mathbb{E}_{v \in X(1)}\left[\mathbb{E}_{t \in X_v(d)}[\mathbf{1}_C(t)] \;\middle|\; v \in B\right] \geqslant \mathbb{P}_{t \in X_v(d)}\left[\mathbb{P}[A \mid t] > \mathbb{P}_{u \in X_v(1)}[A] - \frac{\varepsilon}{2}\right] \geqslant 1 - \delta_2.$$

Putting everything together we have

$$(1 - \delta_2)\,\mathbb{P}[C] \leqslant \mathbb{P}[B] \leqslant \delta_1.$$

and $\mathbb{P}[C] \leqslant \frac{\delta_1}{1-\delta_2}$ as desired. $\qquad\square$

Combining this fact with standard degree lower bounds for sampler graphs [CEG95] gives the following degree lower bound for the 1-skeletons of such complexes:

Corollary 9.11. *Let X be a d-uniform simplicial complex satisfying:*

1. *$(X(d), X(1))$ is a $(\frac{1}{10}, \delta_1)$-sampler*

2. *For all $v \in X(1)$, $(X_v(d), X_v(1))$ is a $(\frac{1}{10}, \delta_2)$-sampler.*

Then X's 1-skeleton has high max-degree:

$$\exists v \in X(1) : \deg^{(1)}(v) \geqslant \min \left\{ \frac{|X(1)|}{2}, \frac{2(1 - \delta_2)}{5\delta_1} \right\}$$

Proof. By Lemma 9.10 the underlying graph of X is a $(\frac{1}{5}, \delta = \frac{2\delta_1}{1 - \delta_2})$-sampler. [CEG95, Theorem 2] states the following relation on a graph $G = (L, R, E)$ that is an (ε, δ)-sampler of max degree $t \leqslant \frac{|X(1)|}{2}$.

$$t|R| \geqslant \frac{|L|(1 - 2\varepsilon)}{2\delta}.$$

Here $L = R = X(1)$ and this inequality gives $t \geqslant \frac{0.8}{2\delta} = \frac{2(1 - \delta_2)}{5\delta_1}$. $\qquad\square$

Notice this bound is substantially stronger than what can be inferred from spectral expansion alone. For instance, when X is a $\lambda = \text{poly}(d^{-1})$-HDX, Alon-Boppana only implies a degree lower bound of $\Omega(\lambda^{-2}) = \text{poly}(d^{-1})$, whereas combining the above with Corollary 7.7 gives degree at least $\exp(-\sqrt{d})$. Given this fact, Theorem 9.7 and Theorem 9.8 follows by recursive application of Corollary 9.11 on each link of the HDX.

Proof of Theorem 9.7 and Theorem 9.8. Similar to the spectral argument, since X is hyper-regular the kth degree of any vertex $v_0 \in X(1)$ can be bounded by:

$$\deg^{(k)}(v_0) \geqslant \frac{1}{(k-1)!} \prod_{i=1}^{k-1} \min_{t \in X(i)} \left\{ \deg^{(1)}(X_t) \right\}.$$

It is therefore sufficient to bound the degree of the 1-skeleton of the links of X. We split the analysis into the partite and TD/SI cases.

Partite Case. Let $k' = \max\{\sqrt{d}, d^c\}$. By Theorem 7.1, there exists a universal constant $c' > 0$ such that for every $t \in X^{\leqslant \frac{d}{2}}$, the link X_t satisfies:

1. $(X_t(d - |t|), X_t(1))$ is a $(\frac{1}{10}, \exp(-c'k'))$-sampler

2. For all $v \in X_t(1)$, $(X_{t \cup \{v\}}(d - |t| - 1), X_{t \cup \{v\}}(1))$ is a $(\frac{1}{10}, \exp(-c'k'))$-sampler.

Setting $t \in X(1)$, Corollary 9.11 implies

$$\deg^{(1)}(X) \geqslant \min \left\{ \frac{|X(1)|}{2}, \exp(-c'k') \right\}.$$

Assume $\exp(-c'k') \leqslant \frac{|X(1)|}{2}$ (else Condition (1) is satisfied and we are done). Under this assumption, we argue by induction that the max degree of any i-link is at least:

$$\forall i \leqslant \frac{d}{2} : \quad \min_{t \in X(i)} \{\deg^{(1)}(X_t)\} \geqslant \frac{1}{2^i} \exp(c'k'). \tag{9.1}$$

We remark the restriction on i is to enforce that links of this level are still $(\frac{1}{10}, \exp(-c'k'))$-samplers. Plugging this back into our bound on degree gives $\exp(-\Omega(k'^2)) \geqslant \exp(-\Omega(\max\{d, d^{2c}\}))$ as desired.

To prove (9.1), observe the base case $i = 1$ is given by assumption. The inductive step simply follows from observing that the minimum number of vertices in any i-link is exactly the minimum degree of the 1-skeleton of any $(i-1)$-link. Thus applying Corollary 9.11, we have:

$$\forall t \in X(i) : \deg^{(1)}(X_t) \geq \min\left\{\frac{|X_t(1)|}{2}, \exp(-c'k')\right\} \geq \frac{1}{2^i}\exp(c'k').$$

as desired.

TD/SI Case. When X is λ-TD or η-SI, we'd like to run the same argument as above but run into a slight problem. We can still bound the kth degree of a vertex v by:

$$\deg^{(k)}(v) \geq \frac{1}{(k-1)!}\prod_{i=1}^{k-1}\min_{t \in X(i)}\left\{\deg^{(1)}(X_t)\right\},$$

but because X satisfies only exponential concentration, (9.1) becomes

$$\forall i \leq \frac{d}{2} : \min_{t \in X(i)}\{\deg^{(1)}(X_t)\} \geq \frac{1}{2^i}\exp(c'\sqrt{d}).$$

This bound is now only non-trivial up to $i \leq O(\sqrt{d})$ levels of X, stopping us from proving super-exponential concentration for X's top level. Nevertheless, when $k = d^c$, plugging the above into the degree calculation gives $\exp(d) = \exp(k^{1/c})$ whenever $c \geq 1/2$, and $\exp(k\sqrt{d}) \geq \exp(k^2)$ whenever $c < 1/2$. $\square$

9.3 Geometric Overlap

The geometric overlap property is a classical notion of high dimensional expansion (see e.g. [Lub18]) which promises every embedding of the complex has a point in space hit by at least a constant fraction of the faces.

Definition 9.12. Let X be a d-uniform simplicial complex and let $k \leq d - 1$. Let $c > 0$. We say that X has (k, c)-geometric overlap if for every $\rho : X(1) \to \mathbb{R}^k$, there exists a point $q_0 \in \mathbb{R}^k$ such that

$$\mathop{\mathbb{P}}_{s \in X(d)}[q_0 \in conv(\rho(s))] \geq c.$$

Here $conv(\rho(s))$ is the convex hull of the points of $\rho(s) = \{\rho(v) \mid v \in s\}$. While we are not aware of any applications of geometric overlap in computer science, it is a popular topic in geometry and topology and related ideas (e.g. an overlap variant of the Borsuk-Ulam theorem) have seen recent use in the study of algorithmic stability [Cha+23].

Boros and Füredi [BF84] showed that the 2-dimensional complete complex has geometric overlap, a result which was later generalized to all dimensions by Bárány [Bár82]. The notion was extended to general simplicial complexes by Gromov [Gro10], who asked whether there are bounded degree complexes that have this property. Followup work by Fox, Gromov, Lafforgue, Naor, and Pach [Fox+12] gave an affirmative answer to this question both via random construction and through [LSV05a]'s Ramanujan complexes. Finally, a series of works [PRT16; Evr17; Opp18] extended their ideas to general spectral high dimensional expanders by leveraging high dimensional expander-mixing properties. We note that Gromov also defined a closely related (but stricter) notion of topological overlap [Gro10], which was later studied by [KKL14; EK16] among others. We focus only on the geometric setting here.

While prior works have succeeded in showing optimal geometric overlap for *random* bounded-degree complexes [Fox+12, Theorem 1.7], it has been open since Gromov's work to construct such a family *explicitly*. In this subsection, building on the approach of these prior works, we show sufficiently strong high dimensional expanders have optimal overlap, resolving this problem.

To state our result, it is useful to first understand the overlap properties of the complete complex. Let $c_{k,n}$ be the largest constant so that $\Delta_n(k+1)$ has $(k, c_{k,n})$-geometric overlap, and let $c(k) = \lim_{n\to\infty} c_{k,n}$. By [Bár82] this limit exists and is positive. Similarly, let $c^p_{d,k,n}$ be the largest constant so that the complete d-partite complex with n-vertices on every side has $(k, c^p_{d,k,n})$-geometric overlap, and let $c^p_{d,k} = \lim_{n\to\infty} c^p_{d,k,n}$.

Existence and positivity of this limit follows from [Pac98]. The former constant, $c(k)$, is known to be the best possible for bounded degree simplicial complexes [Fox+12, Theorem 1.6].

We show that every sufficiently strong (partite) high dimensional expanders has geometric overlap nearly matching the (partite) complete complex:

Theorem 9.13. *For every $\varepsilon > 0$ and $k < d \in \mathbb{N}$, there exist $\lambda, \lambda_p > 0$, and $n_0, n_{0,p} \in \mathbb{N}$ such that*

1. ***Two-sided case:*** *If X is d-uniform λ-two-sided high dimensional expander on at least n_0 vertices so that the measure on the vertices is uniform, then X has $(k, c_k - \varepsilon)$-geometric overlap.*

2. ***Partite case:*** *If X is d-partite λ_p-one-sided HDX on at least $n_{0,p}$ vertices so that the measure on the vertices is uniform, then X has $(k, c_{d,k}^p - \varepsilon)$-geometric overlap.*

We comment that this theorem applies to many of the complexes constructed in [LSV05a] (for the non-partite setting we need to take lower uniform skeletons).

The proof of Theorem 9.13 combines our concentration bounds with tools developed in [Fox+12]. We have not tried to find the optimal dependence on λ or on the degree of the complex. Finally before moving to the proof, we remark that the result really only relies on the weaker 'high-dimensional expander-mixing' type inequality which actually holds for the more general family of splitting-trees such as expander walks (see Section 10.1). Thus it is possible to obtain explicit complexes with geometric overlap approaching that of the complete (partite) complex through this more general family as well, though we omit the details.

Homogeneous tuples. The core of [Fox+12]'s proof of geometric overlap is the following partitioning theorem. We follow their definitions and presentation precisely. A tuple of subsets $S_1, \ldots, S_{k+1} \subseteq \mathbb{R}^k$ is said to be *homogeneous* with respect to a point $q \in \mathbb{R}^k$ if either:

1. All simplices with one vertex in each of the sets $S_1, \ldots, S_{k+1}$ contain q.

2. None of these simplices contain q.

An *equipartition* of a finite set is a partition of the set into subsets whose sizes differ by at most one.

Theorem 9.14 ([Fox+12, Corollary 1.9]). *Let d be a positive integer and $\varepsilon > 0$. There exists a positive integer $K = K(\varepsilon, k) \geqslant k+1$ such that for any $K' \geqslant K$ the following statement holds. For any finite set $P \subseteq \mathbb{R}^k$ and for any point $q \in \mathbb{R}^k$, there is an equipartition $P = P_1 \uplus \cdots \uplus P_{K'}$ such that all but at most an ε-fraction of the $(k+1)$-tuples $P_{i_1}, \ldots, P_{i_{k+1}}$ are homogeneous with respect to q.*

Our proof adapts the idea in [Fox+12, Theorem 1.7].

Proof of Theorem 9.13. We split the proof into the two-sided and partite cases.

Two-sided Case. Let $K = K(\varepsilon/16, k)$ be the constant promised in Theorem 9.14 and $K' \geqslant K$ be the smallest constant so that for any equipartition of the n-vertices to K' parts, the fraction of faces in the complete complex (over the same vertex set) that touch every part at most once is at least $1 - \frac{\varepsilon}{16}$.[17] Let $z_0 = \frac{1}{2K'}$ and let n_0 be such that for any $n \geqslant n_0$ and sets $A_1, A_2, \ldots, A_{k+1} \subset [n]$ of measure at least $\frac{1}{K'} - z_0$ it holds that

$$\mathbb{P}_{\{v_1, v_2, \ldots, v_{k+1}\} \in \Delta_n(k+1)} \left[\forall j = 0, 1, \ldots, d : v_j \in A_j \right] \geqslant \prod_{j=1}^{k+1} \mathbb{P}\left[A_j \right] - \frac{\varepsilon}{16}.$$

Let X be as in the statement and $\rho : X(1) \to \mathbb{R}^k$ an embedding. There is a point $q_0 \in \mathbb{R}^k$ so that q_0 is contained in $c_k - \frac{\varepsilon}{16}$ of the faces of $\Delta_n(k+1)$ embedded via ρ.

Let $P_1, P_2, \ldots, P_{K'}$ be the partition we obtain from Theorem 9.14 with respect to ρ and the complete complex. Let $H = \left\{ \{P_{i_1}, P_{i_2}, \ldots, P_{i_{k+1}}\} \mid \forall s \in E(P_{i_1}, P_{i_2}, \ldots, P_{i_{k+1}}), q_0 \in conv(\rho(s)) \right\}$. Here

---

[17]One can verify as in [Fox+12, Theorem 1.7] that $K' = \max\{K, 16(d+1)/\varepsilon\}$ for large enough n.

$E(P_{i_1}, P_{i_2}, \ldots, P_{i_{k+1}})$ are the faces $s \in X(k)$ such that for every $j = 1, \ldots, k+1$, $\left|s \cap P_{i_j}\right| = 1$. Recall that at most an $\frac{\varepsilon}{16}$-fraction of the tuples are not homogeneous and that at most $\frac{\varepsilon}{16}$ of the faces in $\Delta_n(k+1)$ touch a part in this partition more than once. Thus,

$$\sum_{(P_{i_j})_{j=1}^{k+1} \in H} \prod_{j=1}^{k+1} \mathbb{P}\left[P_{i_j}\right] \geqslant c_k - \frac{3\varepsilon}{16}. \tag{9.2}$$

On the other hand,

$$\mathbb{P}_{s \in X(k+1)}\left[q_0 \in conv(\rho(s))\right] \geqslant \sum_{\{P_{i_1}, P_{i_1}, \ldots, P_{i_{k+1}}\} \in H} \mathbb{P}_{s \in X(k+1)}\left[E(P_{i_1}, P_{i_2}, \ldots, P_{i_{k+1}})\right]. \tag{9.3}$$

By Proposition 7.2 (applied for indicators of the corresponding functions) for small enough $\lambda = \lambda(K', \varepsilon, z_0)$, for every $k+1$ parts $P_{i_1}, P_{i_2}, \ldots, P_{i_{k+1}}$ of size at least $\frac{1}{K'} - z_0$,

$$\mathbb{P}_{\{v_1, v_2, \ldots, v_{k+1}\} \in X(k+1)}\left[\forall j = 1, \ldots, k+1 \; v_j \in P_{i_j}\right] \geqslant (1 - \frac{\varepsilon}{2}) \prod_{j=1}^{k+1} \mathbb{P}\left[P_{i_j}\right].$$

Combining (9.2) and (9.3) gives $\mathbb{P}_{s \in X(k+1)}\left[q_0 \in \rho(s)\right] \geqslant (1 - \frac{\varepsilon}{2})(c_k - \frac{\varepsilon}{4}) \geqslant c_k - \varepsilon$ as desired.

Partite Case. The argument is similar. Let $K = K(\varepsilon/16, k)$ be the constant promised in Theorem 9.14. Let $K' \geqslant K$ be the smallest constant so that for any equipartition over K' parts, the fraction of faces in the complete partite complex that touch every part at most once is at least $1 - \frac{\varepsilon}{16}$.

Let $z = \frac{1}{2K'}$ be as before. As before we first identify the points of X with the complete partite complex with the same number of vertices in every side. Then we find a point q_0 such that $c_{d,k}^p - \frac{\varepsilon}{16}$ of the faces of the complete partite complex contain q_0. We find an equipartition $P_1, P_2, \ldots, P_{K'}$ of $\rho(X(1))$ to K' parts such that $1 - \frac{\varepsilon}{16}$ of the parts are homogeneous with respect to q_0.

We partition further every P_i according to the colors of the vertices to P_i^j for $j = 1, 2, \ldots, d$. Then we remove from the partition every subset P_i^j whose relative size is smaller than $\frac{\varepsilon}{16K' \cdot d^2}$. We note that by taking only the remaining sets in the partition the fraction of vertices we removed from every color is no more than $\frac{\varepsilon}{8}$ vertices in any side. This is because for every P_i, colors j that have that $\mathbb{P}\left[P_i^j\right] \leqslant \frac{\varepsilon}{16K' \cdot d^2}$ can account for no more than $\frac{\varepsilon}{16K' \cdot d}$ out of at least $\frac{1}{K'} - z = \frac{1}{2K'}$. Thus we have not removed more than $\frac{\varepsilon}{8d}$ vertices in total, or $\frac{\varepsilon}{8}$-fraction out of each color.

We now have subsets P_i^j that such that $\frac{\varepsilon}{16K'd^2} \leqslant \mathbb{P}\left[P_i^j\right] \leqslant \frac{3}{2K'}$. The proof now follows as in the two-sided case. $\square$

9.4 Separating MLSI from Reverse Hypercontractivity

As discussed in Section 1, prior techniques for establishing reverse hypercontractivity relied on tensorization or bounded (Modified) Log-Sobolev constant (MLSI). To our knowledge, all known reverse-hypercontractive objects prior to this work also have bounded MLSI. In this section, we give the first separation between these analytic notions in the case where the leading hypercontractive constant C may be greater than 1 (or, alternatively, for the indicator variant). To start, recall the definition of MLSI.

Definition 9.15 (Modified Log-Sobolev Inequality). Let M be a reversible markov chain with stationary distribution π_s. The modified log-sobolev constant of M is:

$$\rho_{\text{MLSI}} = \inf_{f > 0} \frac{\langle f, (I - M) \log f \rangle}{2\text{Ent}_\pi(f^2)}$$

MLSI is classically used to bound the *mixing time* of its associated Markov chain,

$$T_{\text{mix}}(M, \varepsilon) = \min_t \max_{\pi_{init}} \|\pi_{init}^T M^t - \pi_s\|_{TV} \leqslant \varepsilon,$$

due to the following upper bound:

Fact 9.16. *The mixing time of a reversible Markov chain M with stationary distribution π is at most:*

$$T_{mix}(M, \varepsilon) \leqslant \frac{1}{\rho_{MLSI}} \log \left(\frac{\log \pi_*^{-1}}{\varepsilon} \right),$$

where π_* is the minimum stationary probability

High dimensional expanders (at least in the bounded degree setting), are *slow-mixing*, so they cannot have small MLSI constant. In particular, applying this fact to the Ramanujan complexes gives an infinite family of reverse-hypercontractive operators with at least inverse logarithmic dependence on the domain size.

Corollary 9.17. *There exist constants $\ell > 1$ and $C > 0$ and an infinite family $\{T_\rho^{(n)}\}$ of $(\frac{1}{\ell}, \frac{\ell}{1-\ell}, C)$-reverse hypercontractive operators with at most inverse logarithmic MLSI:*

$$\rho_{MLSI} \leqslant C_2 \frac{\log^3 \log(\pi_*)}{\log \pi_*}$$

Proof. Set $\rho = \frac{1}{2}$, let c be as in Theorem 6.3, and let $\{X^{(n)}\}$ be the family of $k = k(n)$-uniform cut-off coset complexes promised in Theorem 9.4 for $k = \log \log n$ and $\lambda = 2^{-k}$, and recall that

1. The degree of every vertex is at most $2^{C_2 k^3}$ for some universal constant $C_2 > 0$

2. The stationary distribution of $T_{1/2}$ on X is uniform

By Theorem 6.3, there exist (universal) constants C, ℓ such that the noise operator $T_{1/2}$ is $(\frac{1}{\ell}, \frac{\ell}{1-\ell}, C)$-reverse hypercontractive for every complex in the family. On the other hand, we will argue that:

$$T_{mix} \left(T_\rho, \frac{1}{2} \right) \geqslant \min \left\{ \frac{2^k}{4}, \frac{C_2' \log(n)}{k^3} \right\},$$

which combined with Fact 9.16 completes the result.

Define the ε-support of a distribution

$$\operatorname{supp}_\varepsilon^\geqslant(\pi) := \{x : \pi(x) \geqslant \varepsilon\}.$$

Setting $\varepsilon = \frac{1}{4n}$, observe that any distribution π with $|\operatorname{supp}_\varepsilon^\geqslant(\pi)| < \frac{n}{4}$ has $\|\pi - \pi_s\|_{TV} \geqslant \frac{1}{2}$. We will argue this holds for $\pi_{\text{init}} T_{\frac{1}{2}}^t$ whenever π_{init} is any indicator and t is sufficiently small.

To see this, we separate out the bounded-degree and stationary components of $T_{\frac{1}{2}}$ as:

$$T_{\frac{1}{2}}^+ := T_{\frac{1}{2}} - 2^{-k}\Pi$$

where Π is the stationary operator, and note that $T_{\frac{1}{2}} = T_{\frac{1}{2}}^+ + 2^{-k}\Pi$. Observe that after t steps, we can write any initial distribution π as:

$$\pi^T T_{\frac{1}{2}}^t = \pi^T (T_{\frac{1}{2}}^+ + 2^{-k}\Pi)^t = \pi^T (T_{\frac{1}{2}}^+)^t + (1 - (1 - 2^{-k})^t)\pi_s$$

For large enough k, as long as $t \leqslant 2^{-k/4}$, the latter term contributes strictly less than $\frac{1}{4}$ to each coordinate. On the other hand for π an indicator, the first term contributes positive weight to at most $deg(T_{\frac{1}{2}}^+)^t$ coordinates, so setting t such that $deg(T_{\frac{1}{2}}^+)^t \leqslant \frac{n}{4}$ completes the result. $\square$

9.5 It Ain't Over Till It's Over

Friedgut and Kalai's 'It Ain't Over Till It's Over' Theorem is a classical result in social choice which states that even after taking a significant random restriction, any balanced function still has some uncertainty with high probability (in other words, it is a tail bound on the random restriction operator). The precursor to this Theorem was a version for the noise operator proved in [MOO05, Appendix C]. We prove the result holds for any LUS, which implies the same for any c-nice complex.

Theorem 9.18. *For any $\rho \in (0,1)$ and $\tau > 0$, let X be a d-uniform τ-LUS for d sufficiently large. Then there exists a constant $q = q(\rho, \tau)$ such that for any $\delta > 0$ and $f : X(d) \to [0,1]$ with density $\mu = \mathbb{E}[f]$:*

$$\mathbb{P}[T_\rho f > 1 - \delta] \leqslant c_{\mu,1}\delta^q$$

where $c_{\mu,1} = \left(\frac{8(1+\mu)^{q+1}}{(1-\mu)^{q+2}} \right)^{\frac{1}{q}}$, and

$$\mathbb{P}[T_\rho f < \delta] \leqslant c_{\mu,2}\delta^q$$

where $c_{\mu,2} = \left(\frac{8(2-\mu)^{q+1}}{\mu^{q+2}} \right)^{\frac{1}{q}}$.

Proof. We prove the former. The latter follows from considering $1 - f$. Let T of the set of elements whose neighborhoods are dense in f:

$$T := \{x : T_\rho f(x) \geqslant 1 - \delta\},$$

and consider the indicator cut-off $h = \mathbf{1}_{f \leqslant \frac{1+\mu}{2}}$. We will argue that on the one hand the correlation of T and h cannot be large, since when $x \in T$, $y \sim_\rho x$ must be mostly above the cutoff. On the other hand, the correlation is at least some power of $\mu(T)$ by reverse hypercontractivity, so $\mu(T)$ must be small.

Formally, observe that for any $x \in T$, Markov's inequality gives that $T_\rho h(x) \leqslant \frac{2\delta}{1-\mu}$ so

$$\mathbb{E}[1_T T_\rho h] \leqslant \mu(T)\delta\frac{2}{1-\mu}.$$

On the other hand, we also have by Markov that $\mathbb{E}[h] \geqslant \frac{1-\mu}{1+\mu}$, so reverse hypercontractivity for indicators (Theorem 6.4) implies there exists some ℓ such that:

$$\mathbb{E}[1_T T_\rho h] \geqslant \frac{1}{4} \left(\mu(T)\frac{1-\mu}{1+\mu} \right)^\ell$$

Combining the equations and setting $q = \ell - 1$ gives the desired result. $\qquad\square$

We note we have made no attempt to optimize the constants, and it may be possible to improve dependence on μ. [MOO05] combine a variant of Theorem 9.18 with the invariance principle to prove the classical form of the conjecture with respect to random restrictions. It is not clear what the correct form of this conjecture should be for HDX, or even for product spaces.[18] We leave this as an open question.

9.6 A Frankl–Rödl Theorem

The Frankl–Rödl Theorem [FR87] is a powerful result from extremal combinatorics that states that the independence number of the graph whose vertices are $\{0,1\}^n$ and whose edges are given by fixed weight intersection k (typically for some $k = \Theta(n)$) is at most $\exp(-\Omega(k))|V|$. Benabbas, Hatami, and Magen [Ben12] proved a variant of this Theorem via reverse hypercontractivity on the cube, later leading to several applications in hardness of approximation [Kau+14]. Their method, which also bounds classical properties such as the chromatic and dominating set numbers, is based on the following claim:

Claim 9.19. Let $p_0, q_0 > 0$ and let $G = (V, E)$ be a weighted graph with minimum vertex weight q_0 such that for every $A, B \subset V$ of relative size at least p_0:

$$\mathbb{P}_{(s_1,s_2)\sim E}[s_1 \in A, s_2 \in B] > 0.$$

Then the maximal independent set in G has size at most $2p_0 + q_0$, the minimal dominating set has relative size at least $1 - (2p_0 + q_0)$, and the chromatic number is at least $\frac{1}{2p_0+q_0}$.

---

[18]We remark it should be possible to give a version of the result for suitably 'global' functions via tools of [Kee+21; GLL22; Baf+22b], but it is not clear such a notion would have significant qualitative meaning

As a corollary we have that the down-up walks in any of the discussed LUS-complexes in Section 6.3 have maximal independent sets of relative size at most $\exp(-\Omega(k))$, a chromatic number of at least $\exp(\Omega(k))$ and minimal dominating sets of size at least $1 - \exp(-\Omega(k))$. As discussed in the introduction, we remark that these graphs are not quite the correct analog of Frankl–Rödl, since they have edges with intersection up to some fixed γk instead of γk exactly. Benabbas, Hatami and Magen [Ben12] handle this by analyzing the closeness of the exact intersection versus noise operator on the cube, using tools from Fourier analysis. It would be interesting to see if a variant of this result holds on HDX. If this could be shown generally, then our reverse hypercontractivity theorem would recover Frankl–Rödl up to constants in the exponent.

Proof of Claim 9.19. We prove the statement for independent sets. The chromatic number statement follows since every coloring is a partition to independent sets, and the statement for dominating sets follows since the complement of any dominating set is an independent set. Let $I \subseteq G$ be an independent set and assume toward contradiction that $\mathbb{P}[I] \geqslant 2p_0 + q_0$. We partition I into I_1, I_2 such that $\mathbb{P}[I_1], \mathbb{P}[I_2] \geqslant p_0$ (we can do so by greedily adding vertices to I_1 until its probability is between p_0 and $p_0 + q_0$). Then by assumption $\mathbb{P}_{(s_1,s_2)\sim E}[s_1 \in I_1, s_2 \in I_2] > 0$, which gives the desired contradiction. $\qquad\square$

Remark 9.20. Although the graphs we discussed in this paper have self loops. The claim above continues to hold after removing these self loops, as is evident from the proof.

10 Codes and Splitting Trees

Error correcting codes, and in particular the powerful notions of *list-decoding* and *local-testability*, are among the earliest successful applications of inclusion samplers [Imp+08; IJK09; IKW12] and HDX [Din+18a; Ale+20; JST21; PK22; Din+22a]. In this section we explore the implications of our tools and related ideas over the weaker family of 'splitting trees' in this classical setting. Our main application is the first construction of constant rate codes over large alphabets which simultaneously have 1) near-optimal distance 2) list-decodability and 3) local-testability. At the end of the section we introduce a (conjectural) HDX-based approach toward 'lossless' distance-amplification of LTCs that experiences no decay in soundness.

10.1 Splitting Trees

We start with a fairly substantial detour into the world of splitting trees, which are a weakening of high dimensional expanders introduced in [AJT19] which only requires certain patterns of swap walks to expand. We will prove a high-dimensional expander mixing lemma and some basic sampling properties for general splitting trees that will be useful for our applications to coding theory.

10.1.1 Splitting Preliminaries

Given a binary tree T, let $\mathcal{L}(T)$ denote T's leaves and $\mathcal{I}(T)$ its internal nodes. We drop T from the notation when clear from context. For every internal node $v \in \mathcal{I}$, denote the left and right children of v by ℓ_u and r_u respectively.

Definition 10.1 (Ordered Binary Tree). A d-uniform ordered binary tree is a pair (T, ρ) where T is a rooted binary tree with d leaves, and $\rho : T \to [d]$ is a labeling such that:

1. The label of every leaf $u \in \mathcal{L}$ is $\rho(u) = 1$.

2. For every interior $u \in \mathcal{I}$, it holds that $\rho(u) = \rho(\ell_u) + \rho(r_u)$.

We call such a ρ well-ordered with respect to T.

A complex is called splittable if it can be decomposed by an ordered tree where every non-leaf node corresponds to an expanding swap walk.

Definition 10.2 (Splitting tree). Let $\gamma > 0$. A d-uniform γ-*splitting tree* is a triple (X, T, ρ) where X is a d-uniform simplicial complex and (T, ρ) is an ordered binary tree such that for every $u \in \mathcal{I}(T)$:

$$\lambda(S_{\rho(\ell_u),\rho(r_u)}) \leqslant \gamma,$$

We will also work with the corresponding partite notion of splitting trees.

Definition 10.3 (Partite Ordered Binary Tree). A d-uniform partite ordered binary tree is a pair (T, ρ) where T is a rooted binary tree with d leaves, and $\rho : T \to \mathbb{P}\{[d]\} \setminus \{\emptyset\}$ is a labeling such that:

1. $\rho|_{\mathcal{L}}$ is a bijection to the singletons of $[d]$.

2. For every interior $u \in \mathcal{I}$, it holds that $\rho(u) = \rho(\ell_u) \uplus \rho(r_u)$.

We remark that when convenient, we may use a general set S of size $d + 1$ as the domain of ρ (generally corresponding to a labeling of the "parts" of an associated d-uniform partite complex).

Definition 10.4 (Tuple splitting tree). Let $\gamma > 0$. A d-uniform γ-*tuple splitting tree* is a triple (X, T, ρ) where X is a d-uniform simplicial complex and (T, ρ) is an ordered partite binary tree such that for every non-leaf vertex $u \in T$:

$$\lambda(S_{\rho(\ell_u), \rho(r_u)}) \leqslant \gamma$$

Since all swap walks on two-sided HDX expand, these give rise to a basic family of splitting trees.

Example 10.5 ([AJT19; DD19; GLL22]). Let $\lambda \geqslant 0$ and let X be a d-uniform two-sided λ-high dimensional expander. Then for any binary tree T and labeling ρ that satisfies the first and second item in Definition 10.2, (X, T, ρ) is a $d\lambda$-splitting tree.

An easy consequence of [DD19] is that partite one-sided HDX also give rise to tuple-splitting trees.

Example 10.6. Let $\lambda \geqslant 0$ and let X be a d-uniform two-sided λ-high dimensional expander. Then for any binary tree T and labeling ρ that satisfy the first and second item in Definition 10.4, (X, T, ρ) is a $d^2\lambda$-tuple splitting tree.

Finally, [Ale+20, Corollary 9.18] also prove that walks on expander graphs are splittable. Let us define the complex formally. Let $G = (V, E)$ be a graph and $k \in \mathbb{N}$. Let W_G be the following k-partite simplicial complex

$$W_G(1) = V \times [k],$$

$$W_G(k) = \{\{(v_1, 1), (v_2, 2), \ldots, (v_k, k)\} \mid (v_1, v_2, \ldots, v_k) \text{ is a walk in } G\}.$$

We choose a face in $W_G(k)$ by choosing a vertex $v_1 \in V$ according to the stationary distribution over vertices in G and then taking a $(k - 1)$-length random walk on that vertex.

Example 10.7. Let G be a λ-expander and T, ρ an ordered partite tree. If every $u \in \mathcal{I}$ satisfies:

$$\max \rho(\ell_u) < \min \rho(r_u),$$

then (W_G, T, ρ) is a λ-splittable tree.

10.1.2 Operations on splitting trees

We will need several further operations on splitting trees for our analysis: a *pruning* method that allow us to analyze higher dimensionalT faces, and a *partitification* technique that allows us to reduce standard splitting trees to the tuple setting.

We start with the latter. Recall X's partitification is the complex

$$P(k) := \left\{ (s, \pi) := \left\{ (s_{\pi(1)}, 1), (s_{\pi(2)}, 2), \ldots, (s_{\pi(k)}, k) \right\} \mid s \in X(k), \pi \in S_k \right\},$$

This notion extends naturally to splitting trees.

Definition 10.8. Let (X, T, ρ) be a d-uniform λ-splitting tree, and $\rho' : T \to [d]$ a well-ordered labeling such that $|\rho'(u)| = \rho(u)$ for every $u \in T$. The ρ'-*partitification* of (X, T, ρ) is defined to be the tuple splitting tree (P, T, ρ').

Any un-ordered splitting tree can be embed into a tuple splitting tree by taking any partitification.

Claim 10.9. Let (X, T, ρ) be a λ-splitting tree. Every ρ'-partitification of (X, T, ρ) is a λ-tuple splitting tree.

Proof. It suffices to show for every internal node $u \in \mathcal{I}$ with $\rho'(\ell(u)) = A$ and $\rho'(r(u)) = B$ that $S_{A,B}(P)$ is a λ-bipartite expander. This follows from the observation that $S_{A,B}(P)$ is isomorphic to the bipartite graph with left vertices $L = X(|A|) \times [|A|!]$, right vertices $R = X(|B|) \times [|B|!]$, and edges $(s, i) \to (t, j)$ for all $i \in A, j \in B$, and (s, t) such that $s \sim t$ in $S_{\rho(\ell_u), \rho(r_u)}$.[19] It is a standard fact that such a bipartite construction inherits its expansion from $S_{\rho(\ell_u), \rho(r_u)}$, and is therefore a λ-bipartite expander. $\qquad\square$

Finally, we note that tuple splitting trees can be pruned to create new splitting trees. Let (X, T, ρ) be a tuple splitting tree with k leaves. Let $u \in \mathcal{I}$ be an internal node whose label is $F = \rho(u)$. Let $P = P(X, F)$ be the partite complex with vertices $P(1) = X^{[k] \setminus F}(1) \uplus X[F]$ and top level faces $t \cup \{s\}$ such that $t \cup s \in X(k)$, inheriting the corresponding densities. In other words, we replace the vertices of X in $X^F(1)$ with the set of faces $X[F]$.

Let $T(u)$ be the sub-tree of T rooted by u. Let $T' = (T \setminus T(u)) \cup \{u\}$. Let $\rho' : T' \to ([k] \setminus F) \cup \{F\}$ be the labeling that replaces the subset F with a single replacement index f:

$$\rho'(v) = \begin{cases} (\rho(v) \setminus F) \cup \{f\} & F \subseteq \rho(v) \\ \rho(v) & otherwise \end{cases}.$$

Definition 10.10 (Pruning Trees). Let (X, T, ρ) be a λ-tuple splitting tree with k leaves. Let $u \in T$ be a non-leaf node whose label is $F = \rho(u)$. The u-pruning of the tree is the triple $(P(X, F), T', \rho')$.

Claim 10.11. Let (X, T, ρ) be a λ-tuple splitting tree. Let $u \in T$ be a non-leaf node. Then the u-pruning is also a λ-splitting tree. $\quad\square$

The proof follows directly from the definitions and we therefore omit it.

Finally, it will be convenient to introduce notation for *sequential pruning*, that is given a collection of disjoint color sets $\mathcal{F} = \{F_1, \ldots, F_m\}$, we denote the sequential pruning of X by $\mathcal{F}$ as $P(X, \mathcal{F}) :=$ $P(\ldots P(X, F_1), F_2), F_m)$. Note that the resulting complex is independent of the order of $\mathcal{F}$, so $P(X, \mathcal{F})$ is well-defined. One can similarly define a sequentially pruned splitting tree as above. It is immediate from Claim 10.11 that these are also λ-tuple splitting.

10.1.3 Hitting set and Bias amplification

In this section, we prove two basic 'sampling-type' properties for splitting trees: hitting set and bias amplification. We start with the former. Recall a complex is (γ, i)-hitting if for any $A \subset X(0)$:

$$\mathbb{P}_{\sigma \in X(i)}[\sigma \subset A] \leqslant \mu(A)^i + \gamma$$

Proposition 10.12 (Hitting Set). *Any depth D λ-(tuple) splitting tree with k leaves is $(3^D \lambda, k)$-hitting.*

Splitting trees also exhibit strong bias amplification, used to build balanced error correcting codes.

Proposition 10.13 (Bias-Amplification). *Let (X, T, ρ) be a λ-tuple splitting tree for $\lambda < \frac{1}{16}$. For any $0 < \varepsilon < \frac{1}{4}$ and family of mean ε functions $\{f_i : X[i] \to \{\pm 1\}\}_{i \in [k]}$, $\prod f_i$ is a $\{\pm 1\}$-valued function with bias at most:*

$$\left| \mathbb{E}_{a \in X(k)} \left[\prod_{i=1}^k f_i(a_i) \right] \right| \leqslant \varepsilon^k + 2\lambda \tag{10.1}$$

While Proposition 10.12 and Proposition 10.13 are elementary, to the authors' knowledge they do not appear in the literature. Prior works on HDX (e.g. [DK17; Ale+20]) use weaker variants of these lemmas which require super-exponential overhead to achieve a given distance or bias, while the above only requires quasi-polynomial overhead. We discuss this in greater detail in Section 10.

These results are immediate corollaries of a functional variant of [DD19]'s "high dimensional expander-mixing lemma" (HD-EML) extended to splitting trees. We give the statement here and include a proof in the appendix for completeness along with the proofs of Proposition 10.12 and Proposition 10.13.

---

[19]Formally, the edge $(s, i) \to (t, j)$ also inherits the weight $\frac{\mathbb{P}(st)}{|A||B|}$ from $S_{\rho(\ell_u), \rho(r_u)}$.

Theorem 10.14 (high dimensional expander Mixing Lemma). *Let (X, T, ρ) be a depth D λ-tuple splitting tree with k leaves. Denote by d_i the depth of the leaf labeled i. Then for any family of functions $\{f_i : X[i] \to \mathbb{R}\}_{i \in [k]}$:*

$$\left| \underset{a \in X(k)}{\mathbb{E}} \left[\prod_{i=1}^{k} f_i(a_i) \right] - \prod_{i=1}^{k} \underset{a_i \in X[i]}{\mathbb{E}} [f_i(a_i)] \right| \leqslant 3^D \lambda \prod_{i=1}^{k} \|f_i\|_{2^{d_i}}. \tag{10.2}$$

If (X, T, ρ) is instead a depth D standard λ-splitting tree, we take $\{f_i : X(1) \to \mathbb{R}\}_{i \in [k]}$ and have:

$$\left| \underset{a \in X(k), \pi \in S_k}{\mathbb{E}} \left[\prod_{i=1}^{k} f_i(a_{\pi(i)}) \right] - \prod_{i=1}^{k} \underset{a_i \in X(1)}{\mathbb{E}} [f_i(a_i)] \right| \leqslant 3^D \lambda \prod_{i=1}^{k} \|f_i\|_{2^{d_i}}. \tag{10.3}$$

In the future, we state our results only for the partite setting unless there is a substantial difference in proof or parameters for the results. Otherwise, we collect analog statements for the unordered case at the end of Appendix D for completeness.

We close the subsection with a few remarks. First it should be noted that for the special case of expander walks, better hitting set and bias amplification methods are known (indeed one can achieve polynomial overhead vs our quasipolynomial bound above). Nevertheless, it is useful to have these lemmas on structures like HDX that admit stronger properties like agreement tests. We remark it is also actually possible to use the high dimensional expander mixing lemma for splitting trees directly to prove a different variant of Chernoff-Hoeffding by reduction to the complete complex. However, this results in exponentially worse parameters than our arguments in Section 5, and we will not need the result in the following.

10.2 Coding Preliminaries

A q-ary error correcting code C over a finite size-q alphabet Σ is a subset $C \subset \Sigma^n$, where $n \in \mathbb{N}$ is called the *block length*. We say a code is *linear* if $\Sigma = \mathbb{F}_p$ for prime power p and C is a subspace, and $\mathbb{F}_p$-*linear* if $\Sigma = \mathbb{F}_p^k$ for some $k \in \mathbb{N}$ and C is a subspace of $(\mathbb{F}_p^k)^n$.

The *rate* of a code C, which measures its overhead, is $r = \frac{\log_{|\Sigma|}(|C|)}{n}$, and its *distance*, which captures the codes error correction capability, is the minimum normalized hamming distance between any two codewords

$$d := \min_{c_1, c_2 \in C} \left\{ \frac{|\{i \in [n] : c_1(i) \neq c_2(i)\}|}{n} \right\}.$$

We remark that for any $\mathbb{F}_p$-linear code, the distance is the minimum weight of a non-zero codeword. For ease of notation, we will call a q-ary code with rate r and distance d an $(r, d)_q$-code.

We will be interested in so-called "good" families of codes, which are infinite families with growing block length and constant rate and distance. We will focus in particular in this section on building large distance $\mathbb{F}_2$-linear codes over $\Sigma = \mathbb{F}_2^k$. In this regime there is a classical upper bound due to McEliece, Rodemich, Rumsey, and Welch [McE+77] stating that good q-ary codes have distance at most $1 - \frac{1}{q}$, or more exactly:

Theorem 10.15 (MRRW Bound [McE+77]). *For all $\varepsilon > 0$ and $q \in \mathbb{N}$, any family of q-ary codes with distance $1 - \frac{1}{q} - \varepsilon$ has rate at most:*

$$r \leqslant O\left(\varepsilon^2 \log(1/\varepsilon) \right).$$

Given a code C, a *unique decoding algorithm* with radius $\gamma \in [0, 1/2]$ is a (possibly randomized) algorithm $Dec : \Sigma^n \to C$ which given any $w \in \Sigma^n$ such that $dist(w, C) \leqslant \gamma$ outputs the uniquely closest $y \in C$ to w. Note that unique decoding is only possible up to radius $\lfloor \frac{nd-1}{2n} \rfloor \leqslant 1/2$, even with large alphabet, as there may be no unique closest codeword beyond this point.

List-decoding is the natural extension of unique decoding beyond the $\frac{1}{2}$ barrier. A code is said to be list-decodable with radius t if there is an algorithm which outputs a *list* of all codewords within radius t of the given word. We first give the combinatorial definition, which simply bounds the number of codewords around any fixed point.

Definition 10.16 (Combinatorial List-decoding). *For $\gamma \in (0, 1)$ and $L \in \mathbb{N}$, a code C is said to be (γ, L) list-decodable if for every $z \in Y$, there are at most L codewords $x \in C$ such that $dist(x, z) < \gamma$.*

In this section, we will be interested in *efficient algorithms* for list-decoding, typically a much more challenging task than the existential definition.

Definition 10.17 (Efficient List-decoding)**.** For $\delta > 0$, $\gamma \in (0,1)$ and $L \in \mathbb{N}$, a family of codes $\{C_n\}$ is said to be efficiently (γ, L) list-decodable with confidence δ if it is combinatorially (γ, L) list-decodable and there exists a (randomized) polynomial time[20] algorithm DEC: $\Sigma^n \to P(C)$ such that for every word $w \in \Sigma^n$:

$$\text{DEC}(w) = \{c \in C : \text{dist}(c, w) < \gamma\}$$

except with probability δ.

Another classical approach to handling the high error regime is *local testability.* A code C is said to be locally testable if there is a (randomized, non-adaptive) algorithm $\mathcal{T}_C$ which on input of a word $w \in \Sigma^n$, queries constantly many symbols of w and rejects with high probability if it is far from the code. In practice, if one receives such a word, instead of trying to list decode they could simply request re-transmission. More formally, we consider the following standard notion of 'strong' local testability:

Definition 10.18 (Locally Testable Code)**.** For any $s > 0$ and $q \in \mathbb{N}$, a code C of blocklength n is said to be (q, s)-locally testable if there exists a function $\mathcal{T}_C : [n]^q \times \Sigma^q \to \{0, 1\}$ and a distribution $\mathcal{D}$ over $[n]^q$ such that on any word $w \in \Sigma^n$, the pair $(\mathcal{T}_C, \mathcal{D})$ (called the tester) satisfies:

1. **Completeness:** If $w \in C$, then $\mathbb{P}_{J \sim \mathcal{D}}[\mathcal{T}_C(J, w_J) = 1] = 1$

2. **Soundness:** $\mathbb{P}_{J \sim \mathcal{D}}[\mathcal{T}_C(J, w_J) = 0] \geqslant s \cdot \text{dist}(w, C)$

We call q the locality, and s the soundness.

Typically we will simply think of $\mathcal{T}_C$ as a randomized algorithm and write $\mathcal{T}_C(w)$ to denote its application on the word w, dropping the distribution of the tester when clear from context. We will work with testers that have one additional constraint: their queries should be marginally uniform.

Definition 10.19 (Uniform LTC)**.** $(\mathcal{T}, \mathcal{D})$ is called *uniform* if every marginal of $\mathcal{D}$ is uniform over $[n]$.

We remark that focusing on uniform LTCs is not particularly restrictive—since codes do not typically have 'preferred' coordinates, most natural constructions are uniform.

Finally, though not strictly necessary, we will assume throughout that all complexes in this section are "homogenous" subsets $X \subset [n]^k$, meaning that their marginal distribution over each part is uniform over $[n]$. We remark that in the context of high dimensional expanders, any partite 'type-regular' complex can be homogenized in this fashion while maintaining expansion and bounded degree (see [FI20] for details). As most known constructions of HDX are type-regular, this is not a particularly restrictive assumption, and it is also known to hold for other classical splitting trees such as expander walks [Ale+20].

10.3 The ABNNR Construction

In the previous section, we proved several 'agreement theorems' on simplicial complexes. These well-studied tests are actually very closely related to both local testability and list decoding of a particular family of codes called *direct product codes*, corresponding exactly to the set of 'global functions' discussed in Section 8. At a broader level, these are themselves a special case of a well-studied family of codes introduced in the seminal work Alon, Bruck, Naor, Naor and Roth [Alo+92] called the 'ABNNR'-construction. We focus on the case of $\mathbb{F}_2$ for simplicity, though our results extend naturally to larger fields.

Definition 10.20 (the ABNNR Construction)**.** Given a right k-regular bipartite graph $G = (L, R, E)$ and a base code $C \subseteq \mathbb{F}_2^L$, the "ABNNR-Encoding" of C with respect to G is the code $\text{Im}(E_G(C))$ where E_G is an "encoding" map $E_G : \mathbb{F}_2^{|L|} \to (\mathbb{F}_2^k)^{|R|}$ defined by setting for each $c \in C$ and $v \in R$:

$$E_G(c)_v = c|_{N(v)}.$$

---

[20]Here we mean polynomial in the block length.

In other words, each right vertex concatenates the symbols of its neighbors to create a new code over R. One can check that in Section 8, the set of global functions is exactly the ABNNR encoding over the inclusion graph $(X(j), X(k))$. For a general k-uniform complex X and $j \leq i \leq k$, we write $E_X^{(j,i)}$ to mean the ABNNR-encoding on the vertex inclusion graph $(X(i), X(j))$. In the special case where $i = k$ and $j = 1$, we write just E_X to denote the ABNNR-encoding over the inclusion graph $([n], X(k))$, where inclusion is viewed taking $X(k) \subseteq [n]^k$ as 'ordered tuples' of $[n]$, rather than as faces of a simplicial complex. This is a minor difference, but is a bit more convenient in the setting of amplifying base codes.

Recent years have seen a great deal of work on instantiating the ABNNR construction on high dimensional expanders and splitting trees [DK17; Din+18a; Ale+20; JST21; BD22; JST23], in particular in the context of efficient list-decoding. Building on tools from these works, we give the first family of large alphabet codes with near-optimal distance that are simultaneously locally testable and list-decodable.

Theorem 10.21. *For all large enough $k \in \mathbb{N}$ and all $\varepsilon > 0$, there exists an explicit family of $\mathbb{F}_2$-linear ABNNR-Codes that are simultaneously:*

1. $\left(\varepsilon^{O(k)}, 1 - 2^{-k} - \varepsilon\right)_{2^k}$-*codes*

2. $\left(O(1), \frac{O(1)}{k\log(1/\varepsilon)}\right)$-*locally testable.*

3. $\left(1 - 2^{-\Omega(k)}, 2^{O(k)}\right)$-*efficiently list-decodable with confidence $2^{-\Omega(n-k)}$ whenever $\varepsilon \leq 2^{-\Omega(k)}$.*

We remark that very recently, Jeronimo, Srivastava, and Tulsiani [JST23] gave a direct list-decoding algorithm for the LTCs of [Din+22a] (which have some small constant distance $d \ll \frac{1}{2}$). To our knowledge, these are the only other examples of codes that are simultaneously c^3-LTC and efficiently list-decodable.

The remainder if this section is split into three parts focused respectively on distance amplification, local-testability, and list-decoding of the ABNNR construction, and a short fourth section devoted to a conjectural approach to removing the soundness decay in Theorem 10.21. Each section is independent, containing both the corresponding component of Theorem 10.21 as well as more general results within the framework. To facilitate this structure we first give a brief proof overview of Theorem 10.21 that doubles as a roadmap for the section.

Proof Overview. The high level idea behind Theorem 10.21 is simple: starting with a binary locally testable 'base code' C with distance near $\frac{1}{2}$, we will argue that the ABNNR-encoding of C on any k-partite splitting tree[21] has distance near $1 - 2^{-k}$ (Corollary 10.24), is list-decodable (Theorem 10.32), and maintains local testability (Theorem 10.25). Instantiating this framework on *expander walks* (the sparsest known family of splitting trees) gives codes with the claimed parameters. We give a brief description of each step.

Distance amplification on splitting trees is fairly elementary, and follows from the hitting-set lemma. This is a standard argument: one observes that the two encoded symbols $E(x), E(x')$ differ in any face which hits the vertex set $1_{x \neq x'}$. Since by assumption the base code has distance roughly $1/2$, the hitting set lemma promises this set is hit except with roughly 2^{-d} probability (see Corollary 10.24). Theorem 10.21 actually uses a slightly stronger distance amplification lemma for expander-walks and is given in Corollary 10.30. The weaker amplification lemma (in particular its bias variant) appears later in the proof as a sub-component of the list-decoding algorithm.

List-Decoding the ABNNR construction, despite being heavily studied on HDX and splitting trees, is surprisingly tricky. Roughly speaking there are two main issues. First, known list-decoding algorithms for this setting are actually either for an $\mathbb{F}_2$-valued variant of ABNNR called the *direct sum* construction [Ale+20; JST21; BD22], or only hold on *two-sided* HDX [Ale+20; Din+18a] (which have exponentially worse rate). To handle this, we build a reduction from list-decoding the ABNNR-encoding on X to list-decoding the direct sum encodings of its projected sub-complexes X^F. The second issue with this strategy is that known list-decoding algorithms for direct sum [JST21; BD22] actually requires the base code to be ε-*biased*, a stronger condition than simply having distance near $\frac{1}{2}$. We will discuss how to construct such base codes shortly. We instantiate this framework in Corollary 10.35 to prove the list-decoding component of Theorem 10.21.

Local Testability of the ABNNR construction, to our knowledge, has not been studied outside the special case of agreement testing. Our test follows a very simple strategy. We first check whether the given word f

---

[21]Formally for list-decoding actually require a slightly stronger notion called 'complete' splittability, c.f. Definition 10.31.

is a direct product (in the sense of Section 8) via a simple 2-query agreement test: pick a random vertex $v \in [n]$, and independently $s, s' \supset v$, and check whether f_s and $f_{s'}$ match on v. If this test passes, we then simulate the tester on the 'direct product decoding' of f. That is, for every location the base tester queries, we pick a random face including that vertex and feed the tester its value in the word f. We prove this procedure is sound in Theorem 10.25, and instantiate the process in Corollary 10.30 to prove the local testability component of Theorem 10.21.

The Base Code for our construction, as discussed above, must be an ε-biased c^3-LTC. c^3-LTCs with distance near $\frac{1}{2}$ are now known due to the breakthrough works of [Din+22a; PK22] combined with LTC distance-amplification techniques of [Kop+17], but the latter amplification causes several core issues that make the tecnique largely useless for our setting: 1) the resulting codes make $\text{poly}(1/\varepsilon)$ queries, 2) may not be ε-biased, and 3) are heavily randomized. We remedy these issues via the ABNNR-construction itself. In particular, starting with the base LTC of [Din+22a], we encode it into $\log(1/\varepsilon)$-uniform walks on an expander graph using ABNNR, and then concatenate with the Hadamard code. This code is ε-biased roughly because concatenating with the perfectly balanced Hadamard code converts the good distance of the ABNNR-encoding into good bias, and the resulting code is locally testable using the Hadamard's classical local decoding and local testing algorithms. We formalize this step in Proposition 10.29

Finally we summarize the construction and proof below:

Code underlying Theorem 10.21:

1. Start with the base c^3-LTC C of [Din+22a]

2. Amplify C's distance via expander walks

3. Reduce alphabet and bias by concatenating with Hadamard

4. Re-amplify the resulting code via a second set of expander walks.

Proof of Theorem 10.21. Bias, rate, and local-testability are proven in Corollary 10.30. List-decodability is proved in Corollary 10.35. □

10.4 Distance Amplification

The ABNNR construction is one of the most classical distance amplification methods in coding theory. The standard result to this effect is that instantiating the framework on a (bounded degree) additive sampler amplifies distance while maintaining rate up to a constant factor.

Lemma 10.22 ([Alo+92])**.** *Let $G = (L, R, E)$ be a right k-regular bipartite graph such that (R, L, E) is an (α, β)-sampler. If $C \subset \mathbb{F}_2^{|L|}$ is a code of distance α and rate r, then $E_G(C)$ is a $(r\frac{|L|}{|R|}, 1 - \beta)_{2^k}$-code.*

This observation was used in [DK17; Din+18a] to give distance amplification for codes defined on HDX. Our sampling lemmas imply an exponential improvement in the rate-distance trade-off of such codes. We choose a fairly arbitrary setting of our parameters for simplicity, but any setup with good sampling works:

Corollary 10.23 (Distance Amplification on HDX)**.** *For any $\alpha, r, c > 0$, let C be a base code of distance α and rate r. Let X be a d-uniform c-nice complex. Then for any $i < j \leqslant d$ and $G = (X(i), X(j))$ the ABNNR code $E_X^{i,j}(C)$ is an $(r', d')_{|\Sigma|}$-code for*

$$r' = r\frac{|X(i)|}{|X(j)|}, \qquad d' = 1 - \frac{1}{\alpha}\exp\left(\Omega_c\left(\alpha^2\frac{j}{i}\right)\right), \qquad |\Sigma| = 2^{\binom{j}{i}}.$$

Note that when $i = 1$, Corollary 10.23 has alphabet size $|\Sigma| = 2^j$ and distance $1 - \text{poly}(|\Sigma|^{-1})$, optimal up to polynomial factors by Theorem 10.15. However, in this special case, it is actually possible achieve optimal amplification directly via the hitting set lemma (Proposition 10.12).

Corollary 10.24 (Distance Amplification for Splitting Trees). *For any $\alpha, r > 0$, let C be a base code of distance α and rate r. If X is a k-uniform λ-(tuple)-splitting tree of depth D, then $E_X(C)$ is an $(r', d')_{|\Sigma|}$-code for*

$$r' = r\frac{n}{|X(k)|}, \qquad d' = 1 - (1-\alpha)^k - 3^D\lambda, \qquad |\Sigma| = 2^k$$

Moreover when $\alpha = \frac{1}{2}(1-\varepsilon)$, then $d' \geqslant 1 - \frac{1}{2^k}(1+\varepsilon)^k - 3^D\lambda$.

Proof. Rate and alphabet size are immediate from construction. Let $x, x' \in C$ be any two distinct non-zero codewords. We wish to bound the distance between their encodings $E_X(x)$ and $E_X(x')$, which is exactly the hitting probability of $1_{x \neq x'}$ in the inclusion graph $([n], X(k))$. By the distance of C we are promised that $\mu(1_{x=x'}) \leqslant 1 - \alpha$, so Proposition 10.12 gives the desired result.[22] Plugging $\alpha = \frac{1}{2}(1-\varepsilon)$ into the resulting bound gives the final line. $\qquad\square$

10.5 Local Testability of ABNNR

In this section, we prove the ABNNR encoding of a locally testable code is locally testable.

Theorem 10.25 (Local Testability of ABNNR). *For any uniform (q,s)-LTC of blocklength n and (k_L, k_R)-regular bipartite graph $G = (L, R, E)$ on $|L| = n$ vertices, $E_G(C)$ has a uniform $(q+2, \frac{cs}{qk_R})$-local tester.*

Our proof relies on the following elementary agreement test. Given a word f, denote by $\mathcal{T}_G(f)$ the following process:

1. Pick a uniformly random vertex $v \in L$

2. Pick two uniformly random neighbors $w, w' \in N(v)$

3. Accept if and only if $f_w(v) = f_{w'}(v)$

In the agreement notation of the previous section, we'd write the word f as an ensemble $\mathcal{F} = \{f_w : N(w) \to \mathbb{F}_2\}_{w \in R}$ and denote the acceptance probability of this test as $Agree_0(\mathcal{F})$. We will move between these two viewpoints interchangeably. $\mathcal{T}_G$ is clearly complete. Its soundness follows from a general result of [GC19]:

Lemma 10.26 ([GC19, Lemma 5.2]). *There exists a universal constant $c \geqslant 1$ such that for all $\varepsilon > 0$ and any ensemble $\mathcal{F}$:*

$$Agree_0(\mathcal{F}) \geqslant 1 - \varepsilon \implies dist(F, g_{maj}) \leqslant ck_R\varepsilon$$

where for $v \in L$ we define $g_{maj}(v) := plurality_{w \in N(v)}\{f_w(v)\}$.[23]

With this in mind, we prove Theorem 10.25 via the following natural tester: given a word $f \in (\mathbb{F}_2^{k_R})^R$, run the DP-tester on w, and the base code tester on the plurality decoding of f if it passes:

1. Run $\mathcal{T}_G$ on f

2. For every $v \in L$ that $\mathcal{T}_C$ would query:

   – Sample a random neighbor of v: $w \sim N(v)$

   – Denote the 'decoded' plurality as $g_{\text{dec}}(v) = f_w(v)$

3. Run $\mathcal{T}_C$ on g_{dec}

Proof of Theorem 10.25. Query complexity and completeness are immediate from construction. The tester is uniform since the base code is uniform, and picking a uniformly random vertex $v \in L$ and random $w \in N(v)$ is uniformly distributed over R.

Toward analyzing soundness, fix any $\delta \in [0,1]$ and f some word at distance $dist(f, E_X(C)) = \delta$. We need to show the test *rejects* on f with probability at least $\delta c'\frac{s}{qk_R}$ for some $c' \geqslant 0$. Observe that we may assume $\mathcal{T}_G(f)$ passes with probability at least $1 - \delta\frac{1}{4c}\frac{s}{qk_R}$, since otherwise Step (1) rejects with probability at least $\delta\frac{1}{4c}\frac{s}{qk_R}$ and we are done. We claim that conditioned on this event, the following two properties hold:

---

[22] Proposition 10.12 is formally for $(X(1), X(k))$, but because X is homogenous the bound carries over by simply setting each partite component of A to be $A_i = 1_{x=x'}$.

[23] Ties can be broken arbitrarily.

1. g_{dec} matches g_{maj} with high probability:

$$\Pr_{S \sim \mathcal{T}_C}[\exists v \in S : g_{\text{dec}}(v) \neq g_{\text{maj}}(v)] \leq \frac{s}{4ck_R}\delta$$

2. g_{maj} is far from the base code:

$$dist(g_{\text{maj}}, C) \geq \frac{1}{2k_R}\delta.$$

Soundness is then immediate: as long as the base tester $\mathcal{T}_C$ receives the 'correct' symbols from g_{maj}, it rejects with probability $s \cdot dist(g_{\text{maj}}) \geq \frac{s}{2k_R}\delta$ by Item 2. Since the former occurs with probability at least $\frac{s}{4k_R}\delta$ by Item 1, the total rejection probability is at least $\frac{s}{2k_R}\delta - \frac{s}{4k_R}\delta = \frac{s}{4k_R}\delta$ as desired.

It is left to prove the claimed properties.

Proof of Item 1. Observe that since the queries are marginally uniform, by a union bound it is enough to show that the decoding of a uniformly random vertex succeeds except with probability $\frac{s}{4cqk_R}\delta$. The proof is essentially a simpler variant of Claim 8.2. For a fixed $v \in L$ define the set of 'bad' neighbors as those whose labeling disagrees with the plurality decoding as:

$$B_v := \{w \in N(v) : f_w(v) \neq g_{\text{maj}}(v)\}.$$

The failure probability of our single-query decoding procedure is exactly $\mathbb{E}_{v \in L}[B_v]$, so it is enough to show

$$\mathbb{E}[B_v] \leq \frac{s}{4cqk_R}\delta.$$

This follows from the fact that the test $\mathcal{T}_G$ can be written as an expectation over complete local distributions. In particular, observe that we can decompose $\mathcal{T}_G$ as:

$$Agree_0(f) = \Pr_{v \sim L, w, w' \sim N(v)}[f_w(v) = f_{w'}(v)] \geq 1 - \frac{s}{4cqk_R}\delta.$$

This implies that the expected probability that $f_w(v) \neq f_{w'}(v)$ in the local distribution is at most:

$$\mathbb{E}_{v \sim L}\left[\Pr_{w, w' \sim N(v)}[f_w(v) \neq f_{w'}(v)]\right] \leq \frac{s}{4cqk_R}\delta.$$

Since w and w' are independent neighbors of v, the inner term is at least

$$\Pr_{w, w' \sim N(v)}[f_w(v) \neq f_{w'}(v)] = 2\mathbb{E}[B_x]\mathbb{E}[\overline{B_v}] \geq \mathbb{E}[B_v]$$

since $\mathbb{E}[\overline{B_v}] \geq \frac{1}{2}$ by construction, which completes the proof.

Proof of Item 2. The proof follows from the fact that conditioned on Step (1) (the agreement test) passing, $E_X(g_{\text{maj}})$ is close to a DP-codeword. Thus if g_{maj} itself is too close to C, it's encoding will be close to $E_G(C)$ violating our original assumption on the distance. Formally, assume for the sake of contradiction that g_{maj} is not $\frac{\delta}{2k_R}$-far from C. Then there exists $x \in C$ such that $dist(g_{\text{maj}}, x) < \frac{\delta}{2k_R}$. Since G is bi-regular, this further implies a bound on the distance of the *encodings* of g_{maj} and x:

$$\begin{aligned}
dist(E_G^k(g_{\text{maj}}), E_G^k(x)) &= \Pr_{w \sim R}[E_G(g_{\text{maj}})(w) \neq E_G(x)(w)] \\
&\leq k_R \Pr_{w \sim R, v \sim N(w)}[E_G(g_{\text{maj}})(w)_v \neq E_G(x)(w)_v] \\
&= k_R dist(g_{\text{maj}}, x) \\
&< \frac{\delta}{2}.
\end{aligned}$$

where the second equality follows from the fact that since G is bi-regular, sampling a uniformly random vertex $v \in R$ and a random $s \in N(v)$ is equidistributed with sampling a random $w \in L$. Finally closeness of the encoded words contradicts the fact that f is δ-far from $E_G(C)$, since $E_G(x)$ is a codeword and by the triangle inequality:

$$dist(f, E_G(x)) \leqslant dist(f, E_G(g_{\text{maj}})) + dist(E_G(g_{\text{maj}}), E_G(x)) < \delta,$$

where we have used the fact that since $Agree_0(f) \geqslant 1 - \frac{s}{4cqk_R}$ by assumption, f is $\frac{s}{4q}\delta$-close to $E_G(g_{\text{maj}})$ by Lemma 10.26. □

We remark that it is likely possible to reduce the query complexity of the above argument to just $q + 1$ by re-using the first query of $\mathcal{T}_C$ for the agreement test and performing a more careful conditioned analysis on the latter passing. Since q for us is some small constant, the difference between $q + 1$ and $q + 2$ is negligible so we only give the simpler independent version above.

To prove Theorem 10.21, we need a good explicit family of binary linear LTCs with near optimal distance to amplify via Theorem 10.25. In fact, as discussed in the proof overview we will need base codes with the stronger property of having small *bias*:

$$bias(C) := \max_{c \in C \setminus \{0\}} \left\{ \left| \mathbb{E}_{i \in [n]}[(-1)^{c_i}] \right| \right\}.$$

Note that an ε-biased linear code always has distance at least $\frac{1-\varepsilon}{2}$ since for every $c \in C$, $\left| \mathbb{E}_{i \in [n]}[(-1)^{c_i}] \right| = 1 - 2\,dist(c, 0)$. Our base codes start with the breakthrough construction of c^3-LTCs by Dinur, Evra, Livne, Lubotzky, and Moses [Din+22a]:

Theorem 10.27 ([Din+22a])**.** *There exists an explicit infinite family of binary, linear, uniform LTCs with constant rate, distance, locality, and soundness, and a linear time unique decoder up to constant radius.*

We remark that while [Din+22a] do not argue their testers are uniform, it is easy to check from their construction that this is the case. As discussed in the overview, it is possible to use the distance amplification techniques of [Kop+17] to achieve c^3-LTCs with distance near $1/2$ (even hitting the GV bound), but these codes may not be ε-biased and incur a heavy $\text{poly}(\varepsilon^{-1})$ cost in the query complexity. We will take a different amplification strategy that avoids these issues almost entirely at the cost of a small polynomial blowup in rate. The idea is to concatenate our amplification procedure in Theorem 10.25 with the *Hadamard code*:

$$Had : \mathbb{F}_2^k \to \mathbb{F}_2^{2^k}, \ \alpha \overset{Had}{\mapsto} \langle \alpha, \cdot \rangle$$

We denote the resulting code (the image of this map), by C_{Had}.

The Hadamard code comes with three properties critical to our construction. First, it is easy to check that the code is binary, linear, and perfectly balanced (0-biased). Second, it is a classical result that the Hadamard code is an excellent LTC:

Theorem 10.28 ([BLR90])**.** *The Hadamard code is a uniform $(3, 1)$-LTC.*

Finally, it is well-known folklore (and an easy exercise) to show the Hadamard code is *locally decodable*. We will not go into the formal details of locally decodable codes, but it suffices to say there is a randomized algorithm $Dec : \mathbb{F}_2^{2^k} \times [k] \to \mathbb{F}_2$ which given a word y such that $d(y, C_{\text{Had}}) < \frac{1}{4}$ and a coordinate i, queries only 2 bits of y and outputs the correct decoding with probability proportional to its distance from the code:

$$\mathbb{P}[Dec(y, i) \neq x_i] \leqslant 2d(y, C_{\text{Had}}) \tag{10.4}$$

where $x = Had^{-1}(y)$ is the pre-image of the unique closest codeword to y.

Note that this is essentially the same property we used in Theorem 10.25 in decoding the ABNNR-encoding. We will show that concatenating with the Hadamard code similarly allows us to amplify the bias and distance of the base code while only incurring constant blow-up in query complexity and logarithmic loss in soundness.

Proposition 10.29 (The Base Code)**.** *For every $\varepsilon > 0$ there is an explicit infinite family of uniform, ε-biased, linear $(\text{poly}(\varepsilon), \frac{1}{2} - \varepsilon)_2$-codes that are $(O(1), \log^{-1}(1/\varepsilon))$-locally testable. Moreover, there exists $t < 1$ such that these codes are uniquely decodable up to radius t in time $n \exp(\text{poly}(\varepsilon^{-1}))$.*

Proof. We apply Theorem 10.25 to [Din+22a]'s LTCs using the complex $X = W_G$ generated by length $k = O(\log(1/\varepsilon))$ random walks on any constant degree, regular $\frac{1}{2}$-spectral expander G (such expanders are known to exist for any n, see e.g. [Alo21]). This results in a $\mathrm{poly}(\varepsilon)$-rate code with distance at least $1 - \frac{\varepsilon}{2}$ by the classical expander hitting set lemma [HLW06] which states that for any set $T \subset X(1)$ of measure ρ

$$\mathop{\mathbb{P}}_{(s_1,\ldots,s_k) \in X(k)} [|\{s_1, \ldots, s_k\} \cap T| = 0] \leqslant \rho \left(\frac{1+\rho}{2}\right)^{k-1}. \tag{10.5}$$

The argument is then as in Corollary 10.24. Finally by Theorem 10.25 this encoding has a uniform $(c, \frac{c'}{\log(1/\varepsilon)})$-local tester for some universal constants $c, c' > 0$.

The key step is now to concatenate this resulting family with the Hadamard code:

$$\mathcal{C} := Had(E_X(C)),$$

where the Hadamard encoding is applied pointwise to each symbol in $E_X(C)$. This encoding maintains linearity and decreases the rate by a factor of at most $2^k = \mathrm{poly}(\varepsilon)$. To compute the bias, observe that since the Hadamard code is perfectly balanced, every non-zero symbol of a word $w \in E_X(C)$ contributes 0 to the bias so in the worst case we have

$$bias(\mathcal{C}) \leqslant \max_{c \in E_X(C) \setminus \{0\}} \{wt(c) \cdot 0 + (1 - wt(c)) \cdot 1\} \leqslant 1 - d(E_X(C)) \leqslant \varepsilon,$$

where $wt(c)$ denotes the normalized hamming weight (fraction of non-zero elements in c).

We now argue that this concatenation has a uniform local tester. Write $q = c$ and $s = \frac{c'}{\log(1/\varepsilon)}$ and recall $E_X(C)$ has a uniform (q, s)-local tester (which we'll refer to as the "base tester") which requests only a single bit of each queried symbol. We'll denote these requests as pairs $(x, i) \in X(k-1) \times [k]$. Our testing procedure follows a similar form as in Theorem 10.25. Namely given a potential word y:

1. Draw a uniformly random $x \in X$, and run the Hadamard test on $y(x)$.

2. For each (x, i) requested by the base tester, run the Hadamard decoder on $y(x)$ at coordinate i.

3. Run the base tester on the output

The test is complete by construction, and makes at most $2q + 3$ queries (two for each run of the local decoder, and three for the initial Hadamard test). Soundness follows along the same lines as Theorem 10.25. We first argue that the word y must be close to a cartesian product of Hadamard codewords (else Step 1 rejects). We then argue that the decoding succeeds with high probability, and the decoded word must be far from the base code (else y itself would be close to $\mathcal{C}$), so the base tester rejects with good probability.

Formally, let $y \in (\mathbb{F}_2^{2^k})^X$ satisfy $dist(y, \mathcal{C}) = \tau$. We argue that the above procedure rejects with probability at least $\frac{s\tau}{4q}$. To see this, first observe we can assume the word y is $1 - \frac{s\tau}{4q}$ close from being a product of Hadamard codewords. Otherwise Step 1 rejects with probability at least $\frac{s\tau}{4q}$ since

$$\mathbb{P}[\text{Step 1 Rejects}] \geqslant \mathbb{E}_x[dist(y(x), C_{Had})] = dist\left(y, \prod_X C_{Had}\right).$$

Fix some $y_{\mathrm{had}} \in \prod_X C_{Had}$ which minimizes the distance from y and denote the set of 'corrupted' coordinates of y to be

$$B := \left\{x \in X : dist\left(y(x), C_{\mathrm{Had}}\right) \geqslant \frac{1}{4}\right\}.$$

Note that the measure of B is at most $\frac{s\tau}{q}$, since the closeness of y to $\prod C_{\mathrm{Had}}$ is violated otherwise.

Define set of 'valid' decodings of y to be strings in $(\mathbb{F}_2^k)^X$ whose encodings match y_{had} on all uncorrupted coordinates:

$$Y_{\mathrm{dec}} := \left\{y_{\mathrm{dec}} \in (\mathbb{F}_2^k)^X : \forall x \in \bar{B}, Had(y_{\mathrm{dec}}(x)) = y_{\mathrm{had}}\right\}.$$

85

Note that when we *re-encode* any $y_{\text{dec}} \in Y_{\text{dec}}$, it will be at distance at most $\frac{s\tau}{q}$ from y since:

$$dist(Had(y_{\text{dec}}), y) \leqslant dist(Had(y_{\text{dec}}), y_{\text{had}}) + dist(Had(y_{\text{had}}), y)$$
$$\leqslant (\mathbb{P}[B] \cdot \frac{1}{2} + \mathbb{P}[\bar{B}] \cdot 0) + \frac{s\tau}{4q}$$
$$\leqslant \frac{s\tau}{q}.$$

We now argue that Step 2 correctly decodes to some $y_{\text{dec}} \in Y_{\text{dec}}$ with high probability. Since decoding always 'succeeds' on coordinates in B (since any decoded value is accepted), the worst case is when $B = \emptyset$ and y_{dec} is the pre-image of the unique closest word in every coordinate. In this case, local decodability of the Hadamard code (10.4) promises the probability any fixed query (x, i) fails to decode to $y_{dec}(x)_i$ is at most

$$\mathbb{E}_x[2dist(y(x), C_{Had})] \leqslant 2dist\left(y, \prod_X C_{\text{Had}}\right) \leqslant \frac{s\tau}{2q}.$$

Union bounding over all q queries, we decode to y_{dec} except with probability $\frac{s\tau}{2}$.

Finally, we argue any $y_{\text{dec}} \in Y_{\text{dec}}$ must be $\frac{3}{4}\tau$-far from $E_X(C)$, in which case the base tester rejects with probability at least $\frac{3}{4}s\tau - \frac{s\tau}{2} = \frac{s\tau}{4}$ by a union bound. This simply follows from the fact that if y_{dec} were $\frac{3}{4}\tau$-close to $E_X(C)$, then there exists $w \in E_X(C)$ such that

$$dist(Had(w), y) \leqslant dist(Had(w), Had(y_{\text{dec}})) + dist(Had(y_{\text{dec}}), y)$$
$$\leqslant dist(w, y) + \frac{s\tau}{q}$$
$$\leqslant \tau$$

where the second inequality comes from observing that the encodings of w and y_{dec} are the same on any symbol on which they agree, and can only improve in distance when they differ.[24]

It is left to show our concatenated family has an efficient unique decoder. It is a classical fact that concatenating codes with efficient unique decoding at radii t_{inner} and t_{outer} gives a code with an efficient unique decoder at radius $t_{\text{inner}}t_{\text{outer}}$ simply by running the inner decoder on every inner symbol, then running the outer decoder on the result [For65]. In particular decoding at this distance is successful since if a true word m is corrupted in at most a $t_{\text{inner}}t_{\text{outer}}$ fraction of its coordinates, no more than a t_{outer} fraction of inner codewords can have a t_{inner} fraction of corruptions. Thus running the inner decoder results in the correct word except in at most a t_{outer} fraction of outer coordinates, which is then successfully decoded by the outer decoder.

We can decode over each Hadamard symbol up to radius $t' < \frac{1}{4}$ by brute force in $\exp(\text{poly}(\varepsilon^{-1}))$ time (and otherwise output some fixed word, say 0, beyond this radius). $E_X(C)$ can be efficiently decoded up to some constant radius t by taking the majority vote over each coordinate, then running the base decoder on C (see e.g. [Din+18a]). This process takes time $n\exp(\text{poly}(\varepsilon^{-1})) + O(n)$, for a total time of $n\exp(\text{poly}(\varepsilon^{-1}))$ as desired. $\qquad\square$

Finally, we argue these base codes can be amplified to large alphabet LTCs with near optimal distance.

Corollary 10.30 (Large Distance c3-LTCs). *For all large enough $k \in \mathbb{N}$ and all $\varepsilon > 0$, there exists an infinite family of $\mathbb{F}_2$-linear $(r, d)_{|\Sigma|}$-codes that are (q, s)-locally testable for:*

$$r = \varepsilon^{O(k)}, \quad d = 1 - 2^{-k} - \varepsilon, \quad |\Sigma| = 2^k, \quad q = O(1), \quad s = O\left(\frac{1}{k\log(\frac{1}{\varepsilon})}\right)$$

Proof. We start with the explicit family of base codes given in Proposition 10.29 with $q = O(1)$ queries, distance $\frac{1}{2} - \frac{\varepsilon}{4}$, and rate $\text{poly}(\varepsilon)$. For a given code C in the family of blocklength n, let X be the collection of length-k walks on any explicit degree $\frac{\varepsilon}{2}$-expander of degree $O(\varepsilon^{-2})$ on n vertices (such graphs exist for

---

[24]Technically we've also used that $q \geqslant 4$ and $s \leqslant 1$ here, which follow from the parameters of [Din+22a].

any large enough n [Alo21]). The size of $X(k-1)$ is $O(\frac{n}{\varepsilon^{2k}})$, and the expander hitting-set lemma (Equation (10.5)) ensures the distance of $E_X(C)$ is at least

$$1 - \left(\frac{1}{2} + \frac{\varepsilon}{2}\right)^k \geqslant 1 - 2^{-k} - \varepsilon.$$

Finally, since the base code's tester is uniform, local testability follows from Theorem 10.25. □

It is worth comparing Theorem 10.25 to the distance amplification for LTCs of Kopparty, Meir, Ron-Zewi, and Saraf [Kop+17], who also analyzes local testability of distance-amplified codes via a related construction of Alon and Luby [AL96]. Their work largely focuses on achieving near-optimal *rate-distance* trade-off: they achieve the singleton bound over large alphabet codes and the GV-bound over $\mathbb{F}_2$. In comparison Theorem 10.25 has worse rate than their construction, but makes up for this with exponentially better alphabet size and local testability. Quantitatively, to achieve distance $1 - \beta$, [Kop+17]'s codes require alphabet size $2^{\mathrm{poly}(1/\beta)}$ and lose $\mathrm{poly}(1/\beta)$ factors in the testing parameters, while ours have the optimal β^{-1}-dependence in the former and lose only *logarithmically* in the latter.

10.6 List-Decoding ABNNR

In this section, we show that the codes in Corollary 10.30 are efficiently list-decodable, completing the proof of Theorem 10.21. Efficient list-decoding of the ABNNR construction was first considered by Dinur, Harsha, Kaufman, Livni-Navon, and TaShma [Din+18a], who gave an algorithm for the construction instantiated on HDX. After [Din+18a]'s work, Alev, Jeronimo, Quintana, Srivastava, and Tulsiani [Ale+20] showed it possible to efficiently list decode a closely related construction called a *direct sum code* over some relaxed families of splitting trees (including HDX and expander walks). Using a reduction to direct sum, they also recover the parameters of [Din+18a]'s ABBNR-codes on HDX as well.

In this section, we give an efficient list-decoding algorithm for the ABNNR-construction on complexes satisfying a slightly strengthened variant of splittability that ensures the complex can be "split" at any coordinate. We refer to this notion as complete splittability.

Definition 10.31 (Complete Splittability). A d-uniform partite complex X is called completely λ-splittable if for all $i \in [d]$:

$$\lambda_2(S_{[1,i],[i+1,d]}) \leqslant \lambda$$

where we recall $S_{[1,i],[i+1,d]}$ is the 'swap walk' adjacency operator between $X[[1,i]]$ and $X[[i+1,d]]$.

Complete splittability sits in-between tuple-splitting trees and partite HDX, and is closely related to notions of splittability defined in more recent works on direct-sum decoding [JST21; BD22] (in particular it is equivalent to the definition introduced in the former, see Appendix F for details). We can now state the main result of the subsection:

Theorem 10.32 (List Decoding on Completely Splittable Complexes). *For any $t < 1/2$, let C_0 be a binary code with blocklength n, $\mathrm{bias}(C_0) \leqslant 1 - 2t$, and which is unique-decodable to radius t in time $\mathcal{T}_0$. Let $X \subset [n]^k$ be a homogenous, completely λ-splittable complex, and $C = E_X(C_0)$ its ABNNR encoding. Let $d(C) = 1 - \varepsilon$ and $\tau > 0$ be any value satisfying*

$$\tau \geqslant \max\left\{\sqrt{\varepsilon}, 4\sqrt{(1-2t)^i + 2\lambda}, \sqrt{2^{21}\lambda k^3}, 8(1-t)^{i/2}\right\}$$

for $i = \frac{k}{2} - \sqrt{k \log(\frac{2}{\tau})}$. There exists a randomized algorithm, which given $\tilde{y} \in (\mathbb{F}_2^k)^X$ recovers the list

$$\mathcal{L}_\tau(\tilde{y}) = \{y \in C : d(\tilde{y}, y) \leqslant 1 - \tau\}$$

with probability $1 - \frac{\tau - \varepsilon}{\varepsilon(\tau^2 - \varepsilon^2)} 2^{-\Omega(n/t^2)}$ in time $\tilde{O}(\frac{\tau - \varepsilon}{\tau^2 - \varepsilon^2} \exp(\log(1/t)\exp(k^3/\tau^2))(|X| + \mathcal{T}_0))$. Moreover, the number of such codewords is at most:

$$|\mathcal{L}_\tau(\tilde{y})| \leqslant \frac{\tau - \varepsilon}{\tau^2 - \varepsilon^2}.$$

The proof of Theorem 10.32 goes via reduction to direct sum decoding. The *direct sum* encoding of a code C on a complex X simply concatenates the ABNNR-encoding with $XOR : \mathbb{F}_2^k \to \mathbb{F}_2$, that is $dsum_X(C) := XOR(E_X(C))$ (where XOR applied pointwise, similar to in the Hadamard concatenation in the prior section). Most of the work in proving Theorem 10.32 is already done via [JST21; BD22]'s direct sum decoder for completely splittable complexes:

Theorem 10.33 (List Decoding Direct Sum [JST21; BD22])**.** *For any $t < 1/2$, let C_0 be a binary code with blocklength n, $bias(C_0) \leqslant 1 - 2t$, and which is unique-decodable to radius t in time $\mathcal{T}_0$. Let $X \subset [n]^k$ be a homogenous, completely λ-splittable complex, and $C = dsum_X(C_0)$ be its direct sum encoding. Let $\varepsilon = bias(C)$ and $\beta > 0$ be such that*

$$\beta \geqslant \max\left\{ \sqrt{\varepsilon}, \sqrt{2^{20}\lambda k^3}, 2(1-t)^{k/2} \right\}.$$

There exists a randomized algorithm which given $\tilde{y} \in \mathbb{F}_2^X$ recovers the list

$$\mathcal{L}_\beta^{X,(sum)}(\tilde{y}) = \left\{ x \in C_0 : d(\tilde{y}, dsum_X(x)) \leqslant \frac{1}{2} - \beta \right\}$$

with probability $1 - \frac{1}{\varepsilon} 2^{-\Omega(n/t^2)}$ in time $\tilde{O}(\exp(\log(t)\exp(\frac{k^3}{\beta^2}))|X| + \mathcal{T}_0)$. Moreover, the list has at most

$$|\mathcal{L}_\beta^{X,(sum)}(\tilde{y})| \leqslant \frac{1}{\varepsilon}$$

codewords.

We remark that the requirement on the unique decoder radius in [JST21; BD22] is slightly more strict as stated, but one can easily modify parameters in their proof to achieve the above. We now prove Theorem 10.32 by reduction to Theorem 10.33.

Proof of Theorem 10.32. We appeal to a simple variant of the reduction of [Ale+20]. Given a potential word $\tilde{y} \in (\mathbb{F}_2^k)^X$ at distance $d(\tilde{y}, C) = 1 - \tau$, we claim the following process is an efficient list-decoder.

1. For every $F \subseteq [d]$, apply Theorem 10.33 to $XOR(\tilde{y}|_F)$ on $dsum_{X^F}(C)$ with $\beta = \tau/4$ to get:

$$\mathcal{L}_\beta^{pot}(\tilde{y}) := \bigcup_{F \subseteq [d+1]} \mathcal{L}_\beta^{X^F,(sum)}(\tilde{y})$$

2. Prune any decodings that are far from $\tilde{y}$:

$$\mathcal{L}_\tau(\tilde{y}) := \left\{ E_X(x) : x \in \mathcal{L}_\beta^{pot}(\tilde{y}),\ d(E_X(x), \tilde{y}) \leqslant 1 - \tau \right\}.$$

To prove this procedure indeed gives a valid list-decoding we will use the fact that any projection of a splittable complex, in particular each X^F above, is also splittable.

Claim 10.34. Let $X \subset [n]^d$ be a homogeneous, completely λ-splittable complex. Then for any $F \subseteq [d]$, X^F is homogeneous and completely λ-splittable.

The proof follows from the inheritance of expansion under projection, and is defered to Appendix F.

We now prove the output list $\mathcal{L}_\tau$ is exactly the set of codewords at distance at most $1 - \tau$ from C. Recall that our base code is assumed to be at most $(1 - 2t)$-biased. By bias amplification lemma for splitting trees (Proposition 10.13) and Claim 10.34, we have that for any X^F, the bias of $dsum_{X^F}(C)$ is at most $(1 - 2t)^{|F|} + 2\lambda$. As such, our choice of $\beta = \tau/4$ satisfies the conditions of Theorem 10.33 whenever $i \leqslant |F| \leqslant k$ (for i as in the Theorem statement). Let $y \in E_X(C)$ be a codeword at distance at most $1 - \tau$ from $\tilde{y}$, then we have

$$\mathbb{P}[\tilde{y}_s = y_s] = \mathbb{E}_s\left[\mathbb{E}_{F \in \mathbb{F}_2^k}[\chi_F(\tilde{y}_s + y_s)] \right] \geqslant \tau,$$

where $\chi_F(y) = (-1)^{\langle F,y \rangle}$ are the standard characters. By averaging and Chernoff, there must exist $|F| \geqslant \frac{k}{2} - \sqrt{k \log(2/\tau)})$ such that

$$\mathbb{E}_s \left[\chi_F(\tilde{y}_s) \chi_F(y_s) \right] \geqslant \tau/2,$$

which is equivalent to the statement that the distance of $\tilde{y}|_F$ from $y|_F$ is at most $\frac{1}{2} - \frac{\tau}{4} = \frac{1}{2} - \beta$ in the direct sum encoding on X^F. Since X^F is homogeneous and completely λ-splittable, by Theorem 10.33 we have that the decoding of y appears in $\mathcal{L}_{pot}(\tilde{y})$ except with probability $\frac{1}{\varepsilon} \exp(n/t^2)$. By the Johnson bound (see e.g. [Din+18a][Theorem 2.15]), there can be at most $\frac{\tau - \varepsilon}{\tau^2 - \varepsilon^2}$ such codewords y close to $\tilde{y}$, so union bounding over these gives that all valid codewords are in the list with the desired probability. Finally, all extraneous decodings are removed in the second step, which completes correctness of the procedure.

Running time follows immediately from the fact that we have run the direct sum decoder at most 2^k times, and the resulting potential list has at most $\exp(k)$ potential codewords by the Johnson bound and our choice of $\beta = \tau/4$, each of which can be pruned in $O(k|X|)$ time. $\qquad\square$

Finally, we instantiate this result on expander walks to complete the proof of Theorem 10.21.

Corollary 10.35. *The family of codes in Corollary 10.30 is $(1 - 2^{-\Omega(k)}, \frac{1}{2^{O(k)}})$ efficiently list-decodable with confidence $2^{-\Omega(n-k)}$.*

Proof. Since [JST21] prove that walks on an λ-expander are homogeneous and completely λ-splittable, it is enough to ensure the parameter τ in Theorem 10.32 can be set to $2^{-\Omega(k)}$. Recalling that the base code of this construction is ε-biased and uniquely decodable in linear time up to constant radius, and that the encoding has distance at least $1 - 2^{-k} - \varepsilon$, as long as $\varepsilon \leqslant \exp(-\Omega(k))$ one can easily check it is possible to take $\tau = \exp(-\Omega(k))$ and $\frac{\tau - \varepsilon}{\tau^2 - \varepsilon^2} = 2^{-O(k)}$ satisfies all conditions of Theorem 10.32. $\qquad\square$

10.7   Toward Lossless Amplification from HDX?

Our expander-based construction experiences polylogarithmic decay in soundness with the alphabet size. This is exponentially better than prior methods, but it is plausible the decay could be avoided entirely. In this section, we propose an approach toward such 'lossless' amplification via the ABNNR-Encoding on an HDX.

Instantiating Theorem 10.25 directly on an HDX does actually slightly improve soundness over Theorem 10.21, but the loss of $\frac{1}{k}$ in the dimension of the complex is inherent to our approach, no matter the instantiation. The reason is that our method only simulates the base tester a single time. As a result, the rejection probability will only scale with the distance of the decoded word from the *base* code. On the other hand, because X amplifies this distance by a factor of k, this results in an unavoidable loss in soundness.

There is a natural approach to fixing this issue: simulate the base tester "*k times in parallel*," analogous to Raz's parallel repetition theorem [Raz95] and its de-randomized variants in agreement testing [IKW12; DM11; DD23b; BM23b]. Indeed on the complete complex, this approach can actually be carried out successfully:

Theorem 10.36 (Parallel LTC Amplification). *Let C be a uniformly (q, s)-locally testable $(r, \frac{1-\varepsilon}{2})_2$-code of blocklength n and $X = \binom{[n]}{k}$. Then $E_X(C)$ is a uniformly (q', s')-locally testable $(r', d')_{|\Sigma|}$-code for:*

$$r' = \frac{r}{n^{k-1}}, \quad d' = 1 - \frac{1}{2^k}(1+\varepsilon)^k - o_n(1), \quad |\Sigma| = 2^k, \quad q' = q + 2, \quad s' = cs$$

where $c > 1$ is some universal constant (independent of k).

Proof. Rate and alphabet are immediate from construction, and distance follows from (near)-independence of $\binom{[n]}{k}$. For local testability, we rely on the following agreement theorem of Dinur and Steurer:

Claim 10.37 ([DS14]). There exists a two query agreement test on $\binom{[n]}{k}$ and a constant $c \geqslant 1$ such that[25]

$$Agree_0(\mathcal{F}) \geqslant 1 - \varepsilon \quad \implies \quad dist(\mathcal{F}, g_{\text{maj}}) \leqslant c\varepsilon$$

---

[25]We remark that formally Dinur and Steurer consider $[n]^k$, but their techniques transfer to this setting (formally, see e.g. [DD19] for the quoted result).

With this in hand the proof is similar to Theorem 10.25. For k independent runs of the base tester $\mathcal{T}_C$, denote the requested symbols by $\{x_j^{(i)}\}_{j \in [q], i \in [k]}$ and write $s_j := \{x_j^{(1)}, \ldots, x_j^{(k)}\} \in \binom{[n]}{k}$. Given a word f, our tester performs the following two-step procedure:

1. Run the Dinur-Steurer test on f

2. For every $i \in [q]$:

    – For each $j \in [k]$, denote the 'decoded' plurality function at $x_j^{(i)}$ as $g_{\mathrm{dec}}(x_j^{(i)}) = f_{s_j}(x_j^{(i)})$

    – Run $\mathcal{T}_C$ on $g_{\mathrm{dec}}(x_1^{(i)}), \ldots, g_{\mathrm{dec}}(x_q^{(i)})$

The tester is complete and makes $q + 2$ queries by construction. The proof of soundness is very similar to Theorem 10.25, so we give a shortened explanation here. Assuming $dist(f, E_X(C)) = \delta$, we need to show our test rejects with probability $\Omega(\delta)$. Assume Step (1) passes with probability at least $1 - \frac{s\delta}{8cq}$ (else we are done). Since the base tester is marginally uniform and the k runs are independent, each face $s_j \in X$ is marginally uniform, and therefore by Claim 10.37 exactly matches plurality except with probability $\frac{s\delta}{8q}$. Union bounding over the q faces gives that g_{dec} matches g_{maj} on all queried points except with probability $\frac{s}{8}\delta$. By the same argument as in Theorem 10.25, g_{maj} must be at least $\frac{\delta}{2k}$-far from C, so conditioned on receiving the correct values, at least one of the base testers rejects with probability $1 - (1 - \frac{s\delta}{2k})^k \geqslant \frac{s\delta}{4}$ for a total rejection probability of $\frac{s\delta}{4} - \frac{s\delta}{8} = \frac{s\delta}{8}$ as desired. □

The issue with Theorem 10.36 is that the resulting code has very poor rate. It is natural to conjecture that, as in the single query variant, this can be resolved by replacing $\binom{[n]}{k}$ with a high dimensional expander. Indeed it is known that the agreement testing step (the analog of Claim 10.37) works in this setting [DK17; DD19; KM20]. Note that this is *not* the case for other splitting trees such as expander walks, which would lose a factor of k. The challenge then mainly lies in extending the parallel simulation. Below, we discuss the current challenges toward porting this second half of the proof to HDX.

The first difficulty in parallel simulation is that the HDX and base code must 'match' in the sense that there is a natural distribution over q-tuples of faces that corresponds to k valid q-query tests for the base code. Recently, Golowich [Gol23] showed how to construct HDX from certain (modified) locally testable codes, and Dinur, Liu, and Zhang [DLZ23] built LTCs (of sub-constant rate) that sit on the triangles of an HDX. It is not clear a priori how to fit either of these constructions into our framework, but it is plausible they or modifications thereof do.

The second difficulty in simulating the parallel tester is that the k runs of the base test will be inherently correlated. There is a long history in the HDX literature of handling this type of behavior (e.g. [DK17; DD19; KM20; GK23b; DD23b; BM23b]), including the agreement tests we discuss in Section 8. Very recently, building upon tools in this work, [DD23b; BM23b] prove that certain strong families of high dimensional expanders admit 1%-regime agreement tests, which are closely related to parallel repetition [IKW12; DM11]. As a start, it would be interesting to extend Theorem 10.36 even to dense complexes known to admit such tests, such as the Grassmann [IKW12] or spherical buildings [DD23b]. To summarize, we give the following informal conjecture:

Conjecture 10.38 (Parallel LTC Amplification). *There exists a family of high dimensional expanders and locally testable codes which admit a "parallel" variant of Theorem 10.25, that is:*

1. *There is a q-step random walk or distribution on X generating k valid 'parallel' tests*

2. *The resulting parallel tester has no (asymptotic) decay in soundness.*

11   Concentration Lower Bounds

In this section we prove our results are essentially optimal with respect to their dependence on the sampling parameters and quantitative expansion of the complex. We further show one cannot hope to achieve our more involved applications such as reverse hypercontractivity using classical samplers such as expander walks.

11.1 Optimality for Inclusion Samplers

In Section 5.2, we proved that the if X is a high dimensional expander, then the inclusion graphs $(X(k), X(i))$ are roughly $(\varepsilon, \exp(-\varepsilon^2 \frac{k}{i}))$-additive samplers. Inspecting the proof of Theorem 5.1, it is easy to show that for any set A of density μ and strong enough local-spectral expansion $\lambda \ll \mu$, one can give the following improved bound:[26]

$$\mathop{\mathbb{P}}_{s \in X(k)} \left[\mathop{\mathbb{P}}_{r \in X(i)} [A \mid r \subseteq s] \geqslant \mathbb{P}[A] + \varepsilon \right] \leqslant \exp \left(-\Omega \left(\frac{\varepsilon^2}{\mathbb{P}[A]} \frac{k}{i} \right) \right).$$

In this section we prove that for most $i \leqslant k$, this is essentially optimal. Recall that a complex X is called (γ, i)-hitting if for any $A \subseteq X(1)$, $\mathbb{P}_{\sigma \in X(i)}[\sigma \subseteq A] \leqslant \mathbb{P}[A]^i + \gamma$. We say that X is (γ, i)-hitting for sets of size μ if $\mathbb{P}_{\sigma \in X(i)}[\sigma \subseteq A] \leqslant \mathbb{P}[A]^i + \gamma$ holds for any $A \subseteq X(1)$ of size $\mathbb{P}[A] = \mu$.

Theorem 11.1. *Let $\varepsilon \in (0, 0.1)$. Let $i < k$ be two integers satisfying $\frac{k}{i} \geqslant 3\varepsilon^{-2}$ and let $n \geqslant ke^k$. Let X be a k-uniform simplicial complex such that every vertex in X has the same probability $\frac{1}{n}$. If X is $(i, \frac{\varepsilon}{3})$-hitting, then $(X(k), X(i))$ is not an (ε, β)-additive sampler for $\beta < \exp(-O(\varepsilon^2 \frac{k}{i}))$.*

In particular, for any $\mu \in (0, 0.2)$, if X is $(i, \frac{\min(\varepsilon, \mu)}{3})$-hitting for sets of size $\frac{\mu}{i}$ there exists $E \subseteq X(i)$ with $\frac{1}{2}\mu \leqslant \mathbb{P}[E] \leqslant \mu$ such that $\mathbb{P}_{s \in X(k)} \left[\mathbb{P}_{r \in X(i)}[E \mid r \subseteq s] > \mathbb{P}[E] + \varepsilon \right] \geqslant \exp(-O(\frac{\varepsilon^2}{\mu} \frac{k}{i}))$.

While the assumption that X is $(i, \frac{\varepsilon}{3})$-hitting is not completely general, it follows from $(X(k), X(i))$ being a modest sampler. We prove this at the end of the subsection.

Claim 11.2. Let $(X(k), X(i))$ be an $(\frac{1}{10i}, \frac{\varepsilon}{6i})$-additive sampler. Then X is a $(i, \frac{\varepsilon}{3})$-hitting for all sets of size at least $\frac{0.2}{i}$.

This results in the variant of the bound stated in the introduction.

Corollary 11.3. *Let $\varepsilon \in (0, 0.1)$. Let k and n be large enough (as a function of ε). Let X be a k-uniform simplicial complex such that every vertex in X has the same probability $\frac{1}{n}$. If $(X(k), X(i))$ is a $(\frac{1}{10i}, \frac{\varepsilon}{6i})$-additive sampler, then $(X(k), X(i))$ is not a (ε, β)-additive sampler for any $\beta < \exp(-O(\varepsilon^2 \frac{k}{i}))$.*

The assumption that the vertices all have the same probability is also not general, see the discussion in the end of this subsection.

11.1.1 Proof of Theorem 11.1

For the rest of the section we restrict our attention to "one-sided" additive sampling, that is:

$$\mathop{\mathbb{P}}_s \left[\mathop{\mathbb{P}}_{v \sim s} [A \mid s] > \mathbb{P}[A] + \varepsilon \right].$$

This is no larger than the probability that $|\mathbb{P}_{v \sim s}[A \mid s] - \mathbb{P}[A]| < \varepsilon$ and is more convenient to handle.

Let us introduce some notation, similar to the notation in Section 7.1 (but for one-sided sampling in different levels of the complex). For a k-uniform simplicial complex, $i \leqslant k$ and constants $\mu, \varepsilon > 0$ let

$$\pi(X, k, i, \mu, \varepsilon) = \max \left\{ \mathop{\mathbb{P}}_{s \in X(k)} \left[\mathop{\mathbb{P}}_{r \in X(i)} [A \mid r \subseteq s] > \mathbb{P}[A] + \varepsilon \right] \,\middle|\, A \subseteq X(i), \mathbb{P}[A] = \mu \right\}.$$

That is the best one-sided sampling bound for when restricting to sets of relative size μ.[27]

For general sampler graphs $G = (L, R, E)$ we use similar notation:

$$\pi(G, \varepsilon, \mu) = \max \left\{ \mathop{\mathbb{P}}_{v \in R} \left[\mathop{\mathbb{P}}_{u \in L} [A \mid v \sim u] > \mathbb{P}[A] + \varepsilon \right] \,\middle|\, A \subseteq L, \mathbb{P}[A] = \mu \right\}.$$

We also denote by Δ_n the complete complex on n vertices.

---

[26]Though formally when $i = \Omega(\frac{k}{\log k})$ one needs to assume ε is a large enough to subsume the $\frac{1}{\varepsilon}$ factor in front of the exponent.

[27]Here let us assume for simplicity that there exists sets of this probability. We note that this assumption is justified because as n grows larger the possible sizes of sets, $\{\frac{k}{n}\}_{k=0}^n$, becomes dense in $[0, 1]$. The use of this claim later on is to find a set of size $\approx \mu$ anyway, so we neglect dealing with this minor point.

Proof of Theorem 11.1. We start by stating a general bound on *all* samplers (not just inclusion samplers), via reduction to the sampling properties of the complete complex.

Claim 11.4. Let $k, n > 0$. Let G be a bipartite graph such that $|L| = n$ and such that it is k-right regular. Then

$$\pi(G, \varepsilon, \mu) \geqslant \pi(\Delta_n, k, 1, \varepsilon, \mu).$$

We defer the proof, and first argue this implies Theorem 11.1.

Fix μ as in the theorem statement. By Claim 11.4 there exists $A \subseteq X(1)$ of measure $\mathbb{P}[A] = \frac{\mu}{i}$ such that

$$\mathbb{P}_{s \in X(k)}\left[\mathbb{P}[A \mid s] > \mathbb{P}[A] + \frac{10\varepsilon}{i}\right] \geqslant \pi(\Delta_n, k, 1, \frac{\mu}{i}, \frac{10\varepsilon}{i}).$$

Let $E = N_i(A)$ be the set of all i-faces that hit A and let $\mu' = \mathbb{P}[E]$. We claim that

$$\mu = i\,\mathbb{P}[A] \geqslant \mathbb{P}[E] \geqslant \frac{i}{2}\,\mathbb{P}[A] = \frac{\mu}{2}$$

The lower bound follows from the hitting property, namely:

$$\mathbb{P}[E] \geqslant 1 - (1 - \frac{\mu}{i})^i - \frac{\mu}{3} \geqslant \frac{\mu}{2} = \frac{i}{2}\mathbb{P}[A]$$

since $\mu < 0.2$.

The upper bound follows because the only way to sample a vertex in A is to sample an i-face in E, then subsample a vertex in A in that i-face. The probability of doing so (conditioned on landing in E) is at least $\frac{1}{i}$.

The proof will follow if we can show that

$$\mathbb{P}_{s \in X(k)}\left[\mathbb{P}_{r \in X(i)}[E \mid r \subseteq s] > \mathbb{P}[E] + \varepsilon\right] \geqslant \pi(\Delta_n, k, 1, \frac{\mu}{i}, \frac{10\varepsilon}{i}) \tag{11.1}$$

since Chernoff's bound is tight on the complete complex.

Lemma 11.5 (Reverse Chernoff, e.g. [KY99, Lemma 5.2])**.** *Let ε, μ and k be such that $\frac{\varepsilon'^2}{\mu'}k \geqslant 3$, and let $n \geqslant ke^k$, then $\pi(\Delta_n, k, 1, \mu', \varepsilon') \geqslant \exp(-10\frac{\varepsilon'^2}{\mu'}k)$.*[28]

Applying this lemma with $\mu' := \frac{\mu}{i}$ and $\varepsilon' := \frac{10\varepsilon}{i}$ and using the fact that $\frac{\varepsilon'^2}{\mu'}k \geqslant 3$ by assumption, implies

$$\mathbb{P}_{s \in X(k)}\left[\mathbb{P}_{r \in X(i)}[E \mid r \subseteq s] > \mathbb{P}[E] + \varepsilon\right] \geqslant \pi(\Delta_n, k, 1, \frac{\mu}{i}, \frac{10\varepsilon}{i}) \geqslant \exp(-O(\frac{\varepsilon^2}{\mu}\frac{k}{i})).$$

Toward proving (11.1) observe that for a fixed $s \in X(k)$, the probability of E inside s depends only on the probability of A inside s.

Assume that $\mathbb{P}[A|s] - \mathbb{P}[A] > \frac{10\varepsilon}{i}$. Then

$$\mathbb{P}_{r \in X(i)}[E \mid r \subseteq s] > 1 - (1 - \mathbb{P}[A \mid s])^i$$

$$\geqslant 1 - (1 - \mathbb{P}[A] - \frac{10\varepsilon}{i})^i$$

$$\geqslant 1 - (1 - \mathbb{P}[A])^i + 2\varepsilon$$

$$\geqslant \mathbb{P}[E] + \varepsilon$$

where the first inequality is by observing the choice of $r \subseteq s$ subsamples i random vertices (without replacement). The probability of hitting A in the latter is then at least $1 - (1 - \mathbb{P}[A \mid s])^i$, the hitting

---

[28]Technically the lemma in [KY99] is for sampling with replacement, instead of sampling without replacement (which is the complete complex case), but for $n \geqslant ke^k$ the two are close enough in TV-distance so we get essentially the same theorem.

probability with replacement. The third inequality is true when $i \geqslant 1$ and $\varepsilon < 0.1$. Combining this with our original guarantee on sampling A itself we have

$$\pi(X, k, i, \mu', \varepsilon) \geqslant \mathop{\mathbb{P}}_{s \in X(k)} \left[\mathop{\mathbb{P}}_{r \in X(i)} [E \mid r \subseteq s] > \mathbb{P}[E] + \varepsilon \right]$$

$$\geqslant \mathop{\mathbb{P}}_{s \in X(k)} \left[\mathbb{P}[A \mid v \in S] \geqslant \mathbb{P}[A] + \frac{10\varepsilon}{i} \right]$$

$$\geqslant \pi(\Delta_n, k, 1, \frac{\mu}{i}, \frac{10\varepsilon}{i})$$

as desired. $\qquad \square$

It remains to prove Claim 11.4. The proof is basically repeating the lower bound in [CEG95]. The latter show any degree-k (ε, δ)-additive sampler satisfies $\delta \geqslant \exp(-\Omega(\varepsilon^2 k))$. Similar to our strategy above, their proof can be interpreted as showing $(\Delta_n(k), \Delta_n(1))$ containment graph is the optimal sampler in terms of right-degree, then reducing the general case to Δ_n. We give a refinement of their argument for sets of specific size, though our definition of samplers is less general than theirs.

Proof of Claim 11.4. Let D_μ be the uniform distribution over all sets $A \subseteq L$ of probability μ. Then

$$\pi(G, \varepsilon, \mu) \geqslant \mathop{\mathbb{P}}_{r \in R, A \sim D_\mu} \left[\mathop{\mathbb{P}}_{v \in L} [A \mid v \sim r] > \mu + \varepsilon \right].$$

Observe that since we randomize over A and G is right regular, the choice of $r \in R$ doesn't affect the inner probability. As such, we may fix any $r_0 \in R$ and instead consider

$$\mathop{\mathbb{P}}_{A \sim D_\mu} \left[\mathop{\mathbb{P}}_{v \in L} [A \mid v \sim r_0] > \mu + \varepsilon \right],$$

that is the probability (over A) that $A \cap r_0$ contains $\geqslant (\mu + \varepsilon)k$ elements.

We claim this is equal to $\pi(\Delta_n, k, 1, \varepsilon, \mu)$. To see this, assume without loss of generality that $L = \{1, 2, \ldots, n\}$ and $r_0 = \{1, 2, \ldots, k\}$. Since we have assumed the vertex weights are uniform, A is a uniform sample of μn points out of L. It is convenient to think of A as being sampled via the following procedure:

1. Draw a random permutation $\sigma \in S_n$

2. Set $A = \sigma(A_0)$, where $A_0 = \{1, 2, \ldots, \mu n\}$

Observe $|r_0 \cap A| \geqslant (\mu + \varepsilon)k$ if and only if $|\sigma^{-1}(r_0) \cap A_0| \geqslant (\mu + \varepsilon)k$. But note that the random variable $\sigma^{-1}(r_0)$ is exactly a random choice of a set $\Delta_n(k)$. Therefore, this is equal to

$$\mathop{\mathbb{P}}_{s \in \Delta_n(k)} \left[\mathop{\mathbb{P}}_{v \in [n]} [A_0 \mid v \in s] > \mu + \varepsilon \right].$$

Since every set of vertices of equal measure has the same sampling in Δ_n by symmetry, this is exactly $\pi(\Delta_n, k, 1, \varepsilon, \mu)$ as desired. $\qquad \square$

11.1.2 Sampling implies hitting

We briefly prove Claim 11.2.

Proof of Claim 11.2. Let $A \subset X(1)$ be a set of size $\frac{\mu}{i}$ and $E \subseteq X(i)$ its set of neighbors. We need to show

$$\mathbb{P}[E] \geqslant 1 - (1 - \mathbb{P}[A])^i - \frac{\varepsilon}{3}.$$

Let $f : X(i) \to [0, 1]$ be $f(s) = \frac{|s \cap A|}{i}$ and observe that $\mathbb{E}_{s \in X(i)}[f(s)] = \frac{\mu}{i}$. Let B be the set of bad k-faces that over-samples f:

$$B = \left\{ t \in X(k) \;\middle|\; \mathop{\mathbb{E}}_{s \subseteq t} [f(s)] \geqslant \frac{\mu}{i} + \frac{1}{10i} \right\},$$

then by our sampling guarantee $\mathbb{P}\left[B\right] < \frac{\varepsilon}{6i}$.

Fix $t \notin B$. Then it holds that $\mathbb{P}_{v \in X(1)}\left[A \mid\ v \in t\right] \leqslant \frac{\mu+1}{i} \leqslant \frac{1.1}{i}$ because $\mathbb{P}_{v \in X(1)}\left[A \mid\ v \in t\right] = \mathbb{E}_{s \subseteq t}\left[f(s)\right]$. Moreover, by the hitting properties of the complete complex, we have that $\mathbb{P}\left[E \mid\ t\right] \geqslant 1 - (1 - \mathbb{P}\left[A \mid\ t\right])^i$ (using the fact that inside the complete complex, the sampling is negatively correlated). Thus

$$
\begin{aligned}
\mathbb{P}\left[E\right] &= \mathbb{E}_{t}\left[\mathbb{P}\left[E \mid\ t\right]\right] \\
&\geqslant \mathbb{P}\left[\neg B\right] \mathbb{E}_{t \notin B}\left[\mathbb{P}\left[E \mid\ t\right]\right] \\
&\geqslant (1 - \frac{\varepsilon}{6})(1 - \mathbb{E}_{t \notin B}\left[(1 - \mathbb{P}\left[A \mid\ t\right])^i\right]) \\
&\geqslant (1 - \frac{\varepsilon}{6})(1 - \mathbb{E}_{t \notin B}\left[(1 - \mathbb{P}\left[A \mid\ t\right] - \frac{\varepsilon}{6i})^i\right]) \\
&\geqslant (1 - \frac{\varepsilon}{6})(1 - (1 - \mathbb{P}\left[A\right])^i - \frac{\varepsilon}{6}) \\
&= 1 - (1 - \mathbb{P}\left[A\right])^i - \frac{\varepsilon}{3}.
\end{aligned}
$$

$\square$

11.1.3 Discussion

We conclude the section with a few remarks.

Remark 11.6.

1. [Theorem 11.1](#) doesn't follow from the tight bounds depending on degree or on $\frac{|X(k)|}{|X(i)|}$ [CEG95]. The degree of a k-face in $(X(k), X(i))$ is $\binom{k}{i}$, and the degree of a set in $X(i)$ can be arbitrarily large, as in the complete complex case. The lower bounds in [CEG95], imply a lower of $\exp(-O(\binom{k}{i}))$ instead of $\exp(-O(\frac{k}{i}))$. In addition, the ratio $\frac{|X(k)|}{|X(i)|}$, which is another source for lower bounds in [CEG95], can be arbitrarily large. That being said, we will *reduce* to the bound in [CEG95] (only for small sets).

2. While most high dimensional expanders can be symmetrized to satisfy the assumed vertex uniformity [FI20], it is worth noting the condition can likely be relaxed to a bound on $\max_{s \in X(k)} \mathbb{P}_{v \in X(1)}\left[v \in s\right]$ with some additional effort. Further, some assumption of this sort is necessary. To see this, take the complete k-partite complex X where $X[i] = \{x_i\}$ are singletons for all $i < k$, and $X[k] = \{y_1, y_2, \ldots, y_n\}$. That is $X(k) = \left\{\{x_0, x_1, \ldots, x_{k-1}\} \cup \{y_j\} \mid\ j = 1, 2, \ldots, n\right\}$ are all possible faces between the different parts. For any $f : X(1) \to [0,1]$ we note that $\mathbb{E}[f] = \frac{\sum_{i=1}^{k-1} f(x_0)}{k} + \frac{1}{k} \mathbb{E}_{y_i \in X(k)}\left[f(y_i)\right]$. Thus for every s, $|\mathbb{E}_{v \subseteq s}\left[f(v)\right] - \mathbb{E}[f]| \leqslant \frac{1}{k}$. In particular, for every $\varepsilon > \frac{1}{k}$, the inclusion graph is an $(\varepsilon, 0)$-sampler.

3. We note that a similar bound can be achieved by essentially the same proof assuming only $\frac{\mu}{3}$-hitting, albeit at the cost of worse dependence on ε. In particular, assuming only that X is $\Theta(1)$-hitting, one can still prove a lower bound against $(\varepsilon, \exp(-O(\varepsilon \frac{k}{i})))$-sampling.

11.2 High Dimensional Expansion

Having shown that the concentration bounds of [Theorem 5.1](#) cannot be quantitatively improved, we turn to understanding the extent to which our requirements on the *expansion* of the underlying complex are necessary for strong concentration. We give two lower bounds to this effect. First, we argue that concentration is not implied by local-spectral expansion at or below the TD-barrier.

Proposition 11.7 (Lower Bounds at the TD-Barrier). *For every $\beta < 1$ and $d \in \mathbb{N}$, there exists a family of 1-TD complexes $\{X_n\}$ such that $(X_n(d), X_n(1))$ is not a $(\frac{1}{2}, \beta)$-additive sampler.*

Thus in this sense our results giving concentration bounds for λ-TD complexes for any $\lambda < 1$ are tight. Note that by [Lemma 4.15](#), the above also implies failure of sampling for these complexes for any (large enough) $k \leqslant d$ (this can also be shown directly by the same method).

94

The proof of Proposition 11.7 is based on the 'product' complexes of Golowich [Gol21], whose construction we quickly describe. Given a weighted graph G defines the product-complex X_G as

$$X_G(d) = \left\{ \{(v_1, s_1), \ldots, (v_d, s_d)\} : \{v_i\}_{i \in [d]} \in E, \{s_i\}_{i \in [d]} \in \binom{[n]}{d} \right\}.$$

The measure of each face $\sigma = \{(v_1, s_1), \ldots, (v_d, s_d)\}$ corresponding to an edge $\{s, t\}$ is proportional to $w(s, t) \cdot f(j)$, where f is some weighting function dependent only on the number $v_i = s$.

Proof of Proposition 11.7. Let $G = K_{n/2}^{(1)} \amalg K_{n/2}^{(1)} \amalg \{e\}$, that is two disjoint copies of $K_{n/2}$ with an additional edge e passing between them, and take the weights of every edge to be uniform. [Gol21] shows that there exists a weighting function f such that X_G is connected and has expansion $\frac{1}{d}$ in every $(d-2)$-link. By construction, the 1-skeleton of X is a $(d+1)$-cover of G with edges $(\{v_0, s_0\}, \{v_1, s_1\})$ whenever $\{v_0, v_1\}$ is an edge in E, and $s_0 \neq s_1$. With this in mind, define the set A to be the indicator of the first clique in each part:

$$A := \{(v, s) : v \in K_{n/2}^{(1)}\}$$

A clearly has measure $\frac{1}{2}$ by construction, while on the other hand any $\sigma \in X(d)$ corresponding to an edge $e' \neq \{e\}$ is either entirely contained in or entirely misses A. Since all edges are evenly weighted, the probability that a d-face corresponds to $\{e\}$ is at most $O(n^{-2})$, which gives the desired bound for large enough n. □

The above proof extends immediately to a lower bound on sampling of X_G under any weight function and any graph G simply by taking A to be any (nearly) balanced subset in G with expansion (near) $\frac{1}{2}$. It is an easy exercise to show that choosing A randomly suffices (and namely that such a set always exists).

Our second lower bound studies a slightly different regime. 1-TD complexes typically have only constant local-spectral expansion at low level links, so Proposition 11.7 does not, for instance, rule out showing $o_d(1)$-local-spectral HDX satisfy strong concentration. Using similar ideas to our degree lower bounds, we show at least some inverse polynomial local-spectral expansion is required to have strong concentration.

Proposition 11.8. *For any $c < \frac{1}{2}$ and large enough $d \in \mathbb{N}$, there exists a family $\{X_n\}$ of d^{-c}-two-sided local spectral expanders such that for any $\beta < 1$, $(X_n(d), X_n(1))$ is not a (d^{-c}, β)-additive sampler.*

The proof of Proposition 11.8 relies on another construction of Golowich [Gol23] building inverse polynomial HDX from Cayley complexes.

Theorem 11.9 ([Gol23]). *For every $c < \frac{1}{2}$ and large enough $d \in \mathbb{N}$, there exists a family of d-uniform, $2^{\Omega_d(\sqrt{\log(n)})}$-degree, $\frac{1}{d^c}$-two-sided local spectral expanders $\{X_n\}$ whose 1-skeletons are Cayley graphs over $\mathbb{F}_2^n$ with two-sided expansion $\lambda = \Theta(\frac{1}{d^c})$.*

We can now use Lemma 9.10 and the Cayley structure of X's 1-skeleton to prove Proposition 11.8.

Proof of Proposition 11.8. Since the 1-skeleton of X is a Cayley graph over $\mathbb{F}_2^n$, there is an eigenvector $g_0 : \mathbb{F}_2^n \to \{\pm 1\}$ whose eigenvalue is λ. Let us consider the indicator g of the set $A = \{v \in \mathbb{F}_2^n \mid g_0(v) = 1\}$ whose measure is $\mathbb{P}[A] = \frac{1}{2}$. This indicator is equal to $g = \frac{1}{2} + \frac{1}{2}g_0$. We note that for every v,

$$\mathbb{P}_{u \sim v}[A] = \mathbb{E}_{u \sim v}[g(u)] = \frac{1}{2} + \frac{1}{2}\mathbb{E}_{u \sim v}[g_0(u)].$$

The function g_0 is an eigenvalue so $\frac{1}{2}\mathbb{E}_{u \sim v}[g_0(u)] = \frac{\lambda}{2}g_0(v) = \pm\frac{\lambda}{2}$. Thus, $\mathbb{P}_{u \sim v}[A] - \mathbb{P}[A] = \pm\frac{\lambda}{2}$ so for every $\varepsilon \leqslant \frac{\lambda}{2}$ and any $\beta < 1$, the underlying graph *is not* a (ε, β)-sampler. By Lemma 9.10, this implies that either the containment graph is not a $(2\varepsilon, \beta)$-sampler, or the containment graph of some link of a vertex is not a $(2\varepsilon, \frac{1}{2})$-sampler. Since the vertex links of $\{X_n\}$ also give an infinite family of $\frac{1}{d^c}$-two-sided local spectral expanders this implies the result. □

This counterexample poses the following question: For $\varepsilon \in [d^{-c}, d^{-c/2}]$ what is the worst λ which still promises that the containment graph is a $(\varepsilon, \exp(-d^c))$-sampler for some $c > 0$?

11.3 Expander-Walks

Finally we take a step back from concentration itself and look at necessity of *local* concentration for our applications (that is concentration of the links of X and in particular the locally nice property). We argue this is at least in some sense necessary: classical samplers such as expander walks which behave poorly under restriction fail reverse hypercontractivity. Recall the expander walk complexes W_G defined in Section 10.1.

Proposition 11.10. *For any $0 < \gamma < 1$ and $\lambda > 0$, there are infinitely many pairs $k, n \in \mathbb{N}$ with a corresponding λ-expanders $G = ([n], E)$ and balanced subset $A \subset W_G(k)$ satisfying:*

$$\mathop{\mathbb{P}}_{t,t' \sim UD_{k,\gamma k}} \left[t' \in X(1) \setminus A \text{ and } t \in A \right] \leqslant 2^{-\Omega_\gamma(k)}$$

Proof. Take k odd and n even. It suffices to take any regular graph G with $\mathrm{girth}(G) > k$. Define A to be the set of walks whose center vertex is in $\{1, 2, \ldots, \frac{n}{2}\}$. Notice that A is indeed a balanced function, as every element is the center of the same number of random walks in a regular graph (see e.g. [Ale+20] for details).

On the other hand, it is easy to see A has very poor expansion due to the girth of G. In particular, the only way to leave A (resp. $X(1) \setminus A$) is if the walk re-samples all of $[k/2 + 1]$, or all of $[k] \setminus [k/2]$. If neither event occurs, there exist indices $i < \lceil \frac{k}{2} \rceil < j$ which were not re-sampled, and the girth of the graph ensures there is only one option for the $\lceil \frac{k}{2} \rceil$th element (namely its starting position). One can compute directly the probability either event occurs is at most $2^{-\Omega_\gamma(k)}$.

To complete the proof, we need an infinite family of regular λ-expanders for any $\lambda > 0$ with super-constant girth. Many such constructions exist in the literature, including the classical Ramanujan graphs of [LPS88]. □

Note that for any fixed $\gamma, C, \ell > 0$, taking k in Proposition 11.10 sufficiently large means we can always find subsets $A, B \subset X(k)$ such that

$$\mathop{\mathbb{P}}_{t,t' \sim UD_{k,\gamma k}} [t \in A \land t' \in B] \leqslant C \, \mathbb{P}[A]^\ell \, \mathbb{P}[B]^\ell,$$

violating reverse hypercontractivity for all parameter settings. We remark that the failure of expander walks in agreement testing is well known, and can be derived e.g. from [GC19].

Acknowledgements

References

[Che52]   Herman Chernoff. "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations." In: *The Annals of Mathematical Statistics* (1952), pp. 493–507.

[For65]   G David Forney. "Concatenated codes." In: (1965).

[Gar73]   Howard Garland. "p-Adic Curvature and the Cohomology of Discrete Subgroups of p-Adic Groups." In: *Ann. of Math.* 97.3 (1973), pp. 375–423. DOI: 10.2307/1970829.

[McE+77] Robert McEliece, Eugene Rodemich, Howard Rumsey, and Lloyd Welch. "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities." In: *IEEE transactions on Information Theory* 23.2 (1977), pp. 157–166.

[Bár82]   Imre Bárány. "A generalization of Carathéodory's theorem." In: *Discrete Mathematics* 40.2-3 (1982), pp. 141–152.

[Bor82]   Christer Borell. "Positivity improving operators and hypercontractivity." In: *Mathematische Zeitschrift* 180 (1982), pp. 225–234.

[BF84]   Endre Boros and Zoltán Füredi. "The number of triangles covering the center of an n-set." In: *Geometriae Dedicata* 17 (1984), pp. 69–77.

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

[FR87]   Peter Frankl and Vojtěch Rödl. "Forbidden intersections." In: *Transactions of the American Mathematical Society* 300.1 (1987), pp. 259–286.

[LPS88]   Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. "Ramanujan graphs." In: *Combinatorica* 8.3 (1988), pp. 261–277.

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

[BLR90]   Manuel Blum, Michael Luby, and Ronitt Rubinfeld. "Self-testing/correcting with applications to numerical problems." In: *Proceedings of the twenty-second annual ACM symposium on Theory of computing.* 1990, pp. 73–83.

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

[Nil91]   Alon Nilli. "On the second eigenvalue of a graph." In: *Discrete Mathematics* 91.2 (1991), pp. 207–210.

[Alo+92]   Noga Alon, Jehoshua Bruck, Joseph Naor, Moni Naor, and Ron M Roth. "Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs." In: *IEEE Transactions on information theory* 38.2 (1992), pp. 509–516.

[AS94]   Shigeki Aida and Daniel Stroock. "Moment estimates derived from Poincaré and logarithmic Sobolev inequalities." In: *Mathematical Research Letters* 1.1 (1994), pp. 75–86.

[Hoe94]   Wassily Hoeffding. "Probability inequalities for sums of bounded random variables." In: *The collected works of Wassily Hoeffding* (1994), pp. 409–426.

[CEG95]   Ran Canetti, Guy Even, and Oded Goldreich. "Lower bounds for sampling algorithms for estimating the average." In: *Information Processing Letters* 53.1 (1995), pp. 17–25. ISSN: 0020-0190. DOI: https://doi.org/10.1016/0020-0190(94)00171-T.

[Raz95]   Ran Raz. "A parallel repetition theorem." In: *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing.* 1995, pp. 447–456.

[SSS95]   Jeanette P Schmidt, Alan Siegel, and Aravind Srinivasan. "Chernoff–Hoeffding bounds for applications with limited independence." In: *SIAM Journal on Discrete Mathematics* 8.2 (1995), pp. 223–250.

[AL96]   Noga Alon and Michael Luby. "A linear time erasure-resilient code with nearly optimal recovery." In: *IEEE Transactions on Information Theory* 42.6 (1996), pp. 1732–1736.

[DS96]   Persi Diaconis and Laurent Saloff-Coste. "Logarithmic Sobolev inequalities for finite Markov chains." In: *The Annals of Applied Probability* 6.3 (1996), pp. 695–750.

[DR96]   Devdatt P Dubhashi and Desh Ranjan. "Balls and bins: A study in negative dependence." In: *BRICS Report Series* 3.25 (1996).

[RS96]   Ronitt Rubinfeld and Madhu Sudan. "Robust Characterizations of Polynomials with Applications to Program Testing." In: *SIAM J. Comput.* 25.2 (1996), pp. 252–271.

[AS97]   Sanjeev Arora and Madhu Sudan. "Improved Low Degree Testing and its Applications." In: *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing.* El Paso, Texas, 1997, pp. 485–495.

[BL97]   Sergey Bobkov and Michel Ledoux. "Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution." In: *Probability Theory and Related Fields* 107 (1997), pp. 383–400.

[Gol97]   Oded Goldreich. "A sample of samplers: A computational perspective on sampling." In: *def* 1 (1997), 2n.

[GS97]   Oded Goldreich and Shmuel Safra. "A Combinatorial Consistency Lemma with Application to Proving the PCP Theorem." In: *RANDOM: International Workshop on Randomization and Approximation Techniques in Computer Science.* LNCS, 1997.

[PS97]   Alessandro Panconesi and Aravind Srinivasan. "Randomized Distributed Edge Coloring via an Extension of the Chernoff–Hoeffding Bounds." In: *SIAM Journal on Computing* 26.2 (1997), pp. 350–368. DOI: 10 . 1137 / S0097539793250767. eprint: https : / / doi . org / 10 . 1137 / S0097539793250767. URL: https://doi.org/10.1137/S0097539793250767.

[RS97]   Ran Raz and Shmuel Safra. "A Sub-constant Error-probability Low-degree Test, and a Sub-constant Error-probability PCP Characterization of NP." In: *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing.* STOC '97. El Paso, Texas, USA: ACM, 1997, pp. 475–484. ISBN: 0-89791-888-6. DOI: 10.1145/258533.258641. URL: http://doi.acm.org/10.1145/258533.258641.

[Gil98]   David Gillman. "A Chernoff Bound for Random Walks on Expander Graphs." In: *SIAM Journal on Computing* 27.4 (1998), pp. 1203–1220. DOI: 10.1137/S0097539794268765. eprint: https://doi.org/10.1137/S0097539794268765. URL: https://doi.org/10.1137/S0097539794268765.

[Pac98]   János Pach. "A Tverberg-type result on multicolored simplices." In: *Computational Geometry* 10.2 (1998), pp. 71–76.

[KY99]   Philip Klein and Neal Young. "On the number of iterations for Dantzig-Wolfe optimization and packing-covering approximation algorithms." In: *International Conference on Integer Programming and Combinatorial Optimization.* Springer. 1999, pp. 320–327.

[Pem01]   Sriram V Pemmaraju. "Equitable coloring extends Chernoff-Hoeffding bounds." In: *International Workshop on Randomization and Approximation Techniques in Computer Science.* Springer. 2001, pp. 285–296.

[Sha04]   Ronen Shaltiel. "Recent developments in explicit constructions of extractors." In: *Current Trends in Theoretical Computer Science: The Challenge of the New Century Vol 1: Algorithms and Complexity Vol 2: Formal Models and Semantics* (2004), pp. 189–228.

[LSV05a]   Alexander Lubotzky, Beth Samuels, and Uzi Vishne. "Explicit constructions of Ramanujan complexes of type $\tilde{A}_d$." In: *European J. Combin.* 26.6 (2005), pp. 965–993. DOI: 10.1016/j.ejc.2004.06.007. eprint: math/0406217.

[LSV05b]   Alexander Lubotzky, Beth Samuels, and Uzi Vishne. "Ramanujan complexes of type $\tilde{A}_d$." In: *Israel J. Math.* 149.1 (2005), pp. 267–299. DOI: 10.1007/BF02772543. eprint: math/0406208.

[MOO05]   Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. "Noise stability of functions with low influences: invariance and optimality." In: *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05).* IEEE. 2005, pp. 21–30.

[DR06]   Irit Dinur and Omer Reingold. "Assignment testers: Towards a combinatorial proof of the PCP theorem." In: *SIAM Journal on Computing* 36.4 (2006), pp. 975–1024.

[HLW06]   Shlomo Hoory, Nathan Linial, and Avi Wigderson. "Expander graphs and their applications." In: *Bulletin of the American Mathematical Society* 43.4 (2006), pp. 439–561.

[LM06]   Nathan Linial and Roy Meshulam. "Homological Connectivity Of Random 2-Complexes." In: *Combinatorica* 26 (2006), pp. 475–487. DOI: 10.1007/s00493-006-0027-9.

[Mos+06]   Elchanan Mossel, Ryan O'Donnell, Oded Regev, Jeffrey E Steif, and Benny Sudakov. "Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality." In: *Israel Journal of Mathematics* 154.1 (2006), pp. 299–336.

[Din07]   Irit Dinur. "The PCP theorem by gap amplification." In: *Journal of the ACM (JACM)* 54.3 (2007), 12–es.

[FKO07]   Uriel Feige, Guy Kindler, and Ryan O'Donnell. "Understanding parallel repetition requires understanding foams." In: *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*. IEEE. 2007, pp. 179–192.

[DG08a]   Irit Dinur and Elazar Goldenberg. "Locally testing direct product in the low error range." In: *2008 49th Annual IEEE Symposium on Foundations of Computer Science*. IEEE. 2008, pp. 613–622.

[DG08b]   Irit Dinur and Elazar Goldenberg. "Locally testing direct products in the low error range." In: *Proc. 49th IEEE Symp. on Foundations of Computer Science*. 2008.

[Hea08]   Alexander D Healy. "Randomness-efficient sampling within nc." In: *Computational Complexity* 17 (2008), pp. 3–37.

[Imp+08]   Russell Impagliazzo, Ragesh Jaiswal, Valentine Kabanets, and Avi Wigderson. "Uniform direct product theorems: simplified, optimized, and derandomized." In: *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*. 2008, pp. 579–588.

[IJK09]   Russell Impagliazzo, Ragesh Jaiswal, and Valentine Kabanets. "Approximate list-decoding of direct product codes and uniform hardness amplification." In: *SIAM Journal on Computing* 39.2 (2009), pp. 564–605.

[She09]   Jonah Sherman. "Breaking the multicommodity flow barrier for $O(\sqrt{logn})$-approximations to sparsest cut." In: *2009 50th Annual IEEE Symposium on Foundations of Computer Science*. IEEE. 2009, pp. 363–372.

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

[Gro10]   M. Gromov. "Singularities, Expanders and Topology of Maps. Part 2: from Combinatorics to Topology Via Algebraic Isoperimetry." In: *Geom. Funct. Anal.* 20 (2010), pp. 416–526. DOI: 10.1007/s00039-010-0073-8.

[DM11]   Irit Dinur and Or Meir. "Derandomized parallel repetition via structured PCPs." In: *computational complexity* 20 (2011), pp. 207–327.

[Ben12]   Siavosh Benabbas. *On hierarchies of strong SDP relaxations for combinatorial optimization problems*. University of Toronto (Canada), 2012.

[Fox+12]   Jacob Fox, Mikhail Gromov, Vincent Lafforgue, Assaf Naor, and János Pach. "Overlap properties of geometric expanders." In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 2012.671 (2012), pp. 49–83.

[IKW12]   Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. "New Direct-Product Testers and 2-Query PCPs." In: *SIAM Journal on Computing* 41.6 (2012), pp. 1722–1768.

[Kel12]   Nathan Keller. "A tight quantitative version of Arrow's impossibility theorem." In: *Journal of the European Mathematical Society* 14.5 (2012), pp. 1331–1355.

[Mos12]   Elchanan Mossel. "A quantitative Arrow theorem." In: *Probability Theory and Related Fields* 154.1-2 (2012), pp. 49–88.

[MR12]   Elchanan Mossel and Miklós Z Rácz. "A quantitative Gibbard-Satterthwaite theorem without neutrality." In: *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. 2012, pp. 1041–1060.

[BLM13]   Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Feb. 2013. ISBN: 9780199535255. DOI: 10.1093/acprof:oso/9780199535255.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780199535255.001.0001.

[MOS13]   Elchanan Mossel, Krzysztof Oleszkiewicz, and Arnab Sen. "On reverse hypercontractivity." In: *Geometric and Functional Analysis* 23.3 (2013), pp. 1062–1097.

[DS14]   I. Dinur and D. Steurer. "Direct Product Testing." In: *2014 IEEE 29th Conference on Computational Complexity (CCC)*. June 2014, pp. 188–196.

[Kau+14]   Manuel Kauers, Ryan O'Donnell, Li-Yang Tan, and Yuan Zhou. "Hypercontractive inequalities via SOS, and the Frankl–Rödl graph." In: *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*. SIAM. 2014, pp. 1644–1658.

[KKL14]   Tali Kaufman, David Kazhdan, and Alexander Lubotzky. "Ramanujan Complexes and Bounded Degree Topological Expanders." In: *Proc.* 55*th IEEE Symp. on Foundations of Comp. Science (FOCS)*. 2014, pp. 484–493. DOI: 10.1109/FOCS.2014.58. eprint: 1408.6351.

[KL14]   Tali Kaufman and Alexander Lubotzky. "High dimensional expanders and property testing." In: *Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12-14, 2014*. 2014, pp. 501–506.

[ODo14]   Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

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

[EK16]   Shai Evra and Tali Kaufman. "Bounded degree cosystolic expanders of every dimension." In: *Proc.* 48*th ACM Symp. on Theory of Computing (STOC)*. 2016, pp. 36–48. DOI: 10.1145/2897518.2897543. eprint: 1510.00839.

[PRT16]   Ori Parzanchevski, Ron Rosenthal, and Ran J Tessler. "Isoperimetric inequalities in simplicial complexes." In: *Combinatorica* 36.2 (2016), pp. 195–227.

[DK17]   Irit Dinur and Tali Kaufman. "High dimensional expanders imply agreement expanders." In: *Proc.* 58*th IEEE Symp. on Foundations of Comp. Science (FOCS)*. 2017, pp. 974–985. DOI: 10.1109/FOCS.2017.94.

[DL17]   Irit Dinur and Inbal Livni Navon. "Exponentially Small Soundness for the Direct Product Z-Test." In: *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*. 2017, 29:1–29:50.

[Evr17]   Shai Evra. "Finite quotients of Bruhat–Tits buildings as geometric expanders." In: *Journal of Topology and Analysis* 9.01 (2017), pp. 51–66.

[KM17]   Tali Kaufman and David Mass. "High dimensional random walks and colorful expansion." In: *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2017.

[KMS17]   Subhash Khot, Dor Minzer, and Muli Safra. "On independent sets, 2-to-2 games, and Grassmann graphs." In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*. 2017, pp. 576–589.

[Kop+17]   Swastik Kopparty, Or Meir, Noga Ron-Zewi, and Shubhangi Saraf. "High-rate locally correctable and locally testable codes with sub-polynomial query complexity." In: *Journal of the ACM (JACM)* 64.2 (2017), pp. 1–42.

[LP17]   David A Levin and Yuval Peres. *Markov chains and mixing times*. Vol. 107. American Mathematical Soc., 2017.

[Mos17]   Dana Moshkovitz. "Low-degree test with polynomially small error." In: *computational complexity* 26 (2017), pp. 531–582.

[Dik+18]   Yotam Dikstein, Irit Dinur, Yuval Filmus, and Prahladh Harsha. "Boolean functions on high-dimensional expanders." In: *arXiv preprint arXiv:1804.08155* (2018).

[Din+18a]   Irit Dinur, Prahladh Harsha, Tali Kaufman, Inbal Livni Navon, and Amnon Ta-Shma. "List Decoding with Double Samplers." In: *CoRR* abs/1808.00425 (2018). arXiv: 1808.00425. URL: http://arxiv.org/abs/1808.00425.

[Din+18b] 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).* (To appear). 2018.

[KO18] Tali Kaufman and Izhar Oppenheim. "Construction of new local spectral high dimensional expanders." In: *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018.* 2018, pp. 773–786.

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

[Lub18] Alexander Lubotzky. "High dimensional expanders." In: *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018.* World Scientific. 2018, pp. 705–730.

[Opp18] Izhar Oppenheim. "Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps." In: *Discrete & Computational Geometry* 59.2 (2018), pp. 293–330.

[AJT19] Vedat Alev, Fernando Jeronimo, and Madhur Tulsiani. "Approximating Constraint Satisfaction Problems on High-Dimensional Expanders." In: *Proc. 59th IEEE Symp. on Foundations of Computer Science.* Nov. 2019, pp. 180–201. DOI: `10.1109/FOCS.2019.00021`.

[Ana+19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. "Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid." In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing.* 2019, pp. 1–12.

[BKS19] Boaz Barak, Pravesh K. Kothari, and David Steurer. "Small-Set Expansion in Shortcode Graph and the 2-to-2 Conjecture." In: *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA.* 2019, 9:1–9:12.

[CGM19] Mary Cryan, Heng Guo, and Giorgos Mousa. "Modified log-Sobolev inequalities for strongly log-concave distributions." In: *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS).* IEEE. 2019, pp. 1358–1370.

[DD19] Yotam Dikstein and Irit Dinur. "Agreement testing theorems on layered set systems." In: *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science.* 2019.

[DFH19] Irit Dinur, Yuval Filmus, and Prahladh Harsha. "Analyzing Boolean functions on the biased hypercube via higher-dimensional agreement tests: [Extended abstract]." In: *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019.* 2019, pp. 2124–2133.

[GC19] Elazar Goldenberg and Karthik CS. "Toward a General Direct Product Testing Theorem." In: *ACM Transactions on Computation Theory (TOCT)* 12.1 (2019), pp. 1–18.

[Ale+20] Vedat Levi Alev, Fernando Jeronimo Jeronimo, Dylan Quintana, Shashank Srivastava, and Madhur Tulsiani. "List decoding of direct sum codes." In: *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms.* SIAM. 2020, pp. 1412–1425.

[AL20] Vedat Levi Alev and Lap Chi Lau. "Improved analysis of higher order random walks and applications." In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing.* 2020, pp. 1198–1211.

[Dik+20] Yotam Dikstein, Irit Dinur, Prahladh Harsha, and Noga Ron-Zewi. *Locally testable codes via high-dimensional expanders.* 2020. eprint: `arXiv:2005.01045`.

[FI20] Ehud Friedgut and Yonatan Iluz. *Hyper-regular graphs and high dimensional expanders.* 2020. eprint: `arXiv:2010.03829`.

[GM20] Heng Guo and Giorgos Mousa. "Local-to-global contraction in simplicial complexes." In: *arXiv preprint arXiv:2012.14317* (2020).

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

[KO20]   Tali Kaufman and Izhar Oppenheim. "High order random walks: Beyond spectral gap." In: *Combinatorica* 40 (2020), pp. 245–281.

[KS20]   Tali Kaufman and Ella Sharakanski. "Chernoff Bound for High-Dimensional Expanders." In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2020, August 17-19, 2020, Virtual Conference*. Ed. by Jaroslaw Byrka and Raghu Meka. Vol. 176. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020, 25:1–25:22. DOI: `10.4230/LIPICS.APPROX/RANDOM.2020.25`. URL: `https://doi.org/10.4230/LIPIcs.APPROX/RANDOM.2020.25`.

[LMY20]   Siqi Liu, Sidhanth Mohanty, and Elizabeth Yang. "High-Dimensional Expanders from Expanders." In: (2020).

[Alo21]   Noga Alon. "Explicit expanders of every degree and size." In: *Combinatorica* (2021), pp. 1–17.

[ALG21]   Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. "Spectral independence in high-dimensional expanders and applications to the hardcore model." In: *SIAM Journal on Computing* 0 (2021), FOCS20–1.

[Che+21]   Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. "Rapid mixing for colorings via spectral independence." In: *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2021, pp. 1548–1557.

[CLV21]   Zongchen Chen, Kuikui Liu, and Eric Vigoda. "Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion." In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. 2021, pp. 1537–1550.

[Gol21]   Louis Golowich. "Improved Product-Based High-Dimensional Expanders." In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021)*. Vol. 207. 2021. DOI: `10.4230/LIPIcs.APPROX/RANDOM.2021.38`.

[JST21]   Fernando Granha Jeronimo, Shashank Srivastava, and Madhur Tulsiani. "Near-linear time decoding of ta-shma's codes via splittable regularity." In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. 2021, pp. 1527–1536.

[KKS21]   Tali Kaufman, Rasmus Kyng, and Federico Soldá. "Scalar and Matrix Chernoff Bounds from ℓ_∞-Independence." In: *arXiv preprint arXiv:2111.02177* (2021).

[KT21]   Tali Kaufman and Ran J Tessler. "Garland's Technique for Posets and High Dimensional Grassmannian Expanders." In: *arXiv preprint arXiv:2101.12621* (2021).

[Kee+21]   Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. "Forbidden intersections for codes." In: *arXiv preprint arXiv:2103.05050* (2021).

[Liu21]   Kuikui Liu. "From coupling to spectral independence and blackbox comparison with the down-up walk." In: *arXiv preprint arXiv:2103.11609* (2021).

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

[Baf+22a] Mitali Bafna, Max Hopkins, Tali Kaufman, and Shachar Lovett. "High Dimensional Expanders: Eigenstripping, Pseudorandomness, and Unique Games." In: *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2022, pp. 1069–1128.

[Baf+22b] Mitali Bafna, Max Hopkins, Tali Kaufman, and Shachar Lovett. "Hypercontractivity on high dimensional expanders." In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. 2022, pp. 185–194.

[BD22]   Guy Blanc and Dean Doron. "New near-linear time decodable codes closer to the GV bound." In: *37th Computational Complexity Conference (CCC 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2022.

[Bla+22]   Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Štefankovič, and Eric Vigoda. "On Mixing of Markov Chains: Coupling, Spectral Independence, and Entropy Factorization." In: *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2022, pp. 3670–3692.

[Dik22]   Yotam Dikstein. "New High Dimensional Expanders from Covers." 2022.

[Din+22a]   Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. "Locally testable codes with constant rate, distance, and locality." In: *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*. Ed. by Stefano Leonardi and Anupam Gupta. ACM, 2022, pp. 357–374. DOI: 10.1145/3519935.3520024. URL: https://doi.org/10.1145/3519935.3520024.

[Din+22b]   Irit Dinur, Prahladh Harsha, Tali Kaufman, and Noga Ron-Zewi. "From Local to Robust Testing via Agreement Testing." In: *Theory of Computing* 18.1 (2022), pp. 1–25. DOI: 10.4086/toc.2022.v018a012. URL: https://doi.org/10.4086/toc.2022.v018a012.

[DW22]   Dean Doron and Mary Wootters. "High-probability list-recovery, and applications to heavy hitters." In: *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2022.

[Fen+22]   Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. "Rapid mixing from spectral independence beyond the Boolean domain." In: *ACM Transactions on Algorithms (TALG)* 18.3 (2022), pp. 1–32.

[Gai+22]   Jason Gaitonde, Max Hopkins, Tali Kaufman, Shachar Lovett, and Ruizhe Zhang. "Eigenstripping, Spectral Decay, and Edge-Expansion on Posets." In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques* (2022).

[GK22]   Roy Gotlib and Tali Kaufman. *List Agreement Expansion from Coboundary Expansion*. 2022. eprint: arXiv:2210.15714.

[GLL22]   Tom Gur, Noam Lifshitz, and Siqi Liu. "Hypercontractivity on high dimensional expanders." In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. 2022, pp. 176–184.

[HL22]   Max Hopkins and Ting-Chun Lin. "Explicit Lower Bounds Against \ Omega (n)-Rounds of Sum-of-Squares." In: *arXiv preprint arXiv:2204.11469* (2022).

[PK22]   Pavel Panteleev and Gleb Kalachev. "Asymptotically good Quantum and locally testable classical LDPC codes." In: *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*. Ed. by Stefano Leonardi and Anupam Gupta. ACM, 2022, pp. 375–388. DOI: 10.1145/3519935.3520017. URL: https://doi.org/10.1145/3519935.3520017.

[AP23]   Vedat Levi Alev and Ori Parzanchevski. "Sequential Sweeps and High Dimensional Expansion." In: *arXiv preprint arXiv:2312.02089* (2023).

[Ana+23]   Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. "Universality of Spectral Independence with Applications to Fast Mixing in Spin Glasses." In: *arXiv preprint arXiv:2307.10466* (2023).

[ABN23]   Anurag Anshu, Nikolas P Breuckmann, and Chinmay Nirkhe. "NLTS Hamiltonians from good quantum codes." In: *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. 2023, pp. 1090–1096.

[BM23a]   Mitali Bafna and Dor Minzer. "Characterizing Direct Product Testing via Coboundary Expansion." 2023. URL: https://arxiv.org/pdf/2308.09668.pdf.

[BM23b]   Mitali Bafna and Dor Minzer. "Characterizing Direct Product Testing via Coboundary Expansion." In: *arXiv preprint arXiv:2308.09668* (2023).

[Bea+23]   Daniel Beaglehole, Max Hopkins, Daniel Kane, Sihan Liu, and Shachar Lovett. "Sampling Equilibria: Fast No-Regret Learning in Structured Games." In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2023, pp. 3817–3855.

[Cha+23]   Zachary Chase, Bogdan Chornomaz, Shay Moran, and Amir Yehudayoff. "Local Borsuk-Ulam, Stability, and Replicability." In: *arXiv preprint arXiv:2311.01599* (2023).

[DD23a]   Yotam Dikstein and Irit Dinur. "Agreement theorems for high dimensional expanders in the small soundness regime: the role of covers." In: *CoRR* abs/2308.09582 (2023). DOI: `10.48550/ARXIV.2308.09582`. arXiv: `2308.09582`. URL: `https://doi.org/10.48550/arXiv.2308.09582`.

[DD23b]   Yotam Dikstein and Irit Dinur. "Agreement theorems for high dimensional expanders in the small soundness regime: the role of covers." In: *arXiv preprint arXiv:2308.09582* (2023).

[DLZ23]   Irit Dinur, Siqi Liu, and Rachel Yun Zhang. "New Codes on High Dimensional Expanders." In: *arXiv preprint arXiv:2308.15563* (2023).

[Gol23]   Louis Golowich. "From Grassmannian to Simplicial High-Dimensional Expanders." In: *arXiv preprint arXiv:2305.02512* (2023).

[GK23a]   Roy Gotlib and Tali Kaufman. "Fine grained analysis of high dimensional random walks." In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2023.

[GK23b]   Roy Gotlib and Tali Kaufman. "List Agreement Expansion from Coboundary Expansion." In: *14th Innovations in Theoretical Computer Science Conference, ITCS 2023*. Ed. by Yael Tauman Kalai. Leibniz International Proceedings in Informatics, LIPIcs. Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, Jan. 2023. DOI: `10.4230/LIPIcs.ITCS.2023.61`.

[JST23]   Fernando Granha Jeronimo, Shashank Srivastava, and Madhur Tulsiani. "List Decoding of Tanner and Expander Amplified Codes from Distance Certificates." In: *arXiv preprint arXiv:2311.02263* (2023).

[Lee23]   Holden Lee. "Parallelising Glauber dynamics." In: *arXiv preprint arXiv:2307.07131* (2023).

[BLM24]   Mitali Bafna, Noam Lifshitz, and Dor Minzer. *Constant Degree Direct Product Testers with Small Soundness*. 2024. eprint: `arXiv:2402.00850`.

[DDL24]   Yotam Dikstein, Irit Dinur, and Alexander Lubotzky. *Low Acceptance Agreement Tests via Bounded-Degree Symplectic HDXs*. 2024. URL: `https://eccc.weizmann.ac.il/report/2024/019/`.

[Hop24]   Max Hopkins. "Improved Hypercontractivity on High Dimensional Expanders from Coordinate-Wise Induction." 2024.

A   Concentration for the Complete Complex

In this section we show near-optimal concentration bounds for the complete complex.

Claim A.1. Let X be the complete complex on n vertices, that is $X(k) = \binom{[n]}{k}$. Let $\alpha, \delta > 0$. Then for any $\ell \leqslant k$, the containment graph $G = (X(k), X(\ell))$ is an $(\alpha, \frac{4}{\alpha\delta}\exp(-\frac{\delta^2}{12}\alpha\lfloor\frac{k}{\ell}\rfloor), \delta)$-multiplicative sampler and $(\varepsilon, \frac{2}{\varepsilon}\exp(-\frac{\varepsilon^2}{8}\lfloor\frac{k}{\ell}\rfloor))$-additive sampler for any $\varepsilon > 0$.

Proof. Fix and $A \subseteq X(\ell)$ such that $\mathbb{P}[A] \geqslant \alpha$. We apply Lemma 5.9 in the following distribution. The distribution we use samples $(s_1, s_2, \ldots, s_m, t) \sim D$ such that $(s_1, s_2, \ldots, s_m)$ are independent and $t \in X(k)$ is a uniform face conditioned on containing all s_i. Obviously, (s_i, t) is a uniform pair of ℓ and k faces where $s_i \subseteq t$. Moreover, by Chernoff's bound for independent sampling

$$\mathbb{P}_{(s_1, s_2, \ldots, s_m, t) \sim D}\left[\left|\frac{|s_i \in A|}{m} - \mathbb{P}[A]\right| > \frac{\delta}{2}\mathbb{P}[A]\right] \leqslant 2\exp(\frac{\delta^2}{12}\alpha m).$$

The claim follows from Lemma 5.9. The additive bound follows by the same argument applying Hoeffding's inequality.   □

Note that we do not restrict the number of vertices n in this claim. Since the k-skeleton of the complete complex is only an $\frac{1}{n-k}$-two sided spectral expander, this claim does not follow from Theorem 5.1.

B   Concentration for the Swap Complex

In this appendix we prove near-optimal concentration of measure for Lipshitz functions on the swap complex. Recall the swap complex $C = C_{\ell,k,n}$ is the $d = \lfloor \frac{k}{\ell} \rfloor$-uniform complex whose vertices are $C(0) = \binom{[n]}{\ell}$ and whose top level faces are $C(d) = \left\{ \{s_1, s_2, \ldots, s_d\} \mid \forall i \neq j, \ s_i \cap s_j = \emptyset \right\}$ endowed with the uniform distribution. We prove subgaussian concentration for the swap complex (under mild assumptions on n) up to a $\log(k)$ factor of optimal.

Theorem B.1 (Theorem 7.15 Restated)**.** *Let $C = C_{\ell,k,n}$ for $n \geqslant 3k$ and $f : C(d) \to \mathbb{R}$ be any ν-bounded function. Then:*

1. ***Upper tail:*** $\mathbb{P}[f - \mathbb{E}[f] > t] \leqslant \exp(\frac{t^2}{\nu} \cdot \frac{1}{2\log(ek)})$

2. ***Lower tail:*** $\mathbb{P}[f - \mathbb{E}[f] < -t] \leqslant \exp(\frac{t^2}{\nu} \cdot \frac{1}{2\log(ek)})$

By the Herbst argument (see e.g. [BLM13] for the variant on the cube or [CGM19] for a version for general Markov chains), Theorem 7.15 is an immediate corollary on the following bound of the log-sobolev constant of the swap complex:

Proposition B.2. *Let $\ell, k, n \in \mathbb{N}$ and $d = \lfloor \frac{k}{\ell} \rfloor$. The Log-Sobolev Constant of the down-up walk on $C_{\ell,k,n}$ satisfies:*

$$\rho_{LS} \geqslant \frac{1}{d+1} \cdot \frac{1}{2\log(ek)}$$

whenever $n \geqslant 3k$.

Proof. Let $m = \lfloor \frac{k}{\ell} \rfloor$ for notational simplicity. We first remark that it is sufficient to bound the log-sobolev constant of the partitification of S, since for every function $f : S \to \mathbb{R}$ we have

$$\frac{\langle f, L_S f \rangle}{Ent(f^2)} = \frac{\langle \tilde{f}, L_{P(S)} \tilde{f} \rangle}{Ent(f^2)}$$

where L_S is the Laplacian on S, $L_{P(S)}$ is the laplacian on $P(S)$, and $\tilde{f}$ is the partitification of f on $P(S)$.

We bound the log-sobolev constant via the following bound on the mixing time of the down-up walk:

$$t_{mix} \leqslant O\left(m \log k\right),$$

where t_{mix} is the minimum number of steps of the down-up walk such that any starting distribution is within $\frac{1}{e}$ of uniform in TV-distance. Diaconis and Saloff-Coste [DS96] showed $\rho_{LS} \geqslant \frac{1}{2t_{mix}}$, so this is sufficient to prove the result.

To bound t_{mix}, we appeal to the coupling method which we recall below for completeness:

Definition B.3. Let $G = (V, E)$ be a graph. A Markovian coupling of G is a Markov chain $(X, Y) = \{(X_i, Y_i)\}_{i=0}^{\infty}$ such that for all $i \geqslant 0$, $X_i, Y_i \in V$ and for all $u_1, u_2, v_1, v_2 \in V$

$$\mathbb{P}\left[X_{i+1} = u_2 \mid X_i = u_1, Y_i = v_1\right] = \mathop{\mathbb{P}}_{v \sim u_1}\left[v = u_2\right],$$

$$\mathbb{P}\left[Y_{i+1} = v_2 \mid X_i = u_1, Y_i = v_1\right] = \mathop{\mathbb{P}}_{v \sim v_1}\left[v = v_2\right],$$

where the right-hand probabilities are over the underlying random walk on G.

In other words, X and Y marginally evolve as the underlying walk on G, but may have any desired dependencies outside this condition. It is well known that the mixing time of G is upper bounded by the best *collision time* over Markovian couplings.

Theorem B.4 (See e.g. [LP17, Corollary 5.5]). *Let G be a graph and (X, Y) a Markovian coupling of G such that for all s, if $X_s = Y_s$ then for all $t \geqslant s$ $X_t = Y_t$. Let $Q = \min\{s \geqslant 1 \mid X_s = Y_s\}$. Let t_{col} be the minimum t such that from all starting conditions x_0, y_0:*

$$\mathop{\mathbb{P}}_{(X,Y)}[X_t = Y_t | X_0 = x_0, Y_0 = y_0] \geqslant \frac{1}{e},$$

then $t_{mix} \leqslant t_{col}$.

Let (X, Y) be copies of the down-up walk on $P(S)$ coupled in the following fashion. Starting at state (s^1, s^2), for each coordinate $i \in [m-1]$ let K_i denote the set of shared free elements of the two states without their ith coordinate, that is:

$$K_i = [n] \setminus \{s^1_{-i} \cup s^2_{-i}\}$$

We couple X and Y by ensuring they draw the same elements from K_i. In particular, consider the following update rule for state (s^1, s^2):

1. Choose a random coordinate $i \in [m]$

2. Draw $w \sim \mathrm{Hyp}(|K_i|, n - k + \ell)$

3. Sample $t \sim \binom{K_i}{w}$, $t_X \sim \binom{[n] \setminus (K_i \cup s^1_{-i})}{\ell - w}$, and $t_Y \sim \binom{[n] \setminus (K_i \cup s^2_{-i})}{\ell - w}$

4. Set $s_{X(i)} = t \cup t_X$ and $s_Y(i) = t \cup t_Y$

Here $\mathrm{Hyp}(a, b)$ is the hypergeomtric distribution on b vertices with a 'marked' values (i.e. sampling without replacement). Observe that, the marginals s^1_i and s^2_i are uniform ℓ-sets sampled from the remaining disjoint support after removing the i-th coordinate, itself uniformly random. Thus (X, Y) is a valid Markovian coupling. Clearly it is also the case that once $s^1 = s^2$, all further updates remain unchanged, so by Theorem B.4 it is enough to show that from any starting conditions the probability that $s^1 = s^2$ after $O(m \log k)$ rounds is at least $\frac{1}{e}$.

We show this fact by tracking the expected disagreement in each coordinate across rounds of the coupled walk. Toward this end, for every $t \geqslant 0$ and $i \in [m]$, define the variable $Z_t^{(i)}$ to denote the fraction of disagreeing vertices between s^1_i and s^2_i in step t, that is

$$Z_t^{(i)} = 1 - \frac{|s^1_i \cap s^2_i|}{\ell},$$

and define Z_t to be the 'total' or 'average' disagreement:

$$Z_t = \frac{1}{m} \sum_{i \in [m]} Z_t^{(i)}.$$

If we can show that $\mathbb{E}[Z_t] < \frac{1}{ek}$ for some $t \leqslant O(m \log k)$ we are done by Markov's inequality, since $Z_t < \frac{1}{k}$ if and only if $s^1 = s^2$ and $t_{mix} \leqslant t_{col} \leqslant O(m \log k)$.

To analyze Z_t, we look at the expected difference $R_t^{(i)} = Z_t^{(i)} - Z_{t-1}^{(i)}$ in disagreement on coordinate j between step t and $t-1$. Define $R_t = \frac{1}{m} \sum_{i=1}^{m} R_t^{(i)}$ to be the average difference, and observe by telescoping and linearity of expectation we have:

$$\mathbb{E}[Z_t] = Z_0 + \sum_{j=1}^{t} \mathbb{E}[R_j]$$

The key is now to observe that we can bound $\mathbb{E}[R_j]$ as a function of $\mathbb{E}[Z_{j-1}]$, setting up a recursion on $\mathbb{E}[Z_t]$. To see this, fix the process up to step $j-1$, and consider the difference on coordinate i over the randomness of step j. If i is not sampled in Step (1), then $Z_j^{(i)} = Z_{j-1}^{(i)}$ and $R_j^{(i)} = 0$. On the other hand with probability $\frac{1}{m}$, coordinate i is sampled and the new disagreement $Z_j^{(i)}$ is drawn from a hypergeometric distribution:

$$\ell \cdot Z_j^{(i)} \sim Hyp(p, n - k + \ell)$$

106

where $p \leqslant k \cdot Z_{j-1}$, the worst-case number of unshared free vertices. As such, we can bound the mean of $Z_j^{(i)}$ in this case by

$$\mathbb{E}[Z_j^{(i)}] \leqslant \frac{k \cdot Z_{j-1}}{n - k + \ell} \leqslant \frac{Z_{j-1}}{2}.$$

Taking the expectation over the first $j - 1$ steps, the expected *difference* is then at most:

$$\mathbb{E}[R_j^{(i)}] \leqslant \frac{1}{m} \left(\frac{\mathbb{E}[Z_{j-1}]}{2} - \mathbb{E}[Z_{j-1}^{(i)}] \right)$$

and combining these bounds across coordinates gives:

$$\mathbb{E}[R_j] = \frac{1}{m} \sum_{i=1}^{m} \mathbb{E}[R_j^{(i)}]$$

$$\leqslant \frac{1}{m} \sum_{i=1}^{m} \frac{1}{m} \left(\frac{\mathbb{E}[Z_{j-1}]}{2} - \mathbb{E}[Z_{j-1}^{(i)}] \right)$$

$$= -\frac{1}{2m} \mathbb{E}[Z_{j-1}]$$

Combining this with our telescoping sum and observing that $Z_0 \leqslant 1$ by definition now gives the recurrence:

$$\mathbb{E}[Z_t] \leqslant 1 - \frac{1}{2m} \sum_{j=0}^{t-1} \mathbb{E}[Z_j] \implies \mathbb{E}[Z_t] \leqslant \left(1 - \frac{1}{2m} \right)^t,$$

and $t = \lceil 2m \log(ek) \rceil$ is therefore sufficient to bound $\mathbb{E}[Z_t] \leqslant \frac{1}{ek}$ as desired. $\qquad \square$

We remark that the above mixing bound is essentially optimal by classical coupon collector bounds when ℓ is at most polynomial in k. It is plausible, however, that the (modified) Log-Sobolev constant can be improved to $\Theta(\frac{\ell}{k})$, which would imply optimal concentration for lifted functions on partite HDX. While this cannot be shown by bounding the mixing time, appealing to the local-to-global entropy contraction framework of [CLV21; GM20] reduces the problem to proving optimal entropy contraction for the well-studied Kneser graphs.[29] We are not aware of such a bound and leave the problem as a question for future work.

C Outstanding Proofs on Samplers and Concentration

C.1 Sampling

Claim C.1 (Claim 4.8 Restated). Let $\beta, \delta > 0$, let $\delta' > \delta$ and $\alpha < \frac{\min\{\delta, 0.5\}}{1 + \delta}$. Then for every (α, β, δ)-sampler $G = (L, R, E)$, it holds that $G_{op} := (R, L, E)$ is a $\left(\frac{1 - \alpha(1+\delta)}{\alpha(\delta' - \delta)} \beta, 2\alpha, \delta' \right)$-sampler.

Proof of Claim 4.8. Let $A \subseteq L$ be such that $\mathbb{P}[A] \geqslant \frac{1 - \alpha(1+\varepsilon)}{\alpha(\varepsilon' - \varepsilon)} \beta$. Let

$$B_S = \left\{ v \in R \,\middle|\, \mathbb{P}_{u \in L}[u \in A \mid u \sim v] < (1 - \varepsilon') \mathbb{P}[A] \right\}$$

and let

$$B_B = \left\{ v \in R \,\middle|\, \mathbb{P}_{u \in L}[u \in A \mid u \sim v] > (1 + \varepsilon') \mathbb{P}[A] \right\}.$$

Showing that $\mathbb{P}[B_B], \mathbb{P}[B_S] \leqslant \alpha$ will prove the claim. Let us begin with B_S. Assume toward contradiction that $\mathbb{P}[B_S] > \alpha$. We denote by $\mathbf{1}_A : L \to \{0, 1\}$ and $\mathbf{1}_{B_S} : R \to \{0, 1\}$ the indicators of A, B_S respectively. On the one hand

$$\mathbb{E}_{uv \in E}[\mathbf{1}_A(u) \mathbf{1}_{B_S}(v)] < \mathbb{P}[B_S] (1 - \varepsilon') \mathbb{P}[A], \tag{C.1}$$

---

[29]In particular, it is enough to show for any distribution ϕ on the edges of the Kneser graph that $Ent(D\phi) \leqslant \left(\frac{1}{2} + O\left(\frac{\ell+1}{k+1} \right) \right) Ent(\phi)$ where D is the down operator, i.e. if entropy contraction roughly matches variance contraction.

since the product of indicators is one iff $v \in B_S$, in which case, there are at most a $(1 - \varepsilon') \, \mathbb{P}\,[A]$ fraction of $u \in A$ adjacent to v by the definition of B_S. On the other hand, by the sampling properties of (L, R, E) there is at most a β-fraction of $u \in L$ such that $\mathbb{P}_{v \in R}\,[v \in B_S \mid \; v \sim u] < (1 - \varepsilon) \, \mathbb{P}\,[B_S]$. Hence

$$\mathop{\mathbb{E}}_{uv \in E}\,[\mathbf{1}_A(u)\mathbf{1}_{B_S}(v)] \geqslant \mathbb{P}\,[B_S]\,(1 - \varepsilon)(\mathbb{P}\,[A] - \beta).^{30} \tag{C.2}$$

Combining (C.1) with (C.2) yields

$$\mathbb{P}\,[B_S]\,\mathbb{P}\,[A]\,(\varepsilon' - \varepsilon) < \beta \, \mathbb{P}\,[B_S]\,(1 - \varepsilon)$$

or

$$\mathbb{P}\,[A] < \frac{1 - \varepsilon}{\varepsilon' - \varepsilon}\beta \overset{(\alpha < \frac{\varepsilon}{1+\varepsilon})}{\leqslant} \frac{1 - \alpha(1 + \varepsilon)}{\alpha(\varepsilon' - \varepsilon)}\beta,$$

a contradiction to the lower bound on the size of A.

Let us bound B_B. Let $\mathbf{1}_{B_B} : R \to \{0, 1\}$ be the indicator of B_B. Similar to (C.1) we have that

$$\mathop{\mathbb{E}}_{uv \in E}\,[\mathbf{1}_A(u)\mathbf{1}_{B_S}(v)] > \mathbb{P}\,[B_B]\,(1 + \varepsilon')\,\mathbb{P}\,[A]. \tag{C.3}$$

By the sampling properties of (L, R, E) there is at most a β-fraction of $u \in L$ such that $\mathbb{P}_{v \in R}\,[v \in B_B \mid \; v \sim u] > (1 + \varepsilon) \, \mathbb{P}\,[B_B]$. So similar to (C.2) we have that

$$\mathop{\mathbb{E}}_{uv \in E}\,[\mathbf{1}_A(u)\mathbf{1}_{B_S}(v)] \leqslant \beta + \mathbb{P}\,[B_B]\,(1 + \varepsilon)(\mathbb{P}\,[A] - \beta). \tag{C.4}$$

Combining (C.3) with (C.4) yields

$$\mathbb{P}\,[B_B]\,\mathbb{P}\,[A]\,(\varepsilon - \varepsilon') < \beta\,(1 - \mathbb{P}\,[B_B]\,(1 + \varepsilon))$$

or

$$\mathbb{P}\,[A] < \frac{1 - \mathbb{P}\,[B_B]\,(1 + \varepsilon)}{\mathbb{P}\,[B_B]\,(\varepsilon' - \varepsilon)}\beta \leqslant \frac{1 - \alpha(1 + \varepsilon)}{\alpha(\varepsilon' - \varepsilon)}\beta,$$

where the last inequality is because the function $x \mapsto \frac{1 - (1+\varepsilon)x}{x(\varepsilon' - \varepsilon)}\beta$ is monotone increasing when $x \in [0, \frac{0.5}{1+\varepsilon}]$ and α is in this interval. This is a contradiction to the lower bound on the size of A. $\qquad\square$

Claim C.2 (Claim 4.9 Restated). Let $G = (L, R, E)$ be a bipartite graph.

1. If G is a (β, δ)-additive sampler then G is a $(C\delta, \beta, \frac{1}{C})$-multiplicative sampler for any $C > 1$.

2. If G is a (α, β, δ)-multiplicative sampler for $\alpha \leqslant \frac{1}{2}$. Then G is a (β, δ)-additive sampler, where $\delta = \max\{\delta, (1 + \delta)(\alpha + p)\}$ and $p = \max_{v \in R} \mathbb{P}\,[v]$.

Proof of Claim 4.9. The first item follows immediately from the definition of multiplicative samplers. If $\mathbb{P}\,[A] \geqslant C\varepsilon$ then $\frac{1}{C}\,\mathbb{P}\,[A] \geqslant \varepsilon$ so by the promise of additive sampling it holds that

$$\mathop{\mathbb{P}}_{v \in L}\left[\left|\mathop{\mathbb{P}}_{u \in R}\,[u \in A \mid \; u \sim v] - \mathbb{P}\,[A]\right| > \frac{1}{C}\,\mathbb{P}\,[A]\right] \leqslant \mathop{\mathbb{P}}_{v \in L}\left[\left|\mathop{\mathbb{P}}_{u \in R}\,[u \in A \mid \; u \sim v] - \mathbb{P}\,[A]\right| > \varepsilon\right] \leqslant \beta.$$

We turn to the second item. Let $A \subseteq R$. For every $u \in L$ it holds that

$$\left|\mathop{\mathbb{P}}_{u \in R}\,[u \in A \mid \; u \sim v] - \mathbb{P}\,[A]\right| = \left|1 - \mathop{\mathbb{P}}_{u \in R}\,[u \in A \mid \; u \sim v] - (1 - \mathbb{P}\,[A])\right| = \left|\mathop{\mathbb{P}}_{u \in R}\,[u \in R \setminus A \mid \; u \sim v] - \mathbb{P}\,[R \setminus A]\right|,$$

so we assume that without loss of generality $\mathbb{P}\,[A] \leqslant \frac{1}{2}$. We need to show that

$$\mathop{\mathbb{P}}_{v \in L}\left[\left|\mathop{\mathbb{P}}_{u \in R}\,[u \in A \mid \; u \sim v] - \mathbb{P}\,[A]\right| > \delta\right] \leqslant \beta. \tag{C.5}$$

---

[30] $\mathbb{P}\,[A] > \beta$ by assumption.

If $\mathbb{P}[A] \geqslant \alpha$ then this holds from the multiplicative sampling guarantee. Hence it suffices to show that this inequality *always holds* for sets of relative size smaller that δ.

First we note that if $u \in L$ is such that $\mathbb{P}_{u \in R}[u \in A \mid u \sim v] < \mathbb{P}[A]$ then $|\mathbb{P}_{u \in R}[u \in A \mid u \sim v] - \mathbb{P}[A]| \leqslant \mathbb{P}[A] \leqslant \delta$ so

$$\mathbb{P}_{v \in L}\left[\left|\mathbb{P}_{u \in R}[u \in A \mid u \sim v] - \mathbb{P}[A]\right| > \delta\right] = \mathbb{P}_{v \in L}\left[\mathbb{P}_{u \in R}[u \in A \mid u \sim v] > \delta + \mathbb{P}[A]\right].$$

In this case we find a subset $B \supseteq A$ such that $\mathbb{P}[B] \leqslant \alpha + p$ (we can do so by adding vertices to A one-by-one which is where p comes into play). Then

$$\mathbb{P}_{v \in L}\left[\mathbb{P}_{u \in R}[u \in A \mid u \sim v] > \delta + \mathbb{P}[A]\right] \leqslant \mathbb{P}_{v \in L}\left[\mathbb{P}_{u \in R}[u \in B \mid u \sim v] > \delta + \mathbb{P}[A]\right]$$

$$\leqslant \mathbb{P}_{v \in L}\left[\mathbb{P}_{u \in R}[u \in B \mid u \sim v] > (1 + \varepsilon)\mathbb{P}[B]\right]$$

$$\leqslant \beta.$$

The claim follows. $\qquad\square$

One can also remove the dependence on δ in the second item.

Claim C.3 ([Claim 4.10](#) Restated). Let $\beta, \alpha_0 > 0$. If for every $\alpha > \alpha_0$ it holds that G is an $(\alpha, \beta, \frac{\alpha_0}{\sqrt{\alpha}})$-multiplicative sampler, then G is a $(\beta, 2(\alpha + p))$-additive sampler where $p = \max_{v \in R} \mathbb{P}[v]$.

Proof of [Claim 4.10](#). The proof is the same as the proof of the second item in [Claim 4.9](#). the only difference is that for *large* sets, we still have that

$$\mathbb{P}_{v \in L}\left[\left|\mathbb{P}_{u \in R}[u \in A \mid u \sim v] - \mathbb{P}[A]\right| > 2\alpha_0 + p\right] \leqslant \beta \tag{C.6}$$

because if $\mathbb{P}[A] = \alpha$ then $\frac{\alpha_0}{\sqrt{\alpha}}\alpha \leqslant \alpha_0$. $\qquad\square$

Claim (Restatement of [Claim 4.11](#)). Let $G = (L, R, E)$ be an (ε, β)-additive sampler such that every $r \in R$ has degree at least k and every $v \in L$ has probability at most $\frac{1}{k}$. Assume that $\exp(-0.01\varepsilon^2 k)) < \frac{1}{4}$. Then G is also a $(4\beta, 2\varepsilon)$-function additive sampler.

Proof of [Claim 4.11](#). Fix $f : L \to [0, 1]$. We prove a $(2\beta, 2\varepsilon)$ upper tail bound. The lower bound is similar. Let $A \sim \mathcal{P}(L)$ be a random subset where every vertex is inserted into A independently with probability $p_v = f(v)$. Since G is a sampler, for every possible outcome A we have $\mathbb{P}_{r \in R}[\mathbb{P}_{v \sim r}[v \in A] - \mathbb{P}[A] > \varepsilon] < \beta$. Let $\mathbf{1}(A, r)$ be the indicator for this event. In particular, we have that $\mathbb{E}_{A, r \sim R}[\mathbf{1}(A, r)] \leqslant \beta$. By Markov, the fraction of $r_0 \in R$ such that $\mathbb{E}_A[\mathbf{1}(A, r_0)] > \frac{1}{2}$ is at most 2β. On the other hand, we will show that if

$$\mathbb{E}_{v \sim r_0}[f(v)] \geqslant \mathbb{E}_{v \in L}[f(v)] + 2\varepsilon$$

then $\mathbb{E}_A[\mathbf{1}(A, r_0)] > \frac{1}{2}$. Indeed, fix such an r_0. Observe that if $\mathbb{P}_{v \in L}[A] < \mathbb{E}_{v \in L}[f] + \frac{1}{2}\varepsilon$ and $\mathbb{P}_{v \sim r_0}[A] > \mathbb{E}_{v \sim r_0}[f(v)] - \frac{1}{2}\varepsilon$, then

$$\mathbb{P}_{v \sim r_0}[A] > \mathbb{E}_{v \sim r_0}[f(v)] - \frac{1}{2}\varepsilon \geqslant \mathbb{E}_{v \in L}[f(v)] + \frac{3}{2}\varepsilon \geqslant \mathbb{P}[A] + \varepsilon.$$

It is direct to check that $\mathbb{E}_A[\mathbb{P}[A]] = \mathbb{E}[f]$ and that $\mathbb{E}_A[\mathbb{P}_{v \sim r_0}[A]] = \mathbb{E}_{v \sim r_0}[f(v)]$. Hoeffding's inequality therefore bounds the probability that either $\mathbb{P}_{v \in L}[A] < \mathbb{E}_{v \in L}[f] + \frac{1}{2}\varepsilon$ or $\mathbb{P}_{v \sim r_0}[A] > \mathbb{E}_{v \sim r_0}[f(v)] - \frac{1}{2}\varepsilon$ by $2\exp(-0.01\varepsilon^2 k) < \frac{1}{2}$ and $\mathbb{E}_A[\mathbf{1}(A, r_0)] > \frac{1}{2}$ follows as desired. $\qquad\square$

Finally we prove the basic Chebyshev-type sampling bound from expansion.

Claim (Restatement of [Claim 7.3](#)). Let $G = (L, R, E)$ be a λ-bipartite expander and let A be its bipartite adjacency operator. Let $f : L \to [0,1]$ be a function with $\mathbb{E}[f] \geq \mu$ and let $0 < \varepsilon < \mu$. Then

$$\mathbb{P}_{u \in R}[Af(u) < \mu - \varepsilon] < \frac{\lambda^2 \mu}{\varepsilon^2}.$$

Proof of [Claim 7.3](#). Dinur and Kaufman observe the following bound

$$\mathbb{P}_{u \sim R}[|Af(u) - \mu| > \varepsilon] < \frac{\lambda^2 Var(f)}{\varepsilon^2}.$$

Let $f^\perp = f - \mu$. Then $Af(v) - \mu = Af^\perp$. By Chebyshev's inequality, $\mathbb{P}[T] \leq \frac{\mathbb{E}[(Af^\perp)^2]}{\varepsilon^2}$ and by λ-expansion $\mathbb{E}[(Af^\perp)^2] \leq \lambda^2 \mathbb{E}[(f^\perp)^2] = \lambda^2 Var(f)$. The claim follows.

Now by the above we can write

$$\mathbb{P}[Af(u) < \mu - \varepsilon] \leq \mathbb{P}[|Af(u) - \mathbb{E}[f]| > \mathbb{E}[f] - \mu + \varepsilon] \leq \frac{\lambda^2 Var(f)}{(\varepsilon + (\mathbb{E}[f] - \mu))^2}.$$

Noting that $Var(f) \leq \mathbb{E}[f] = \mu + (\mathbb{E}[f] - \mu)$ and denoting $x = \mathbb{E}[f] - \mu$, this is at most $\frac{\lambda^2(\mu+x)}{(\varepsilon+x)^2} \leq \frac{\lambda^2\mu}{\varepsilon^2}$. $\square$

C.2 Concentration

Lemma C.4 ([Lemma 4.15](#) Restated)**.** *Let X be a d-uniform simplicial complex, $k \leq d$, and $f : X(k) \to [0,1]$ a function satisfying*

1. **Upper Tail**: $\displaystyle \mathbb{P}_{X(k)}[f - \mathbb{E}[f] > t] \leq f_{up}(t)$

2. **Lower Tail**: $\displaystyle \mathbb{P}_{X(k)}[f - \mathbb{E}[f] < -t] \leq f_{low}(t).$

for some functions $f_{up}, f_{low} : \mathbb{R}_+ \to [0,1]$. Then the d-lift $U_{k,d}f : X(d) \to \mathbb{R}$ satisfies:

1. **Upper Tail**: $\displaystyle \mathbb{P}_{X(d)}[U_{k,d}f - \mathbb{E}[f] > t] \leq f_{up}(\tfrac{t}{2})(1 - \pi^{d,k,f}_{low}(\tfrac{t}{2}))^{-1}$

2. **Lower Tail**: $\displaystyle \mathbb{P}_{X(d)}[U_{k,d}f - \mathbb{E}[f] < -t] \leq f_{low}(\tfrac{t}{2})(1 - \pi^{d,k,f}_{up}(\tfrac{t}{2}))^{-1}.$

Proof. We give the argument for the upper tail. The argument for the lower tail is analogous. Assume for the sake of contradiction that

$$\mathbb{P}_{X(d)}\left[U_{k,d}f - \mu > 2t\right] > f_{up}(t)\left(1 - \pi^{f,k,d}_{low}(t)\right)^{-1}.$$

Since a k-face can be drawn by first drawing a d-face, then a uniformly random k-subface, we can use concentration of the complete complex to derive a contradiction. In particular, let E_1 denote the event that $U_{k,d}f - \mu > 2t$. Then:

$$\mathbb{P}_{\{v_0,\dots,v_k\} \in X(k)}\left[U_{k,d}f - \mu > t\right] \geq \mathbb{P}[E_1]\,\mathbb{P}_{r \subset s}[f(r) - \mu > t \mid s \in E_1]$$

$$\geq \mathbb{P}[E_1](1 - \mathbb{P}_{r \subset s}[f(r) - \mathbb{E}[f] < t \mid s \in E_1])$$

$$\geq \mathbb{P}[E_1](1 - \mathbb{P}_{r \subset s}[f(r) - U_{k,d} < -t \mid s \in E_1])$$

$$> f_{up}(t)$$

$\square$

Lemma C.5 ([Lemma 4.16](#) Restated)**.** *Let X be a d-uniform simplicial complex and $k \leq d$. Assume there exist functions $f_{up}(t,\nu)$ and $f_{low}(t,\nu)$ such that any ν-Lipschitz $f : X(d) \to \mathbb{R}$ satisfies:*

1. **Upper Tail**: $\mathbb{P}[f - \mathbb{E}[f] > t] \leqslant f_{up}(t, \nu)$

2. **Lower Tail**: $\mathbb{P}[f - \mathbb{E}[f] < -t] \leqslant f_{low}(t, \nu)$.

Then any function $f' : X(k) \to \mathbb{R}$ *with* ν-*bounded difference satisfies:*

1. **Upper Tail**: $\mathbb{P}[f' - \mathbb{E}[f'] > t] \leqslant f_{up}(\frac{t}{2}, \frac{k}{d}\nu) + e^{-\frac{t^2}{4\nu}}$

2. **Lower Tail**: $\mathbb{P}[f' - \mathbb{E}[f'] < -t] \leqslant f_{low}(\frac{t}{2}, \frac{k}{d}\nu) + e^{-\frac{t^2}{4\nu}}$.

Proof. Let $f : X(k) \to \mathbb{R}$ be a function with ν-bounded difference and g its d-lift $g = U_{k,d}f$. Since a face $r \in X(k)$ can equivalently be sampled by first drawing a face $s \in X(d)$, then subsampling r uniformly from s, we can bound the concentration of f as:

$$\mathbb{P}[f(r) - \mathbb{E}[f] \geqslant t] \leqslant \underset{r \subset s}{\mathbb{P}}[g(s) - \mathbb{E}[f] \geqslant t/2 \vee f(r) - g(s) \geqslant t/2]$$

$$\leqslant \underset{r \subset s}{\mathbb{P}}[g(s) - \mathbb{E}[g] \geqslant t/2] + \underset{r \subset s}{\mathbb{P}}\left[f(r) - \underset{r \subset s}{\mathbb{E}}[f(r)] \geqslant t/2\right]$$

where we have used the fact that $\mathbb{E}[g] = \mathbb{E}[f]$. Informally, the idea is then to argue the first term is small due to exponential concentration of $X(d)$ and the fact that the lift g is itself more concentrated than f, and the second term is small by subgaussian concentration of the complete complex.

We first argue that if f has ν-bounded difference, g is $\frac{k}{d}\nu$-Lipschitz. Given $s \in X(d)$ and $i \in [d]$, let $s^{(i)}$ denote a copy of s with the ith coordinate re-sampled. Expanding ν_g we have:

$$\nu_g \leqslant \sum_{i=1}^{d} \underset{s, s^{(i)}}{\mathbb{E}}\left[\left(g(s) - g(s^{(i)})\right)_+^2\right]$$

$$= \sum_{i=1}^{d} \underset{s, s^{(i)}}{\mathbb{E}}\left[\left(\underset{t \subset s}{\mathbb{E}}[f] - \underset{t \subset s^{(i)}}{\mathbb{E}}[f]\right)_+^2\right]$$

$$= \frac{1}{\binom{d}{k}^2} \sum_{i=1}^{d} \underset{s, s^{(i)}}{\mathbb{E}}\left[\left(\sum_{T \subseteq [d]} f(s_T) - f(s_T^{(i)})\right)_+^2\right]$$

where we recall s_T is the k-face consisting of the elements of s indexed by T. Observe that the inner term is non-zero only when $i \in T$, since otherwise $s_T = s_T^{(i)}$. Thus by Cauchy-Schwarz, we have:

$$\nu_f \leqslant \frac{\binom{d-1}{k-1}}{\binom{d}{k}^2} \sum_{i=1}^{d} \underset{s, s^{(i)}}{\mathbb{E}}\left[\sum_{T \subseteq [d]}\left(f(s_T) - f(s_T^{(i)})\right)^2\right]$$

$$\leqslant \frac{\binom{d-1}{k-1}}{\binom{d}{k}^2} \sum_{i=1}^{d} \underset{s, s^{(i)}}{\mathbb{E}}\left[\sum_{i \in T \subseteq [d]} \frac{\nu}{k}\right]$$

$$= \frac{k}{d}\nu$$

where we have used the fact that $s(T)$ and $s'_{(i)}(T)$ are neighboring k-faces in the down-up walk on level k. We can now bound the first term of our upper tail by assumption as

$$\underset{r \subset s}{\mathbb{P}}[g(s) - \mathbb{E}[g] \geqslant t/2] \leqslant f_{up}\left(\frac{t}{2}, \frac{k}{d}\nu\right).$$

For the second term, observe that since f has ν-bounded difference as a function of $X(k)$, the restriction of f to $\binom{s}{k} \subset X(k)$ also has ν-bounded difference (since any edge (r, r') in the down-up walk restricted to s

111

is also an edge in the walk on $X(k)$). Thus using subgaussian concentration for Lipschitz functions on the complete complex (see e.g. [CGM19, Corollary 2]) we also have

$$\Pr_{r \subset s}\left[f(r) - \mathbb{E}_{r \subset s}[f(r)] \geqslant t/2\right] \leqslant e^{-\frac{t^2}{4\nu}}.$$

Combining the two gives the result. $\qquad\square$

D High Dimensional Expander-Mixing Lemma

In this section, we give the proof of our variant of the high-dimensional expander mixing lemma and its various corollaries.

Theorem D.1 (high dimensional expander Mixing Lemma (Theorem 10.14 Restated))**.** *Let (X, T, ρ) be a depth D λ-tuple splitting tree with k leaves. Denote by d_i the depth of the leaf labeled i. Then for any family of functions $\{f_i : X[i] \to \mathbb{R}\}_{i \in [k]}$:*

$$\left|\mathbb{E}_{a \in X(k-1)}\left[\prod_{i=1}^{k} f_i(a_i)\right] - \prod_{i=1}^{k}\mathbb{E}_{a_i \in X[i]}[f_i(a_i)]\right| \leqslant 3^D \lambda \prod_{i=1}^{k}\|f_i\|_{2^{d_i}}.$$

If (X, T, ρ) is instead a depth D standard λ-splitting tree, we take $\{f_i : X(1) \to \mathbb{R}\}_{i \in [k]}$ and have:

$$\left|\mathbb{E}_{a \in X(k-1), \pi \in S_k}\left[\prod_{i=1}^{k} f_i(a_{\pi(i)})\right] - \prod_{i=1}^{k}\mathbb{E}_{a_i \in X(1)}[f_i(a_i)]\right| \leqslant 3^D \lambda \prod_{i=1}^{k}\|f_i\|_{2^{d_i}}.$$

Proof. We prove the partite case directly, then argue (10.3) follows from partitification. It is convenient to introduce the following notation. Let $\bar{f} = (f_1, f_2, \ldots, f_k)$ such that $f_i : X[i] \to \mathbb{R}$. Let $S \subseteq [k]$. Let (T, ρ) be an ordered tree. Without loss of generality we identify every leaf $u \in \mathcal{L}$ with its label $\rho(u)$.

– Denote by $\bar{f}_S$ the sub-tuple containing only functions whose index is in S.

– Let $\pi_{\bar{f}_S} : X[S] \to \mathbb{R}$ be the product of functions in $\bar{f}_S$, that is $\pi_{\bar{f}_S}(s) = \prod_{i \in S} f_i(a_i)$ (where $s = \{a_i\}_{i \in S}$). We note that when S_1, S_2 are disjoint then $\pi_{\bar{f}_{S_1}} \cdot \pi_{\bar{f}_{S_2}} = \pi_{\bar{f}_{S_1 \cup S_2}}$.

– Let $E_{\bar{f}_S} = \prod_{i \in S} \mathbb{E}[f_i]$.

– Let T_L and T_R be the sub-trees rooted by the left and right children of the root of T, respectively. Let L, R be the leaves of T_L, T_R respectively.

– For T, T_L, T_R and tuple of functions $\bar{f}$ we denote by

$$\mathcal{E}(\bar{f}_{[k]}, T) = \prod_{i=1}^{k}\|f_i\|_{2^{d_i}}, \ \mathcal{E}(\bar{f}_L, T_L) = \prod_{i \in L}\|f_i\|_{2^{d_i-1}}, \ \mathcal{E}(\bar{f}_R, T_R) = \prod_{i \in R}\|f_i\|_{2^{d_i-1}}.$$

In this notation we need to show:

$$\left|\mathbb{E}[\pi_{\bar{f}_{[k]}}] - E_{\bar{f}_{[k]}}\right| \leqslant 3^D \lambda \mathcal{E}(\bar{f}_{[k]}, T).$$

The proof is by induction on the depth D. The base case (depth 1) is trivial. Assume that the theorem holds for any set of functions on any tree of depth $\leqslant D - 1$. By adding and subtracting $\mathbb{E}[\pi_{\bar{f}_L}] \cdot \mathbb{E}[\pi_{\bar{f}_R}]$, observe that

$$\left|\mathbb{E}[\pi_{\bar{f}_{[k]}}] - E_{\bar{f}_{[k]}}\right| \leqslant \underbrace{\left|\mathbb{E}[\pi_{\bar{f}_{[k]}}] - \mathbb{E}[\pi_{\bar{f}_L}] \cdot \mathbb{E}[\pi_{\bar{f}_R}]\right|}_{\text{I}} + \underbrace{\left|\mathbb{E}[\pi_{\bar{f}_L}] \cdot \mathbb{E}[\pi_{\bar{f}_R}] - E_{\bar{f}_{[k]}}\right|}_{\text{II}} \tag{D.1}$$

by the triangle inequality. We bound I and II separately, starting with I. Since the swap walk $S_{L,R}$ is a λ-bipartite expander, applying (4.2) to $\pi_{\bar{f}_L}, \pi_{\bar{f}_R}$ we get that the first term in the right-hand side of (D.1) is

$$\left| \mathbb{E}[\pi_{\bar{f}_{[k]}}] - \mathbb{E}[\pi_{\bar{f}_L}] \mathbb{E}[\pi_{\bar{f}_R}] \right| = \left| \mathbb{E}[\pi_{\bar{f}_L} \cdot \pi_{\bar{f}_R}] - \mathbb{E}[\pi_{\bar{f}_L}] \mathbb{E}[\pi_{\bar{f}_R}] \right| \leqslant \lambda \|\pi_{\bar{f}_L}\|_2 \|\pi_{\bar{f}_R}\|_2 \leqslant \lambda \prod_i \|f_i\|_{2^{d_i}}$$

where we have used the fact that

$$\|\pi_{\bar{f}_L}\|_2 \leqslant \prod_{i \in L} \|f_i\|_{2^{d_i}}, \quad \text{and} \quad \|\pi_{\bar{f}_R}\|_2 \leqslant \prod_{i \in R} \|f_i\|_{2^{d_i}} \tag{D.2}$$

by repeatedly applying Cauchy-Schwarz along the splitting trees of L and R respectively.

It remains to bound II. By the triangle inequality, we can factor this term into components dependent on the left and right subtrees:

$$\left| \mathbb{E}[\pi_{\bar{f}_R}] \cdot \mathbb{E}[\pi_{\bar{f}_L}] - E_{\bar{f}_{[k]}} \right| \leqslant \left| \mathbb{E}[\pi_{\bar{f}_L}] \cdot \mathbb{E}[\pi_{\bar{f}_R}] - \mathbb{E}[\pi_{\bar{f}_L}] \cdot E_{\bar{f}_R} \right| + \left| \mathbb{E}[\pi_{\bar{f}_L}] \cdot E_{\bar{f}_R} - E_{\bar{f}_{[k]}} \right| \tag{D.3}$$
$$= \left| \mathbb{E}[\pi_{\bar{f}_L}] \right| \left| \mathbb{E}[\pi_{\bar{f}_R}] - E_{\bar{f}_R} \right| + \left| E_{\bar{f}_R} \right| \left| \mathbb{E}[\pi_{\bar{f}_L}] - E_{\bar{f}_L} \right|.$$

Since $(X^L, T_L, \rho|_{T_L}), (X^R, T_R, \rho|_{T_R})$ are both λ-tuple splitting trees of depth $\leqslant D - 1$,[31] the inductive hypothesis implies

$$\left| \mathbb{E}[\pi_{\bar{f}_R}] - E_{\bar{f}_R} \right| \leqslant 3^{D-1} \lambda \mathcal{E}(\bar{f}_R, T_R), \quad \text{and} \quad \left| \mathbb{E}[\pi_{\bar{f}_L}] - E_{\bar{f}_L} \right| \leqslant 3^{D-1} \lambda \mathcal{E}(\bar{f}_L, T_L)$$

so we can upper bound (D.3) by

$$3^{D-1} \lambda \left(\left| \mathbb{E}[\pi_{\bar{f}_L}] \right| \mathcal{E}(\bar{f}_R, T_R) + \left| E_{\bar{f}_R} \right| \mathcal{E}(\bar{f}_L, T_L) \right).$$

Finally, observe that

$$\left| E_{\bar{f}_R} \right| = \prod_{i \in R} |\mathbb{E}[f_i]| \leqslant \prod_{i \in R} \|f_i\|_1 \leqslant \prod_{i \in R} \|f_i\|_{2^{d_i}}$$

by monotonicity of expectation norms and similarly

$$\left| \mathbb{E}[\pi_{\bar{f}_L}] \right| \leqslant \|\pi_{\bar{f}_L}\|_2 \leqslant \prod_{i \in L} \|f_i\|_{2^{d_i}}$$

by (D.2). Altogether, this gives an upper bound of

$$\text{I} + \text{II} \leqslant \lambda \mathcal{E}(\bar{f}_{[k]}, T) + 2 \cdot 3^{D-1} \lambda \mathcal{E}(\bar{f}_{[k]}, T) \leqslant \lambda 3^D \mathcal{E}(\bar{f}_{[k]}, T)$$

as desired.

For the unordered case, take an arbitrary partitification (X', T, ρ') as in Definition 10.8. (X', T, ρ') is a λ-tuple splitting tree by Claim 10.9. Define $\tilde{f}_i : X[\rho'(u_i)] \to \mathbb{R}$ by $\tilde{f}_i((v, j)) = f_i(v)$. Viewed as random variables, $\tilde{f}_i$ and f_i are equidistributed, so $\mathbb{E}[f_i] = \mathbb{E}[\tilde{f}_i]$ and $\|f_i\|_{2^{d_i}} = \|\tilde{f}_i\|_{2^{d_i}}$. Moreover, by construction

$$\mathop{\mathbb{E}}_{a \in X(k-1), \pi \in S_k} \left[\prod_{i=1}^k f_i(a_{\pi(i)}) \right] = \mathop{\mathbb{E}}_{(a,\pi) \in X'(k-1)} \left[\prod_{i=1}^k \tilde{f}_i(a_{\pi(i)}, \pi(i)) \right],$$

so can write

$$\left| \mathop{\mathbb{E}}_{a \in X(k), \pi \in S_k} \left[\prod_{i=1}^k f_i(a_{\pi(i)}) \right] - \prod_{i=1}^k \mathop{\mathbb{E}}_{a_i \in X(1)} [f_i(a_i)] \right| = \left| \mathop{\mathbb{E}}_{(a,\pi) \in X'(k)} \left[\prod_{i=1}^k \tilde{f}_i((a, \pi)_i) \right] - \prod_{i=1}^k \mathop{\mathbb{E}}_{(a,\pi)_i \in X'(i)} [\tilde{f}_i(a_i, i)] \right|$$

---

[31] Recall that X^L, X^R are the simplicial sub-complexes induced by the faces in $X[L]$ and $X[R]$ respectively.

$$\leqslant 3^D \lambda \prod_{i=1}^{k} \|\tilde{f}_i\|_{2^{d_i}}$$

$$= 3^D \lambda \prod_{i=1}^{k} \|f_i\|_{2^{d_i}}$$

as desired. □

We now record a few useful corollaries from which hitting set and bias amplification are essentially immediate.

Corollary D.2. *Let X be as in [Theorem 10.14](#) and $A_i \subseteq X[i]$. Then*

$$\left| \underset{a \in X(k-1)}{\mathbb{P}} \left[\bigwedge_{i=1}^{k} a_i \in A_i \right] - \prod_{i=1}^{k} \mathbb{P}\left[A_i\right] \right| \leqslant 3^D \lambda \prod_{i=1}^{k} \mathbb{P}\left[A_i\right]^{2^{-d_i}} \leqslant 3^D \lambda. \tag{D.4}$$

Corollary D.3. *Let (X, T, ρ) be a λ-tuple splitting tree. Let $u_1, u_2, \ldots, u_\ell \in T$ be nodes such that for every $i \neq j$, u_i is not an ancestor of u_j, and let d_i be the depth of u_i in T. Let $f_1, f_2, \ldots, f_\ell$ be functions such that $f_i : X(\rho(u_i)) \to \mathbb{R}$. Then*

$$\left| \underset{a \in X(k-1)}{\mathbb{E}} \left[\prod_{i=1}^{\ell} f_i(a_{\rho(u_i)}) \right] - \prod_{i=1}^{\ell} \mathbb{E}[f_i(a_{\rho(u_i)})] \right| \leqslant 3^D \lambda \prod_{i=1}^{\ell} \|f_i\|_{2^{d_i}}, \tag{D.5}$$

where $D = \max\{d_i\}$.

Proof. We sequentially prune the tuple-splitting tree using nodes $u_1, u_2, \ldots, u_\ell$ as in [Definition 10.10](#). If there are nodes of depth greater than D we also prune their depth D ancestor. By [Claim 10.11](#), this results in a pruned λ-tuple splitting tree (X', T', ρ') where by construction $u_1, u_2, \ldots, u_\ell$ are leaves in T' corresponding to $X'[i] = X[\rho(u_i)]$ (we note X' may have other parts not associated with these leaves as well). Setting $f_j = 1$ for the latter, applying [Theorem 10.14](#) to (X', T', ρ') gives the desired bound. □

Proposition D.4 (Hitting Set ([Proposition 10.12](#) Restated))**.** *Any depth D λ-(tuple) splitting tree with k leaves is $(3^D \lambda, k)$-hitting.*

Proof. Any subset $A \subset X(1)$ can be divided into partite components $A = A_1 \amalg \ldots \amalg A_k$ where $A_i \subset X[i]$. [Corollary D.2](#) then implies

$$\underset{\sigma \in X(k)}{\mathbb{P}}[\sigma \subset A] \leqslant \prod_{i=1}^{k} \mu(A_i) + 3^D \lambda \leqslant \mu(A)^k + 3^D \lambda$$

where the righthand inequality follows from the fact that $\frac{1}{k} \sum_{i=1}^{k} \mu(A_i) = \mu(A)$. □

Proposition D.5 (Bias-Amplification ([Proposition 10.13](#) Restated))**.** *Let (X, T, ρ) be a λ-tuple splitting tree for $\lambda < \frac{1}{16}$. For any $0 < \varepsilon < \frac{1}{4}$ and family of mean ε functions $\{f_i : X[i] \to \{\pm 1\}\}_{i \in [k]}$, $\prod f_i$ is a $\{\pm 1\}$-valued function with bias at most:*

$$\left| \underset{a \in X(k-1)}{\mathbb{E}} \left[\prod_{i=1}^{k} f_i(a_i) \right] \right| \leqslant \varepsilon^k + 2\lambda \tag{D.6}$$

Proof. The proof is similar to [Theorem 10.14](#), and we adopt the same notational conventions. The base case $k = 1$ is trivial. For the inductive step, we have

$$|\mathbb{E}[\pi_{\tilde{f}_{[k]}}]| \leqslant \left| \mathbb{E}[\pi_{\tilde{f}_{[k]}}] - \mathbb{E}[\pi_{\tilde{f}_L}] \cdot \mathbb{E}[\pi_{\tilde{f}_R}] \right| + \mathbb{E}[\pi_{\tilde{f}_L}] \cdot \mathbb{E}[\pi_{\tilde{f}_R}]$$

$$\leqslant \lambda + (\varepsilon^{|L|} + 2\lambda)(\varepsilon^{|R|} + 2\lambda)$$

$$\leqslant \varepsilon^k + 2\lambda$$

by our assumptions on λ and ε. □

Finally we list for completeness the unordered analogs of these results, which follow immediately from partitification as in Theorem 10.14.

Corollary D.6 (Expander-Mixing (Corollary D.2)). *Let X be a depth D λ-splitting tree. Then for any $A_1, \ldots, A_k \subset X(1)$:*

$$\left| \mathop{\mathbb{P}}_{a \in X(k), \pi \in S_k} \left[\bigwedge_{i=1}^{k} a_i \in A_i \right] - \prod_{i=1}^{k} \mathbb{P}\left[A_i\right] \right| \leqslant 3^D \lambda \prod_{i=1}^{k} \mathbb{P}\left[A_i\right]^{2^{-d_i}} \leqslant 3^D \lambda. \quad \square \tag{D.7}$$

Corollary D.7 (Extended HD-EML (Corollary D.3)). *Let (X, T, ρ) be a λ-splitting tree, $u_1, u_2, \ldots, u_\ell \in T$ nodes such that for every $i \neq j$, u_i is not an ancestor of u_j, and let d_i be the depth of u_i in T. Let $f_1, f_2, \ldots, f_\ell$ be functions such that $f_i : X(\rho(u_i)) \to \mathbb{R}$. Then*

$$\left| \mathop{\mathbb{E}}_{a \in X(k-1), \pi \in S_k} \left[\prod_{i=1}^{\ell} f_i(a_{\rho(u_i)}) \right] - \prod_{i=1}^{\ell} \mathbb{E}[f_i(a_{\rho(u_i)})] \right| \leqslant 3^D \lambda \prod_{i=1}^{\ell} \|f_i\|_{2^{d_i}}, \tag{D.8}$$

where $D = \max\{d_i\}$.

Corollary D.8 (Hitting Set (Proposition 10.12)). *Any depth D λ splitting tree with k leaves is $(3^D \lambda, k)$-hitting.*

Corollary D.9 (Bias-Amplification (Proposition 10.13)). *Let (X, T, ρ) be a λ splitting tree for $\lambda < \frac{1}{16}$. For any $0 < \varepsilon < \frac{1}{4}$ and family of mean ε functions $\{f_i : X(1) \to \{\pm 1\}\}_{i \in [k]}$, $\prod f_i$ is a $\{\pm 1\}$-valued function with bias at most:*

$$\left| \mathop{\mathbb{E}}_{a \in X(k-1), \pi \in S_k} \left[\prod_{i=1}^{k} f_i(a_{\pi(i)}) \right] \right| \leqslant \varepsilon^k + 2\lambda \tag{D.9}$$

E Reverse Hypercontractivity: From Boolean to All Functions

We now give a generic reduction from reverse hypercontractivity for general functions to the Boolean case.

Theorem (Restatement of Theorem 6.12). *For every $\ell \geqslant 1$ and $\varepsilon > 0$ there exists $\kappa' \geqslant \left(\frac{1}{5}\right)^{2\ell(1+\varepsilon)} \left(\frac{1}{18 + \frac{12}{\varepsilon}}\right)^{2(\ell-1)}$ such that the following holds. Let V be a finite probability space and D a monotone linear operator such that for every $A, B \subseteq V$,*

$$\langle 1_A, D1_B \rangle \geqslant \kappa \, \mathbb{P}\left[A\right]^\ell \mathbb{P}\left[B\right]^\ell. \tag{E.1}$$

Then for every two arbitrary functions $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ it holds that

$$\langle f_1, Df_2 \rangle \geqslant \kappa \kappa' \|f_1\|_{\frac{1}{\ell(1+\varepsilon)}} \|f_2\|_{\frac{1}{\ell(1+\varepsilon)}}. \tag{E.2}$$

That is, D is $\left(\frac{1}{\ell(1+\varepsilon)}, \frac{1}{1-\ell(1+\varepsilon)}, \kappa\kappa'\right)$-reverse hypercontractive.

Before moving directly to the proof, we give a brief overview of the main idea. Starting with arbitrary f_1, f_2, our goal is to build discrete approximations of the f_i that remain close to the original functions in expectation and don't take too many unique values. Once this is the case, we can write the functions as a weighted sum over indicators for each value, and apply the boolean result without too much loss.

The key to the proof really lies in these approximations, which we build by discretizing the functions to powers of 2 and applying a careful iterative zero-ing procedure to components that fall too far from the mean. Naively this approach seems problematic, as f_i is arbitrary and may not have a well-behaved tail to truncate. We handle this by exploiting the fact that no function can have more than a small constant fraction of its mass beyond its expectation times (an appropriate power of) its $(1 \to (1 + \varepsilon))$-norm.

More formally, the proof of Theorem 6.12 is split into two intermediate reductions, corresponding to the 'approximation step' and the 'indicator step' above. The first of these is a reduction from general reverse hypercontractivity to a slightly richer set of functions that are *balanced* and *discrete*.

Definition E.1 (Nice Functions). For any $\alpha \in [0,1]$, and $\beta \in [1, \infty)$, we say the function $f : V \to \mathbb{R}_{\geqslant 0}$ is (α, β)-nice if it is:

1. **Balanced:** all non-zero outputs of f satisfy

$$\alpha \mathbb{E}[f] \leqslant f(x) \leqslant \beta \mathbb{E}[f].$$

2. **Discrete:** the range of f is entirely contained in $\{0\} \cup \{2^j : j \in \mathbb{Z}\}$.

To state the reduction, we require a setting of parameters α and β based on the $(1 \to (1+\varepsilon))$-norm of the functions in question due to the tail truncation strategy described above. In particular, for a fixed non-negative function f_i, define

$$\eta_i = \frac{\mathbb{E}[f_i^{1+\varepsilon}]}{\mathbb{E}[f_i]^{1+\varepsilon}} = \|f_i\|_{1 \to (1+\varepsilon)}^{1+\varepsilon}.$$

With this in mind, we state our two intermediate reductions and prove Theorem 6.12.

Proposition E.2 (From General to Nice). *Let $\ell, \varepsilon, \kappa, V$, and D be as in Theorem 6.12. Suppose that for any two functions $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ that are $\left(\frac{1}{16\eta_i}, (25\eta_i)^{\frac{2(1+\varepsilon)}{\varepsilon}}\right)$-nice*

$$\langle f_1^{\ell(1+\varepsilon)}, D f_2^{\ell(1+\varepsilon)} \rangle \geqslant \kappa \, \mathbb{E}[f_1]^{\ell(1+\varepsilon)} \, \mathbb{E}[f_2]^{\ell(1+\varepsilon)}. \tag{E.3}$$

Then for every two functions $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ it holds that

$$\langle f_1^{\ell(1+\varepsilon)}, D f_2^{\ell(1+\varepsilon)} \rangle \geqslant \left(\frac{1}{5}\right)^{2\ell(1+\varepsilon)} \kappa \, \mathbb{E}[f_1]^{\ell(1+\varepsilon)} \, \mathbb{E}[f_2]^{\ell(1+\varepsilon)}. \tag{E.4}$$

The 'indicator step' then reduces reverse hypercontractivity for nice functions to the Boolean case.

Proposition E.3 (From Nice to Boolean). *Let ℓ, ε, V, and D be as in Theorem 6.12. If for every $A, B \subseteq V$,*

$$\langle 1_A, D 1_B \rangle \geqslant \kappa \, \mathbb{P}[A]^\ell \, \mathbb{P}[B]^\ell, \tag{E.5}$$

then every two $\left(\frac{1}{16\eta_i}, (25\eta_i)^{\frac{2(1+\varepsilon)}{\varepsilon}}\right)$-nice functions $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ satisfy

$$\langle f_1^{\ell(1+\varepsilon)}, D f_2^{\ell(1+\varepsilon)} \rangle \geqslant \left(\frac{1}{18 + \frac{12}{\varepsilon}}\right)^{2(\ell-1)} \kappa \, \mathbb{E}[f_1]^{\ell(1+\varepsilon)} \, \mathbb{E}[f_2]^{\ell(1+\varepsilon)}. \tag{E.6}$$

The proof of Theorem 6.12 is now essentially immediate.

Proof of Theorem 6.12. We have by assumption that all $A, B \subseteq V$ satisfy $\langle 1_A, D 1_B \rangle \geqslant \kappa \, \mathbb{P}[A]^\ell \, \mathbb{P}[B]^\ell$, so we may apply Proposition E.3 to get that every two $\left(\frac{1}{16\eta_i}, (25\eta_i)^{\frac{2(1+\varepsilon)}{\varepsilon}}\right)$-nice functions $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ satisfy

$$\langle f_1^{\ell(1+\varepsilon)}, D f_2^{\ell(1+\varepsilon)} \rangle \geqslant \left(\frac{1}{18 + \frac{12}{\varepsilon}}\right)^{2(\ell-1)} \kappa \, \mathbb{E}[f_1]^{\ell(1+\varepsilon)} \, \mathbb{E}[f_2]^{\ell(1+\varepsilon)}.$$

Applying Proposition E.2 then implies *arbitrary* non-negative g_1, g_2 satisfy

$$\langle g_1^{\ell(1+\varepsilon)}, D g_2^{\ell(1+\varepsilon)} \rangle \geqslant \left(\frac{1}{5}\right)^{2\ell(1+\varepsilon)} \left(\frac{1}{18 + \frac{12}{\varepsilon}}\right)^{2(\ell-1)} \kappa \, \mathbb{E}[g_1]^{\ell(1+\varepsilon)} \, \mathbb{E}[g_2]^{\ell(1+\varepsilon)}.$$

Finally, for f_1, f_2 arbitrary non-negative functions applying the above to $g_i = f_i^{\frac{1}{\ell(1+\varepsilon)}}$ gives the result. $\qquad \square$

We now move to the formal proofs.

116

Proof of Proposition E.2. Given arbitrary $f_1, f_2 : V \to \mathbb{R}_{\geqslant 0}$ we will construct corresponding f'_1, f'_2 satisfying

1. $0 \leqslant f'_i \leqslant f_i$

2. $\mathbb{E}[f'_i] \geqslant \frac{1}{5}\mathbb{E}[f]$

3. f'_i is $\left(\frac{1}{16\eta_i}, (25\eta_i)^{\frac{2(1+\varepsilon)}{\varepsilon}}\right)$-nice

Combining the above with monotonicity of D and reverse hypercontractive inequality for nice functions gives

$$\langle f_1, Df_2 \rangle \geqslant \langle f'_1, Df'_2 \rangle \qquad \text{(Item (1) and Monotonicity)}$$
$$\geqslant \kappa \mathbb{E}[f'_1]^{\ell(1+\varepsilon)} \mathbb{E}[f'_2]^{\ell(1+\varepsilon)} \qquad \text{(Item (3) and Equation (E.3))}$$
$$\geqslant \left(\frac{1}{5}\right)^{2\ell(1+\varepsilon)} \kappa \mathbb{E}[f'_1]^{\ell(1+\varepsilon)} \mathbb{E}[f'_2]^{\ell(1+\varepsilon)} \qquad \text{(Item (2))}$$

as desired.

We construct f'_i in two separate steps: a simple discretization procedure, and a more involved balancing procedure that iteratively zeroes out unbalanced parts of the function. Starting with the former, given a non-negative f define its *rounding down* $\tilde{f}$ by setting for $j \in \mathbb{Z}$:

$$\tilde{f}(x) = \begin{cases} 2^j & \text{for } f(x) \in [2^j, 2^{j+1}) \\ 0 & \text{if } f(x) = 0. \end{cases}$$

Note that, by construction, $\frac{f_i}{2} \leqslant \tilde{f}_i \leqslant f_i$.

We now define an iterative balancing procedure outputting a sequence of functions $\tilde{f}_i = f_i^{(0)} \geqslant f_i^{(1)} \geqslant \dots \geqslant f_i^{(k_i)} = f'_i$ for $k_i \in \mathbb{N}$ some finite stopping time. For these (soon-to-be-defined) functions, define $\eta_i^{(j)} = \frac{\mathbb{E}[(f_i^{(j)})^{1+\varepsilon}]}{\mathbb{E}[(f_i^{(j)})]^{1+\varepsilon}}$ and denote the set of 'balanced' inputs of $f_i^{(j)}$ as:

$$B_i^{(j)} = \left\{ x \;\middle|\; \frac{1}{4\eta_i^{(j)}} \mathbb{E}[f_i^{(j)}] \leqslant f_i^{(j)}(x) \leqslant \left(4\eta_i^{(j)}\right)^{\frac{2(1+\varepsilon)}{\varepsilon}} \mathbb{E}[f_i^{(j)}] \right\}.$$

If $f_i^{(0)}$ is already $\left(\frac{1}{16\eta_i}, (25\eta_i)^{\frac{2(1+\varepsilon)}{\varepsilon}}\right)$-nice, set $k_i = 0$ and output $f'_i = \tilde{f}_i$. Otherwise for $j = 1, 2, \dots$ define $f_i^{(j)}$ by zeroing out all non-balanced inputs:

$$f_i^{(j)} := f_i^{(j-1)} \cdot \mathbf{1}_{B_i^{(j-1)}},$$

and define the stopping time k_i to be the first index j such that $\eta_i^{(j)} \geqslant \frac{1}{4}\eta_i^{(j-1)}$. This procedure terminates (i.e. k_i is finite) since each $\eta_i^{(j)} \geqslant 1$ by convexity. Every step the process doesn't stop decreases $\eta_i^{(j)}$ by a factor of 4, so there can be a total of $O(\log(\eta_i^{(0)}))$ steps.

It is immediate from definition that $0 \leqslant f'_i \leqslant f_i$, so we just need to show Item 2 and Item 3. The key is to observe that our balancing process cannot remove too much mass in each step:

$$\mathbb{E}[f_i^{(j-1)}] \geqslant \mathbb{E}[f_i^{(j)}] \geqslant \left(1 - \frac{1}{2\eta_i^{(j-1)}}\right) \mathbb{E}[f_i^{(j-1)}]. \tag{E.7}$$

We defer the proof, and first show Item 2 and Item 3 given this assumption.

Proof of Item 2. The proof is essentially immediate by the exponential decay of $\eta_i^{(j)}$. Namely by iterated application of (E.7) we have

$$\frac{\mathbb{E}[f_i^{(k)}]}{\mathbb{E}[f_i^{(0)}]} \geqslant \prod_{j=0}^{k-1}\left(1 - \frac{1}{2\eta_i^{(k-1-j)}}\right) \geqslant \prod_{j=0}^{\infty}\left(1 - \frac{1}{2\cdot 4^j}\right) \geqslant \frac{2}{5}$$

since by construction $\eta_i^{(j)} > 4^{k-1-j}\eta_i^{(k-1)} \geqslant 4^{k-1-j}$. Combining with the fact $\tilde{f}_i = f_i^{(0)} \geqslant \frac{f_i}{2}$ gives the result.

Proof of Item 3. Recall that for any j we have by construction the following 'pseudo'-balance condition:

$$\forall f_i^{(j)}(x) \neq 0 : \frac{1}{4\eta_i^{(j-1)}} \mathbb{E}[f_i^{(j-1)}] \leq f_i^{(j)}(x) \leq \left(4\eta_i^{(j-1)}\right)^{\frac{2(1+\varepsilon)}{\varepsilon}} \mathbb{E}[f_i^{(j-1)}].$$

When $j = k_i$ is the stopping time, observe that by (E.7) and construction respectively we have

1. $\mathbb{E}[f_i^{(k_i)}] \leq \mathbb{E}[f_i^{(k_i-1)}] \leq 2\mathbb{E}[f_i^{(k_i)}]$

2. $4\eta_i^{(k_i)} \geq \eta_i^{(k_i-1)}$.

Substituting into the above gives

$$\frac{1}{16\eta_i^{(k_i)}} \mathbb{E}[f_i^{(k_i)}] \leq f_i^{(k_i)}(x) \leq \left(25\eta_i^{(k_i)}\right)^{\frac{2(1+\varepsilon)}{\varepsilon}} \mathbb{E}[f_i^{(k_i)}]$$

as desired.

Proof of (E.7). The lefthand inequality is by construction. Toward the righthand, fix $f = f_i^{(j-1)}$ and $\eta = \eta_i^{(j-1)}$ for notational simplicity and define the set of 'terrible' inputs zeroed in the jth step as:

$$T_1 := \left\{ x \;\middle|\; f(x) < \frac{1}{4\eta} \mathbb{E}[f] \right\} \qquad T_2 := \left\{ x \;\middle|\; f(x) > (4\eta)^{\frac{2(1+\varepsilon)}{\varepsilon}} \mathbb{E}[f] \right\}$$

Since $T_1 \cup T_2 = V \setminus B_i^{(j-1)}$ and $f_i^{(j)} = \mathbf{1}_{B_i^{(j-1)}} f$, it is enough to show that

$$\mathbb{E}[f \cdot \mathbf{1}_{T_1}], \mathbb{E}[f \cdot \mathbf{1}_{T_2}] \leq \frac{1}{4\eta} \mathbb{E}[f]. \tag{E.8}$$

It is obvious that $\mathbb{E}[f \cdot \mathbf{1}_{T_1}] \leq \frac{1}{4\eta} \mathbb{E}[f]$ since f is upper bounded by this value within T_1. To bound the upper tail we use Hölder and Markov:

$$\mathbb{E}[f \cdot \mathbf{1}_{T_2}] \leq \|f\|_{1+\varepsilon} \mathbb{P}\left[f \geq (4\eta)^{\frac{2(1+\varepsilon)}{\varepsilon}} \mathbb{E}[f] \right]^{\frac{\varepsilon}{1+\varepsilon}} \qquad \text{(Hölder's inequality)}$$

$$\leq \|f\|_{1+\varepsilon} \frac{1}{16\eta^2} \qquad \text{(Markov's inequality)}$$

$$\leq \left(\frac{1}{4} \eta^{-\frac{\varepsilon}{1+\varepsilon}} \right) \frac{1}{4\eta} \mathbb{E}[f] \qquad (\|f\|_{1+\varepsilon} = \eta^{\frac{1}{1+\varepsilon}} \mathbb{E}[f])$$

$$\leq \frac{1}{4\eta} \mathbb{E}[f] \qquad (\eta^{-\frac{\varepsilon}{1+\varepsilon}} \leq 1)$$

as desired. □

Finally we prove reverse hypercontractivity for 'nice' functions (Proposition E.3).

Proof of Proposition E.3. Since f_i is $\left(\frac{1}{16\eta_i}, (25\eta_i)^{\frac{2(1+\varepsilon)}{\varepsilon}} \right)$-nice, it can attain at most

$$n_i = \left(3 + \frac{2}{\varepsilon} \right) \log \eta_i + \frac{10}{\varepsilon} + 15$$

non-zero values. With this in mind, decompose $f_i = \sum_{j \in \mathbb{Z}} 2^j \mathbf{1}_{f_i(x)=2^j}$ into level sets and observe that, by linearity of D, we can decompose the inner product itself into boolean sub-components and apply our assumed reverse hypercontractive inequality (Equation (E.6)):

$$\langle f_1^{(1+\varepsilon)\ell}, D f_2^{(1+\varepsilon)\ell} \rangle = \sum_{i,j \in \mathbb{Z}} 2^{(1+\varepsilon)\ell j} 2^{(1+\varepsilon)\ell i} \langle \mathbf{1}_{f_1(x)=2^j}, D \mathbf{1}_{f_2(x)=2^i} \rangle$$

$$\geqslant \kappa \sum_{i,j \in \mathbb{Z}} 2^{(1+\varepsilon)\ell j} 2^{(1+\varepsilon)\ell i}\, \mathbb{P}\left[f_1(x) = 2^j\right]^\ell \mathbb{P}\left[f_2(x) = 2^j\right]^\ell$$

$$= \kappa \left(\sum_{j \in \mathbb{Z}} (2^{(1+\varepsilon)j}\, \mathbb{P}\left[f_1(x) = 2^j\right])^\ell\right) \cdot \left(\sum_{j \in \mathbb{Z}} (2^{(1+\varepsilon)j}\, \mathbb{P}\left[f_2(x) = 2^j\right])^\ell\right).$$

Now since the sums have only n_1, n_2 non-zero terms respectively, applying $(\ell, \frac{\ell}{\ell-1})$-Hölder's inequality lower bounds this quantity by

$$\kappa(n_1 n_2)^{1-\ell} \left(\sum_{j \in \mathbb{Z}} 2^{(1+\varepsilon)j}\, \mathbb{P}\left[f_1(x) = 2^j\right]\right)^\ell \cdot \left(\sum_{j \in \mathbb{Z}} 2^{(1+\varepsilon)j}\, \mathbb{P}\left[f_1(x) = 2^j\right]\right)^\ell$$

$$= \kappa n_1^{1-\ell}\, \mathbb{E}[f_1^{1+\varepsilon}]^\ell \cdot n_2^{1-\ell}\, \mathbb{E}[f_2^{1+\varepsilon}]^\ell$$

$$\geqslant \kappa \cdot (n_1^{1-\ell}\eta_1^\ell) \cdot (n_2^{1-\ell}\eta_2^\ell) \cdot \mathbb{E}[f_1]^{\ell(1+\varepsilon)}\, \mathbb{E}[f_2]^{\ell(1+\varepsilon)}.$$

It can be checked directly that

$$\min_{\eta_i \geqslant 1} n_i \eta_i \geqslant \frac{1}{18 + \frac{12}{\varepsilon}}$$

which gives the desired result. $\qquad\square$

F   Splittability

In this section we discuss the relations between our notion of complete splittability, the splittability notion of [JST21], and the "τ-sampling" gaurantee of [DW22], and the inheritance of splittability under projection. We start by defining the latter two notions.

Definition F.1 (JST-Splittability [JST21]). $X \subset [n]^d$ is said to be λ-JST-splittable if for all $1 \leqslant a \leqslant t \leqslant b \leqslant d$:

$$\lambda_2(S_{[a,t],[t+1,b]}) \leqslant \lambda$$

In their later work, [DW22] introduced the weaker notion of τ-sampling which suffices for list-decoding.

Definition F.2. $X \subset [n]^d$ is said to be τ-sampling if for $t \in [d]$, $S \subset [n]$ and $W \subset [n]^t$:

$$\operatorname*{Cov}_{w \sim X}\left(1[w_{t+1} \in S], 1[w_{[t]} \in W]\right) \leqslant \tau.$$

Before arguing relations between these notions, we give the key lemma that underlies all proofs in this section: the expansion of swap walks is inherited under projection.

Lemma F.3 (Swap Walk Inheritance). *Let X be a partite d-uniform complex, and A and B disjoint subsets of $[d]$, and F any subset of $[d]$. Then:*

$$\lambda_2(S_{A \cap F, B \cap F}) \leqslant \lambda_2(S_{A,B})$$

Proof. By the variational characterization it is enough to show

$$\max_{\mathbb{E}[f^F] = 0} \frac{\|S_{A \cap F, B \cap F} f^F\|}{\|f^F\|_2} \leqslant \lambda_2(S_{A,B})$$

Given any such function $f^F : X[A \cap F] \to \mathbb{R}$ define its extension to $X[A]$ by $f(x) := f^F(x_F)$. Here $x_F = \{v \in x : col(v) \in F\}$ is the projection into X^F. and observe that this extension satisfies:

1. $\mathbb{E}_{X[A]}[f] = \mathbb{E}_{X[A \cap F]}[f^F] = 0$

2. $\|f\|_{2,X[A]} = \|f^F\|_{2,X[A \cap F]}.$

119

The result then follows from convexity, in particular:

$$
\begin{aligned}
\|S_{A\cap F, B\cap F} f^F\|_2 &= \mathbb{E}_{y\sim X[A\cap F]}\left[\left(\mathbb{E}_{x\in X(d)}[f^F(x_{B\cap F})|x_{A\cap F}=y]\right)^2\right]^{\frac{1}{2}} \\
&= \mathbb{E}_{y\sim X[A\cap F]}\left[\left(\mathbb{E}_{x\in X(d)}[f(x_B)|x_{A\cap F}=y]\right)^2\right]^{\frac{1}{2}} \\
&= \mathbb{E}_{y\sim X[A\cap F]}\left[\left(\mathbb{E}_{y'\sim X_y[A\setminus F]}\left[\mathbb{E}_{x\in X(d)}[f(x_B|x_{A\cap F}=y \wedge x_{A\setminus F}=y')]\right]\right)^2\right]^{\frac{1}{2}} \\
&\leqslant \mathbb{E}_{y\sim X[A]}\left[\left(\mathbb{E}_{x\in X(d)}[f(x_B|x_A=y]\right)^2\right]^{\frac{1}{2}} \qquad\qquad\text{(Jensen's Inequality)} \\
&= \|S_{A,B}f\|_2 \\
&\leqslant \lambda_2(S_{A,B})\|f\|_2 \qquad\qquad\qquad\qquad\qquad\qquad (\mathbb{E}[f]=\mathbb{E}[f^F]=0) \\
&= \gamma\|f^F\|_2
\end{aligned}
$$

$\square$

With this in mind, we prove the following relations between complete splittability, JST-splittability, and τ-sampling.

Proposition F.4. $X \subset [n]^t$ *is λ-JST-splittable if and only if it is completely λ-splittable. Moreover any completely λ-splittable complex is λ-sampling.*

Proof. If X is λ-JST-splittable it is clearly completely λ-splittable setting $a=1, b=d$. If X is completely λ-splittable, then for any $1\leqslant a\leqslant i\leqslant b\leqslant d$ we have:

$$
S_{[a,i],[i+1,b]} \leqslant S_{[1,i],[i+1,d]} \leqslant \lambda
$$

as desired. Finally, if X is completely λ-splittable, then for any $S\subset [n]$ and $X\subset[n]^t$ write $1_S := \mathbf{1}[w_{t+1}\in S]$ and $1_X := \mathbf{1}[w_{[t]}\in X]$

$$
\begin{aligned}
\text{Cov}_{w\sim X}(1_S, 1_X) &= \mathbb{E}_{y\sim X^{t+1}}[1_S(y)S_{t+1,[t]}1_X(y)] - \mathbb{E}[1_S][E][1_X] \\
&\leqslant \lambda_2(S_{t+1,[t]})\|1_S\|\|1_X\| \\
&\leqslant \lambda
\end{aligned}
$$

where we have used that by Hölder duality and Lemma F.3 $\lambda_2(S_{t+1,[t]}) = \lambda_2(S_{[t],t+1})\leqslant\lambda_2(S_{[t],t+1})\leqslant\lambda$. $\square$

Finally, we show that both complete splittability and standard tuple splittability are inherited under projections. Recall that given a d-partite complex X and $F\subseteq[d]$, X^F is the projection of X given by drawing a d-face from X, and projecting the resulting face onto the parts in F.

Proposition F.5 (Claim 10.34 Restated)**.** *Let $X\subset[n]^d$ be a homogeneous, completely λ-splittable complex. Then for any $F\subseteq[d]$ X^F is homogeneous and completely λ-splittable.*

Proof. The proof is essentially immediate from definition and Lemma F.3. Recall X is homogenous if the projection onto each part is uniform on $[n]$. Since projecting onto X^F and then a coordinate $i\in F$ is the same as just projecting onto i, homogeneity is inherited. Toward splittability, write $F=\{i_1,\ldots,i_{|F|}\}$ where $i_j<i_{j+1}$. Then for any i_j, applying Lemma F.3 gives

$$
\begin{aligned}
\lambda_2(S^{X^F}_{\{i_1,\ldots,i_j\},\{i_{j+1},\ldots,i_{|F|}\}}) &\leqslant \lambda_2(S^X_{\{i_1,\ldots,i_j\},\{i_{j+1},\ldots,i_{|F|}\}}) \\
&\leqslant \lambda_2(S_{[i_j],[i_j+1,d]})
\end{aligned}
$$

$$\leqslant \lambda$$

where we have used the fact that $\lambda_2(S^{X^F}_{\{i_1,\ldots,i_j\},\{i_{j+1},\ldots,i_{|F|}\}}) = \lambda_2(S^X_{\{i_1,\ldots,i_j\},\{i_{j+1},\ldots,i_{|F|}\}})$ (indeed they are the same operator). $\qquad\square$

Finally, while not strictly necessary for the results in this work, we show for completeness that inheritance of splittability also holds in the setting of general tuple-splitting trees as well.

Proposition F.6 (Projected Splitting Trees)**.** *Let X be any d-uniform partite complex with a λ-tuple splitting tree of depth D. For every $F \subseteq [d]$, X^F also has a λ-tuple splitting tree of depth at most D.*

The proof of Proposition F.6 relies on a simple inductive algorithm for projecting partite orderered binary trees (T, ρ) onto a subset of their coordinates. In particular, given such a tree and a subset $F \subseteq [d]$, we define the projection operation via the following inductive algorithm that modifies the tree T in place starting from a specific node $v \in T$.

Algorithm 1: $\text{Project}(T, \rho, F, v)$

Result: Projected partite ordered binary tree (T_v^F, ρ_v^F) rooted at v

1 **if** *v is a leaf* **then**
2 $\quad$ **return**
3 **else if** $\rho(\ell_v) \cap F = \emptyset$ **then**
4 $\quad$ $\text{Delete}(T(\ell_v))$;
5 $\quad$ $\text{Merge}(v, r_v)^a$;
6 $\quad$ $\text{Project}(T, \rho, F, r_v)$
7 **else if** $\rho(r_v) \cap F = \emptyset$ **then**
8 $\quad$ $\text{Delete}(T(r_v))$;
9 $\quad$ $\text{Merge}(v, \ell_v)$;
10 $\quad$ $\text{Project}(T, \rho, F, \ell_v)$
11 **else**
12 $\quad$ $\text{Project}(T, \rho, F, r_v)$;
13 $\quad$ $\text{Project}(T, \rho, F, \ell_v)$

---

^{a}Merge(v,r_v) deletes the node v and replaces r_v as the respective child of v's parent.

In other words, Algorithm 1 is the result of intersecting ρ with F and deleting the resulting nodes with empty or repeated labelings. We define the F-projection of a tree by applying this process at the root.

Definition F.7 (Projected Trees)**.** Let (T, ρ) be a d-uniform partite ordered binary tree and $F \subset [d+1]$ any coordinate subset. The F-projection of (T, ρ), denoted (T^F, ρ^F) is given by

$$T^F = \text{Project}(T, \rho, F, \text{root}(T)), \qquad \rho^F = (\rho \cap F)|_{T^F}.$$

We are now ready to prove Proposition F.6.

Proof of Proposition F.6. We will show that (X^F, T^F, ρ^F) is a λ-tuple splitting tree of depth at most D. We first argue that (X^F, T^F, ρ^F) is a partite ordered binary tree of depth at most D. First, observe that every operation in Algorithm 1 maps binary trees to binary trees, and cannot increase depth. In particular, the only modification of the tree structure occurs when a branch is deleted and the relevant root is contracted with its other child, maintaining the invariant that every internal node has one parent and two children, and that the depth is at most D. To see that the leaves are in bijection with F, observe that all leaves labeled by F survive the projection, all leaves labeled by $[d] \setminus F$ are deleted, and no new leaves are introduced by construction. The first two claims hold since a node v is deleted if and only if either $\rho(v) \cap F = \emptyset$, or the labeling is repeated later down the tree. Finally, to see that the children of any internal node in T^F partition its labeling in ρ^F, observe that this property holds trivially for T under the intersected labeling $\rho \cap F$ (albeit the partition may be trivial), and that this invariant is preserved by every operation in Algorithm 1.

Finally we argue the projected tree inherits the splittability of X^F. By Lemma F.3, we have that for any internal node u:

$$\lambda_2(S_{\rho^F(\ell_u),\rho^F(r_u)}) = \lambda_2(S_{\rho(\ell_u)\cap F,\rho(r_u)\cap F}) \leqslant \lambda_2(S_{\rho(\ell_u),\rho(r_u)}) \leqslant \lambda$$

as desired. □