Electronic Colloquium on Computational Complexity, Report No. 119 (2023)

Agreement theorems for high dimensional expanders in the small soundness regime: the role of covers

Yotam Dikstein* and Irit Dinur†

August 18, 2023

Abstract

Given a family X of subsets of $[n]$ and an ensemble of local functions $\{f_s : s \to \Sigma \mid s \in X\}$, an agreement test is a randomized property tester that is supposed to test whether there is some global function $G : [n] \to \Sigma$ such that $f_s = G|_s$ for many sets s. For example, the V-test chooses a random pair of k-element subsets that intersect on $\sqrt{k}$ elements, and accepts if the local functions agree on the common elements.

The small soundness (or 1%) regime is concerned with the structure of ensembles $\{f_s\}$ that pass the test with small but non-negligible probability $\mathrm{Agree}(\{f_s\}) \geqslant \varepsilon > 0$. A "classical" small-soundness agreement theorem is a list-decoding (LD) statement, saying that

$$\mathrm{Agree}(\{f_s\}) > \varepsilon \quad \Longrightarrow \quad \exists G^1, \dots, G^\ell, \quad \mathop{\mathbb{P}}_s[f_s \overset{0.99}{\approx} G^i|_s] \geqslant \mathrm{poly}(\varepsilon),\ i = 1, \dots, \ell. \qquad (LD)$$

Such a statement is motivated by PCP questions and has been shown in the case where $X = \binom{[n]}{k}$, or where X is a collection of low dimensional subspaces of a vector space. In this work we study small soundness behavior of agreement tests on high dimensional expanders X. Such complexes are known to satisfy agreement tests in the high soundness (99%) regime, and it has been an open challenge to analyze their small soundness behavior.

Surprisingly, the small soundness behavior turns out to be governed by the topological covers of X. We show that:

1. If X has no connected covers, then a "classical" small soundness theorem as in (LD) holds, provided that X satisfies an additional expansion property.
2. If X has a connected cover, then "classical" small soundness as in (LD) necessarily fails.
3. If X has a connected cover (and assuming the additional expansion property), we replace the failed (LD) by a slightly weaker statement, which we call lift-decoding:

$$\mathrm{Agree}(\{f_s\}) > \varepsilon \Longrightarrow \quad \exists \text{ cover } \rho : Y \twoheadrightarrow X, \text{ and } G : Y(0) \to \Sigma, \text{ such that} \qquad (LFD)$$
$$\mathop{\mathbb{P}}_{\tilde{s} \twoheadrightarrow s}[f_s \overset{0.99}{\approx} G|_{\tilde{s}}] \geqslant \mathrm{poly}(\varepsilon),$$

where $\tilde{s} \twoheadrightarrow s$ means that $\rho(\tilde{s}) = s$.

The additional expansion property is cosystolic expansion of a complex derived from X. This property holds for the spherical building and for quotients of the Bruhat-Tits building, giving us new examples for set systems with small soundness agreement theorems, a la (LFD).

*Weizmann Institute of Science, ISRAEL. email: yotam.dikstein@weizmann.ac.il.

†Weizmann Institute of Science, ISRAEL. email: irit.dinur@weizmann.ac.il. Both authors are supported by Irit Dinur's ERC grant 772839, and ISF grant 2073/21.

ISSN 1433-8092

Contents

1 Introduction

A function $G : [n] \to \Sigma$ can be specified by a truth table, or, alternatively, by providing its restrictions to a pre-determined family of subsets of $[n]$. Namely, given a family of subsets $X = \{s \subset [n]\}$, the function G is represented by an ensemble of "local functions" $\{f_s : s \to \Sigma \mid s \in X\}$ such that $f_s = G|_s$. Such a representation has built-in redundancy which potentially can be used for amplifying distances while providing local testability. This is why such encodings are prevalent within PCP reductions, where a global PCP proof is broken up into many possibly-overlapping pieces and each piece is further encoded by an inner PCP gadget (see below for more details and references). The PCP verifier is tasked, among other things, with testing whether the ensemble of pieces corresponds to a coherent global proof.

The agreement testing question is as follows. Given an ensemble $\{f_s \mid s \in X\}$, test whether there is some global function $G : [n] \to \Sigma$ such that $f_s = G|_s$ for most s. The PCP setting imposes a stringent requirement that the number of queries be as small as possible, with two queries being the golden standard, as dictated by applications to tight hardness of approximation. The natural two-query test is to select a pair of overlapping subsets s_1, s_2 according to a pre-specified distribution[1] and to check whether f_{s_1}, f_{s_2} agree. Namely, accept iff $f_{s_1}(i) = f_{s_2}(i)$ for all $i \in s_1 \cap s_2$. Such a test is called an *agreement test* and we denote the success probability of the test by $\mathrm{Agree}(\{f_s\})$.

Where agreement tests are concerned, there are two main regimes of interest. The "99% regime" which is natural in the world of property testing, and the "1% regime" which is the regime of interest in the world of PCPs.

- **The 99% regime**. We are given an ensemble $\{f_s\}_{s\in X}$ that passes the agreement test with probability close to 1. We want to conclude that there is a single $G : [n] \to \Sigma$ such that $\{G|_s\}_s \approx \{f_s\}_s$. Formally, an agreement testing theorem describes a distribution $\mathcal{D}$ over pairs of subsets from X for which

$$\mathrm{Agree}(\{f_s\}) > 0.99 \quad \Longrightarrow \quad \mathbb{P}\left[f_s = G|_s\right] \geqslant 1 - O(\varepsilon) \text{ for some } G : [n] \to \Sigma. \tag{1.1}$$

- **The 1% regime**. Here we assume that the test succeeds with much smaller probability $\varepsilon > 0$ for a small positive constant ε, say $\varepsilon = 0.01$, or even $\varepsilon = o_k(1)$. The hope is that even this modest success probability is sufficiently non-negligible and implies that the ensemble $\{f_s\}$ has some global structure. The 1% regime, also called the small-soundness regime, is especially important in PCP settings, where small ε translates to a larger gap between completeness and soundness of the PCP. It is well known that in the 1% regime, one cannot expect to find a single global function g, but at best a *short list* of functions $G^1, \ldots, G^\ell : [n] \to \Sigma$ such that each is supported on a $\mathrm{poly}(\varepsilon)$ fraction of the sets $s \in X$, namely,

$$\mathrm{Agree}(\{f_s\}) > \varepsilon \quad \Longrightarrow \quad \exists G^1, \ldots, G^\ell, \quad \mathbb{P}_s[f_s \overset{0.99}{\approx} G^i|_s] \geqslant \mathrm{poly}(\varepsilon),\ i = 1, \ldots, \ell. \tag{LD}$$

  Note that the conclusion allows approximate agreement between f_s and $G^i|_s$ (meaning that they agree on almost all of s[2]). This is unavoidable, see discussion in [DG08].

  In words, (LD) says that 1% agreement implies *list-decoding*.

The local testing problem is quite non-trivial even when $X = \binom{[n]}{k}$, where we think of k as a sufficiently large constant while $n \to \infty$. In this case the function $G : [n] \to \Sigma$ is represented by its restriction to all possible k-element subsets. The ensemble $\{G|_s\}_s$ is called the (symmetrized) direct product encoding of G. This encoding has been studied both in the context of hardness amplification (see, e.g. [Imp+08]), and as a generalization of PCP low degree tests, as initiated by [GS97]. Following a sequence of works, [GS97; DR06; DG08; IKW12; DS14b; DL17], it is known, by now, that $\binom{[n]}{k}$ satisfies (LD).

Families X that are sparser than $\binom{[n]}{k}$ were considered in later works [IKW12; DK17; DD19]. Such families give rise to so-called derandomized direct product encodings. The size of the family X is important since

[1]Throughout the introduction we suppress the precise distribution of the two-query test. For concreteness, think of the distribution given by selecting two k-sets that intersect on some k' elements, and for example $k' = \sqrt{k}$. The distributions use are precisely defined in Section 2.2.

[2]The notation $f_s \overset{1-\delta}{\approx} G|_s$ throughout the paper means that $f(v) = G(v)$ on a $1-\delta$-fraction of s.

it controls the efficiency of the encoding. A major motivation for studying short encodings is to be able to construct more length-efficient PCPs that still have a large gap between the completeness and soundness cases. A good agreement test can often be transformed into a PCP construction, although this is not immediate. [IKW12] showed a black-box way in which their agreement testing theorem for $X = \binom{[n]}{k}$ implies a PCP theorem with large gap; which, in a way, is an alternative to Raz's parallel repetition theorem [Raz98]. This work also constructed a derandomized family X with a good agreement test, which took more work [DM11] to transform it into an efficient PCP theorem with a large gap. Here efficiency pertains to the length of the PCP proof. This parameter is important for various reasons including actual implementations of PCPs as well as hardness of approximation consequences (see for example discussion in [MR10]).

The recent "discovery" of bounded-degree high dimensional expanders by the theoretical computer science community has lead to a hope that these might provide highly derandomized families X that support good agreement tests. In [DK17] (and later improved in [DD19]) it was shown that if X is a spectral high dimensional expander (see Section 2 for precise definitions), then it supports a 99% agreement theorem as in (1.1). This gave the first family X of subsets whose size is linear in the size n of the ground set and that supports a 99% agreement theorem. Naturally, this raised the question of 1%:

Question 1.1. Does spectral high dimensional expansion suffice for 1% agreement a la (LD)?

The short answer is no (and the longer answer is yes after some modification). First, we show that many complexes X that satisfy (1.1), fail to satisfy (LD) in quite a strong way:

Lemma 1.2. *Let $\delta < \frac{1}{2}$, $k \in \mathbb{N}$, and let $\lambda \leqslant \exp(-7k)$. Let X be a k-dimensional λ-two-sided high dimensional expander, and assume X has a connected 2-cover. Then there exists an ensemble of functions $\mathcal{F} = \{f_r : r \to \{0,1\} \mid\ r \in X(k)\}$ such that Agree$(\mathcal{F}) \geqslant \frac{1}{2}$, and yet for every $G : X(0) \to \{0,1\}$ it holds that*

$$\mathop{\mathbb{P}}_{r \in X(k)}\left[f_r \overset{1-\delta}{\approx} G|_r\right] \leqslant \exp(-\Omega_\delta(k)).$$

This lemma reveals connected 2-covers (and more generally connected t-covers) to be an obstruction to 1% agreement theorems. The notion of a connected cover is a standard topological notion, which generalizes the notion of a lift of a graph to higher dimensions. In a nutshell a 2-cover is a 2-to-1 homomorphism $\rho : Y \twoheadrightarrow X$ such that the preimage of every face of X is a pair of disjoint faces in Y, (see Section 2 for details).

In order to see how to address this obstruction, let us take a detour to describe the counter-example driving Lemma 1.2.

1. Starting with a 2-cover $Y \twoheadrightarrow X$, we fix a global function $H : Y(0) \to \{0,1\}$ on its vertices, with the property that every pair of vertices $(v,0),(v,1)$ that cover the same $v \in X(0)$, are given distinct values $H(v,0) \neq H(v,1)$.

2. Every face $s \in X(k)$ is covered by two faces $\tilde{s}_1, \tilde{s}_2 \in Y(k)$, i.e. $\rho(\tilde{s}_i) = s$. Thus, we naturally get a list of two possible assignments to f_s that biject down from either $h|_{\tilde{s}_1}$ or $h|_{\tilde{s}_2}$ (we remark that by construction, these two assignments are negations of one another).

   Moreover, the lists of intersecting faces "agree" with each other in a way that forms a bijection. This bijection perfectly recovers the structure of the cover.

3. Our final ensemble for X is obtained by choosing, for each s independently at random, one of the two assignments from the list. It is now easy to see that the agreement of this ensemble is about $1/2$. Moreover, sampling arguments (relying on the high dimensional expansion of X, Y) will show that no function $G : X(0) \to \{0,1\}$ can agree with more than $\exp(-k)$ of the subsets.

We remark that this example can be viewed as a generalization of an example of Bogdanov [Bog05] showing that gap amplification fails below $1/2$.

Observe that at the end of step 2, we have a unique games instance on our hands. The instance has vertices $X(k)$ and an edge connects s, s' if they intersect as subsets. Every s has two possible assignments and the unique games constraint between s and s' is given by the bijection between the corresponding lists. As with any unique games instance with alphabet size 2, the value is at least $1/2$, which corresponds to the agreement of our ensemble being at least $1/2$. A similar unique games instance also arises naturally in the proof of our agreement theorem below.

How can we reconcile Lemma 1.2 with previous works [DG08; IKW12; DS14b; DL17] showing that the complete complex $X = \binom{[n]}{k}$ supports a 1% agreement theorem a la (LD) while being an excellent high dimensional expander? The reason Lemma 1.2 doesn't apply is that the complete complex *doesn't have connected covers* at all. This suggests cover-freeness (equivalently, simply-connectedness) as a sufficient condition, in addition to high dimensional expansion, for (LD).

Supporting this hypothesis, we prove a new 1% agreement theorem for spherical buildings. These are spectral high dimensional expanders that are simply connected and hence have no connected covers.

Theorem 1.3 (1% agreement implies list-decoding for the spherical building)**.** *Let $k \in \mathbb{N}$, and let $\varepsilon > \Omega(\log 1/k)$. Let $d > k$ be sufficienty large and let X be a d-dimensional spherical building that is a $\lambda = 2^{-7d}$ high dimensional expander. Then X supports a 1% agreement theorem as in* (LD).

This theorem gives a new record for the shortest derandomized direct product encoding that supports a 1% agreement theorem. The number of k faces in a spherical building X is at most roughly quadratic in the size of the ground set. (Compare $|X(d)| \approx q^{d^2/2}$ to $n = |X(0)| \approx q^{d^2/4}$, and for $k < d$ clearly $|X(k)| \leqslant 2^d \cdot |X(d)|$.) The previously best derandomization is given in [IKW12], where the number of subsets is at least n^{25} where n is the size of the ground set.

Question 1.4. Are there linear-size families that support a 1% agreement theorem a la (LD)?

A major hope has been that high dimensional expanders provide such families. However, known high dimensional expanders are not cover-free (and this seems inherent as they are constructed by a sequence of quotients), so Lemma 1.2 poses an obstruction. Instead, we show that when the complex X has an additional expansion property, the counterexample described in Lemma 1.2 is the only obstruction to a 1% agreement theorem. More elaborately, we show under the additional property that if an ensemble $\{f_s\}_{s\in X}$ satisfies $\mathrm{Agree}(\{f_s\}) > \varepsilon$, then there must be an ℓ-cover $\rho : Y \twoheadrightarrow X$ and a global function $G : Y(0) \to \{0,1\}$ such that $\ell = \mathrm{poly}(1/\varepsilon)$ and for $\mathrm{poly}(\varepsilon)$ fraction of the sets $s \in X(k)$, f_s is explained by G. Namely, f_s is one of the functions in the list $G|_{\tilde{s}_1}, \dots G|_{\tilde{s}_\ell}$ (after applying the appropriate bijection from $\tilde{s}_i \leftrightarrow s$).

In other words, we show that every agreeing ensemble must come from a global function, perhaps not on the original complex, but on a cover (aka lift) of the complex. This revises the decades-long paradigm of "1% agreement implies list-decoding" to "1% agreement implies li**f**t-decoding.".

The additional property is cosystolic expansion of some associated complex described in the proof overview and in Theorem 1.7. This property holds in complexes whose links are spherical buildings, which includes the bounded-degree high dimensional expanders of [LSV05b; LSV05a]. Hence we get the following lift-decoding theorem.

Theorem 1.5. *Let $k \in \mathbb{N}$, and let $\varepsilon > \Omega(\log 1/k)$. Let $d > k$ be sufficienty large and let X be a d-dimensional $\lambda = 2^{-7d}$ high dimensional expander, whose links are spherical buildings.*

For any ensemble $\{f_s\}_{s\in X(k)}$ that satisfies $\mathrm{Agree}(\{f_s\}) > \varepsilon$, there must exist a $\mathrm{poly}(1/\varepsilon)$-cover $\rho : Y \twoheadrightarrow X$, and a global function $G : Y(0) \to \Sigma$, such that

$$\mathbb{P}_s[f_s \text{ is explained by } G] \geqslant \mathrm{poly}(\varepsilon).$$

Observe that Theorem 1.3 is a special case of this theorem, since whenever X is simply connected, the only possible ℓ-cover Y of X consists of ℓ disjoint isomorphic copies of X. In this case, the global function on $Y(0) = X(0) \times [\ell]$ can be interpreted as a list of ℓ global functions on $X(0)$.

Finally, we point out that there is still some hope for answering Question 1.4 positively. Even when X is not simply connected, there is a quantitative aspect connecting the agreement parameter ε and the size ℓ of the cover, which may imply a 1% list agreement theorem. I.e., if X has the additional property that all of its ℓ-covers for $\ell < \mathrm{poly}(1/\varepsilon)$ are necessarily disconnected, then we immediately get a 1% agreement theorem with a list-agreement conclusion, a la (LD). This motivates the study and construction of complexes with no small connected covers. A construction of such a family of complexes would give a linear-size family of complexes supporting a 1% list-agreement theorem.

Proof overview

We next turn to describing the technical ingredients that go into the proof. Like many proofs in the world of high dimensional expansion, our proof has a local-to-global structure. We essentially show how to "lift" agreement theorems from the complete complex to high dimensional expanders.

An agreement test on the complete complex is a distribution over k-element subsets of $[n]$. Let us focus for concreteness on the V-test distribution from [IKW12], in which two k-sets are sampled uniformly, conditioned on their intersection size being $\sqrt{k}$.

Moving to a high dimensional expander X, the corresponding distribution $\mathcal{D}_X$ is obtained by selecting a face of size $2k-\sqrt{k}$, and inside it two faces r_1, r_2 whose sizes are k and such that they intersect on a set of size $\sqrt{k}$. Our proof goes through the following steps.

Constructing a local list for every face

Fix $k \ll d_1 \ll d$. First we use the previous agreement theorem in [IKW12] to construct lists $\left\{\overline{L}(s) \mid s \in X(d_1)\right\}$ that list-decode the local ensembles of functions $\{f_r\}_{r\subseteq s} \subseteq \{f_r\}_{r\in X(k)}$, as in (LD). Here we treat every $s \in X(d_1)$ as a copy of the complete complex along with its own local ensemble. To do so, we describe $\mathcal{D}_X$ in an equivalent local way:

1. Choose $s \in X(d_1)$ at random.
2. Choose $r_1, r_2 \subset s$ according to the original test distribution $\mathcal{D}$ on the complete complex on $|s|$-vertices.

Given an ensemble with $\operatorname{Agree}(\{f_s\}) > \varepsilon$, a standard sampling argument will show that for a typical face $s \in X(d_1)$ the success of the test conditioned on the two faces being contained in s, is still about ε for nearly all sets $s \in X(d_1)$. So by considering these local ensembles, we can use the previous agreement theorem from [IKW12] to obtain a list decoding $\overline{L}(s)$ for $\{f_r\}_{r\subseteq s}$, for almost all $s \in X(d_1)$. We also find lists for all faces $t \in X(2d_1+1)$ and $u \in X(3d_1+2)$.

Showing these lists match each other

Next, we carefully modify the previous argument so that the obtained lists will be compatible with each other. More precisely, we show that one can take lists such that there is some integer ℓ so that $|\overline{L}(s)| = \ell$ for nearly all s. Moreover, denoting $\overline{L}(s) = \{L_s^1, L_s^2, \dots, L_s^\ell\}$, for most $s \subseteq t$, there is a permutation $\pi_{s,t} : [\ell] \to [\ell]$ where $\pi_{s,t}(i) = j$ if and only if $L_s^i \approx L_t^j|_s$.

We remark that these two steps are similar to [Din+18a], where lists of assignments were derived for every face. However, while it is easy to see that a noticeable fraction of the lists match (which was what [Din+18a] showed), this is not enough for us, and we must show that they match with probability 99% for the next steps to work. This causes significant technical complications but turns out to be possible.

Unique games on the *faces* complex

Before stating the next steps, we shift our point of view from X to its *faces complex*.

Definition 1.6 (The Faces Complex)**.** Let X be a d-dimensional simplicial complex. Let $d_1 \leqslant d$. We denote by $FX = F(X, d_1)$ the simplicial complex whose vertices are $FX(0) = X(d_1)$ and whose faces are all $\left\{\{s_0, s_1, ..., s_j\} \mid s_0 \mathbin{\dot\cup} s_1 \mathbin{\dot\cup} \cdots \mathbin{\dot\cup} s_j \in X\right\}$.

The 1-skeleton of FX is a graph whose vertices are the d_1-sets. Each vertex $s_1 \in FX$ has a list of assignments, and (almost) every edge between s_1 and s_2 carries a bijection between the assignments of s_1 and the assignments of s_2. The bijection is just the composition $\psi_{s_1,s_2} = \pi_{s_2,t}^{-1} \circ \pi_{s_1,t}$ where $t = s_1 \mathbin{\dot\cup} s_2$. This can be viewed as a unique games instance over FX. We show that for a most triangles $s_1 s_2 s_3 \in FX$, there are no local contradictions. Namely, if we denote by ψ_{s_i,s_j} the permutation from the list of s_i to the list of s_j, then

$$\psi_{s_2 s_3} \circ \psi_{s_1 s_2} = \psi_{s_1 s_3} \tag{1.2}$$

nearly always. In topological language this collection of bijections gives us a 1-cochain in $C^1(FX, Sym(\ell))$ where $Sym(\ell)$ is the groups of permutations over ℓ elements. The consistency implies that this 1-cochain is nearly a cocycle.

At this point, we wish to tweak the collection of bijections that are nearly consistent, and create a new collection that is consistent everywhere. This is possible if and only if the faces complex is a cosystolic expander (in fact, this is precisely the definition of a cosystolic expander). Unfortunately, we do not know whether the faces complex is always a sufficiently strong cosystolic expander, even if we assume that X itself is a cosystolic expander. However, we prove that FX is a cosystolic expander whenever X is a high dimensional expander whose links are spherical buildings. This appears in a companion paper [DD23b].

Constructing the cover for FX

It is well known that any collection of permutations $\{\pi_{uv}\}_{uv\in FX(1)}$ with *no* local contradictions[3] corresponds to a cover $\rho_{FX} : \widetilde{FX} \to FX$ (see e.g. [Sur84]). Let us describe the underlying graph of the cover. Its vertices are pairs (s, i) for every $s \in FX(0)$ and $i \in [\ell]$. We connect (s_1, i) to (s_2, j) by an edge in $\widetilde{FX}$ if $s_1 s_2 \in FX(1)$ and $\pi_{s_1,s_2}(i) = j$. The fact that (1.2) holds for every triangle is the reason that this underlying graph is indeed a graph skeleton of a cover $\rho_{FX} : \widetilde{FX} \to FX$ (see Definition 2.19 for the formal construction).

The cover we get from this step is a cover of the faces complex FX. However, we show that every cover of $\widetilde{FX}$ of FX must come from a cover $\rho : Y \to X$, where "comes from" in this context means that there is an isomorphism from the faces complex of Y to $\widetilde{FX}$, $\iota : FY \xrightarrow{\sim} \widetilde{FX}$. This complex Y is the cover on which we define the global function.

The use of cosystolic expansion for constructing covers from permutations with *few local contradictions* was developed in [DM22]. Here we instantiate this idea, using the permutations defined by the first two steps.

The idea behind this last step is close in spirit to list decoding argument in [GK22]. Their work build upon [DM22] to list-decode ensembles of functions that have a 99%-regime list-agreement guarantee, using a closely related notion of coboundary expansion. They derive a complex similar (but not identical) to the faces complex, which they call the representation complex. Then they use their list-agreement test to construct permutations that have almost no local contradictions, and use coboundary expansion to transform this to a list of global functions on X. Unfortunately, we could not have adapted their framework to our purposes; the notion of 99%-regime list-agreement they require is too strong to hold in our case. In addition, coboundary expansion does not capture complexes that are not simply connected, and we require a method to deal with such complexes.

From a cover to a global function via 99%-agreement

We are nearly done, but we still need to construct the global function $G : Y(0) \to \Sigma$. We define an ensemble of functions $\{h_{\tilde{s}}\}_{\tilde{s}\in Y(d_1)}$, where $h_{\tilde{s}}$ colors the vertices according to $L_s^i \in \overline{L}(s)$, such that $\iota(\tilde{s}) = (s, i)$. Opening up the definition of the isomorphism ι and the permutations $\pi_{s_1,t}$ above, we show that there is a agreement distribution $\mathcal{D}_Y$ where

$$\mathrm{Agree}_{\mathcal{D}_Y}(\{h_{\tilde{s}}\}) \geqslant 99\%.$$

Then all that remains is to use a known 99%-agreement theorem a la (1.1) to obtain a global function $G : Y(0) \to \Sigma$ that agrees with most $h_{\tilde{s}}$. This is (finally!) the global function that explains the original distribution $\{f_r\}$.

General criterion for agreement

Theorem 1.3 and Theorem 1.5 are written for specific complexes X, requiring their links to be spherical buildings. As observed in the proof overview, the only property we need from X besides high dimensional expansion is cosystolic expansion of the faces complex $F(X, d_1)$ for a sufficiently high level d_1.

Theorem 1.7 (Informal, see Theorem 3.1 for a formal statement)**.** *Let $k \in \mathbb{N}$, and let $\varepsilon > \Omega(\log 1/k)$. Let $d > k$ be sufficiently large and let X be a d-dimensional $\lambda = 2^{-7d}$ high dimensional expander. Suppose that there exists some $k^3 \leqslant d_1 \leqslant \frac{k \log d}{\mathrm{poly}(\varepsilon)}$ such that $F(X, d_1)$ is an $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$-cosystolic expander.*

[3] I.e. (1.2) holds for every triangle.

For any ensemble $\{f_s\}_{s\in X(k)}$ that satisfies Agree$(\{f_s\}) > \varepsilon$, there must exist a $\mathrm{poly}(1/\varepsilon)$*-cover $\rho : Y \twoheadrightarrow X$, and a global function $G : Y(0) \to \{0,1\}$, such that*

$$\mathop{\mathbb{P}}_{s}[f_s \textit{ is explained by } G] \geqslant \mathrm{poly}(\varepsilon).$$

Both our theorems follow from this more general statement, along with a theorem that states that the faces complex of spherical buildings (and of complexes whose links are spherical buildings) are cosystolic expanders; such a theorem is proven in a companion paper [DD23b].

Note that the quantitative cosystolic expansion needed in Theorem 1.7 is rather modest; it is allowed to decay to zero exponentially in d_1/k.

Related work

Agreement tests in the 1% regime appear in several PCP constructions, such as in the plane-versus-plane [RS97] and line-versus-line low degree tests [AS97] that lead to PCP theorems with large gap between completeness and soundness. PCPs with large gaps are used as outer verifiers in hardness of approximation applications.

The large-gap PCP that is the most widely used in hardness of approximation reductions is the label cover PCP which is based on the parallel repetition theorem [Raz98]. Parallel repetition is essentially the direct product encoding, and the analysis is highly related to analysis of agreement tests on $X = \binom{[n]}{k}$, as can be seen both in [IKW12] who give an agreement-based proof of a parallel repetition theorem, and also in [DS14b; DS14a].

More recently agreement tests appear in the proof of the 2:2 theorem [Din+18b; KMS18], where X is the collection of all ℓ-dimensional subspaces of a vector space (the Grassmannian), and the agreement test compares two subspaces that overlap on an $\ell - 1$ dimensional subspace. This test constitutes the 2:2 inner verifier. It was shown [BKS19] that in this setting agreement follows from small set expansion of the underlying complex. Here as well as in other low degree tests the local functions in the ensembles are restricted to having additional structure, namely being linear or having low degree. This gives the theorems a slightly different flavor (for example, it is much easier to get separation between list elements).

Moshkovitz and Raz [MR10] constructed a label cover PCP with nearly-linear length $n \cdot 2^{(\log n)^{0.99}}$, motivated by applications to hardness of approximation (efficient-length reductions give a much stronger connection between exponential hardness assumptions on exact and approximation problems). Getting the length further down (possibly to $n \cdot \mathrm{poly}\log n$) is an open question. Agreement tests for very short families of subsets may lead to progress in this direction.

Agreement tests were first formulated in [GS97]. Ensuing works focused mainly on the 1% regime for its importance in PCP applications, but the 99% regime is even more basic and natural from a property testing perspective. This was studied in [GS97; DK17; DD19; DFH19; GK20; GK22].

Finally, we have learned of very recent independent work by Bafna and Minzer on 1% agreement theorems on high dimensional expanders [BM23].

Acknowledgements

We thank Roy Meshulam and Amitay Kamber for insightful early discussions on the topic.

2 Preliminaries

2.1 High dimensional expanders

Most of the definitions in this subsection are standard, with the exception of the definition of the non-lazy up down walk Definition 2.5.

A pure d-dimensional simplicial complex X is a hypergraph that consists of an arbitrary collection of sets of size $(d+1)$ together with all their subsets. The sets of size $i+1$ in X are denoted by $X(i)$. The vertices of X are denoted by $X(0)$ (we identify between a vertex v and its singleton $\{v\}$). We will sometimes omit set brackets and write for example $uvw \in X(2)$ instead of $\{u,v,w\} \in X(2)$. As a convention $X(-1) = \{\emptyset\}$. Let X be a d-dimensional simplicial complex. Let $k \leqslant d$. We denote the set of oriented k-faces in X by $\vec{X}(k) = \{(v_0, v_1, ..., v_k) \mid \{v_0, v_1, ..., v_k\} \in X(k)\}$.

For $k \leqslant d$ we denote by $X^{\leqslant k} = \bigcup_{j=-1}^{k} X(j)$ the k-skeleton of X. When $k = 1$ we call this complex *the underlying graph of* X, since it consists of the vertices and edges in X (as well as the empty face).

A *clique complex* is a simplicial complex such that if $s \subseteq X(0)$ has that if s is a clique, that is, for every two vertices $v, u \in s$ the edge $vu \in X(1)$, then $s \in X$.

A $(d+1)$-partite d-dimensional simplicial complex is a generalizeation of a bipartite graph. It is a complex X such that one can decompose $X(0) = A_0 \mathbin{\dot\cup} A_1 \mathbin{\dot\cup} \cdots \mathbin{\dot\cup} A_d$ such that for every $s \in X(d)$ and $i \in [d]$ it holds that $|s \cap A_i| = 1$.

Probability over simplicial complexes

Let X be a simplicial complex and let $\mathbb{P}_d : X(d) \to (0,1]$ be a density function on $X(d)$ (that is, $\sum_{s \in X(d)} \mathbb{P}_d(s) = 1$). This density function induces densities on lower level faces $\mathbb{P}_k : X(k) \to (0,1]$ by $\mathbb{P}_k(t) = \frac{1}{\binom{d+1}{k+1}} \sum_{s \in X(d), s \supset t} \mathbb{P}_d(s)$. We can also define a probability over directed faces, where we choose an ordering uniformly at random. Namely, for $s \in \vec{X}(k)$, $\mathbb{P}_k(s) = \frac{1}{(k+1)!} \mathbb{P}_k(set(s))$ (where $set(s)$ is the set of vertices participating in s). When it's clear from the context, we omit the level of the faces, and just write $\mathbb{P}[T]$ or $\mathbb{P}_{t \in X(k)}[T]$ for a set $T \subseteq X(k)$.

Links an high dimensional expansion

Let X be a d-dimensional simplicial complex and let $s \in X$ be a face. The link of s is the $d' = d - |s|$-dimensional complex

$$X_s = \{t \setminus s \mid t \in X, t \supseteq s\}.$$

For a simplicial complex X with a measure $\mathbb{P}_d : X(d) \to (0,1]$, the induced measure on $\mathbb{P}_{d', X_s} : X_s(d - |s|) \to (0,1]$ is

$$\mathop{\mathbb{P}}_{d', X_s}(t \setminus s) \propto \frac{\mathbb{P}_d(t)}{\sum_{t' \supseteq s} \mathbb{P}_d(t')}.$$

We denote by $\lambda(X_s)$ to be the (normalized) second largest eigenvalue of the adjacency operator of $X_s^{\leqslant 1}$. We denote by $|\lambda|(X_s)$ to be the (normalized) second largest eigenvalue of the adjacency operator of $X_s^{\leqslant 1}$ *in absolute norm.*

Definition 2.1 (High dimensional expander)**.** Let X be a d-dimensional simplicial complex and let $\lambda \in (0,1)$. We say that X is a λ-*one sided high dimensional expander* if for every $s \in X^{\leqslant d-2}$ it holds that $\lambda(X_s) \leqslant \lambda$. We say that X is a λ-*two sided high dimensional expander* if for every $s \in X^{\leqslant d-2}$ it holds that $|\lambda|(X_s) \leqslant \lambda$.

We stress that this definition includes $s = \emptyset$, which also implies that $X^{\leqslant 1}$ should have a small second largest eigenvalue.

Walks on high dimensional expanders

Let X be a d-dimensional simplicial complex. Let $\ell \leqslant k \leqslant d$. The (k,ℓ)-containment graph $G_{k,\ell} = G_{k,\ell}(X)$ is the bipartite graph whose vertices are $L = X(k), R = X(\ell)$ and whose edges are all (t,s) such that $t \supseteq s$. The probability of choosing such an edge is as in the complex X.

Theorem 2.2 ([KO20])**.** *Let X be a d-dimensional λ-one sided high dimensional expander. Let $\ell \leqslant k \leqslant d$. Then the second largest eigenvalue of $G_{k,\ell}(X)$ is upper bounded by $\lambda(G_{k,\ell}(X)) \leqslant \frac{\ell+1}{k+1} + O(k\lambda)$.*

A corollary proven in [DK17] is that this graph is also a good sampler.

Corollary 2.3 ([DK17])**.** *Let $A \subseteq X(k)$ and let $\delta > 0$. Let*

$$B(A) = \left\{ s \in X(\ell) \,\middle|\, \left| \mathop{\mathbb{P}}_{t \supseteq s}[A] - \mathbb{P}[A] \right| > \delta \right\}.$$

Then $\mathbb{P}[B(A)] = O(\frac{\ell+1}{(k+1)\delta^2}\mathbb{P}[A])$.

A related walk is the swap walk. Let k, ℓ, d be integers such that $\ell + k \leqslant d - 1$. The k,ℓ-swap walk $S_{k,\ell} = S_{k,\ell}(X)$ is the bipartite graph whose vertices are $L = X(k), R = X(\ell)$ and whose edges are all (t, s) such that $t \mathbin{\dot\cup} s \in X$. The probability of choosing such an edge is the probability of choosing $u \in X(k+\ell+1)$ and then uniformly at random partitioning it to $u = t \mathbin{\dot\cup} s$. This walk has been defined and studied independently by [DD19] and by [AJT19], who bounded its spectral expansion.

Theorem 2.4 ([DD19; AJT19])**.** *Let X be a λ-two sided high dimensional expander. Then the second largest eigenvalue of $S_{k,\ell}(X)$ is upper bounded by $\lambda(S_{k,\ell}(X)) \leqslant (k+1)(\ell+1)\lambda$.*

We also define another random walk call the non-lazy up down walk.

Definition 2.5 (Non lazy up down walk)**.** Let X be a d-dimensional simplicial complex. Let $d_1 \leqslant d_2 \leqslant d$ such that $2d_1 \geqslant d_2 - 1$. The (d_1, d_2)-non lazy up down walk is the distribution $(s_1, s_2) \sim DU_n$ of $X(d_1)$ where s_1, s_2 are chosen by first choosing $t \in X(d_2)$ and then uniformly sampling $s_1, s_2 \in X(d_1)$ such that $s_1 \cup s_2 = t$.

The reason this is called non-lazy, is because we do not allow $s_1 \cup s_2 \subsetneq t$ as in the usual up-down walk defined in [Dik+18] for example. We note that $|s_1 \cap s_2| = 2d_1 - d_2 + 1$, i.e. the intersection size doesn't depend on the pair chosen. We can decompose this walk also as sampling $r = s_1 \cap s_2 \in X(2d_1 - d_2)$, then sampling $p_1, p_2 \in X_r(d_2 - d_1)$ according to the swap walk, and then outputting $s_1 = r \cup p_1, s_2 = r \cup p_2$.

The spherical building

Let $d \in \mathbb{N}$ and q be a prime power.

Definition 2.6. The spherical building (sometimes called the $SL_d(\mathbb{F}_q)$-spherical building), is the complex X whose vertices are all non-trivial linear subspaces of $\mathbb{F}_q^d$. It's higher dimensional faces are all flags $\left\{W_0 \subseteq W_1 \subseteq \cdots \subseteq W_m \,\middle|\, W_0, W_1, \ldots, W_m \subseteq \mathbb{F}_q^d\right\}$.

This complex is $(d-2)$-dimensional.

Claim 2.7 ([EK16]). Let X be a $SL_d(\mathbb{F}_q)$-spherical building. Then X is a $O(\frac{1}{\sqrt{q}})$-one sided high dimensional expander. Moreover, $X^{\leqslant k}$ is a $\max\{O(\frac{1}{\sqrt{q}}), \frac{1}{d-k}\}$-two sided high dimensional expander.

2.2 Agreement tests

Let $k < d$ and let X be a d-dimensional simplicial complex. Let Σ be some fixed alphabet and suppose we have an ensemble of functions $\mathcal{F} = \{f_r : r \to \Sigma \,|\, r \in X(k)\}$.

An two-query agreement test is a distribution $\mathcal{D}$ over pairs $r_1, r_2 \in X(k)$. The agreement of an ensemble is

$$\text{Agree}_{\mathcal{D}}(\mathcal{F}) = \mathop{\mathbb{P}}_{r_1, r_2 \sim D}[f_{r_1} = f_{r_2}]. \tag{2.1}$$

When we write $f_{r_1} = f_{r_2}$ we mean that $f_{r_1}(v) = f_{r_2}(v)$ for every $v \in r_1 \cap r_2$.

More generally, a q-ary agreement test is a distribution of $s_1, s_2, ..., s_q \in X(k)$, where the agreement of the ensemble

$$\text{Agree}_{\mathcal{D}}(\mathcal{F}) = \mathop{\mathbb{P}}_{r_1, r_2, \ldots r_q \sim D}\left[\forall i, j\ f_{r_i} = f_{r_j}\right]. \tag{2.2}$$

Let $\Delta_k(d)$ be the k-dimensional complete complex over d vertices. Let $\mathcal{D}$ be a q-ary agreement test on $\Delta_k(d)$ and assume that it is symmetric[4]. Let X be another d-dimensional simplicial complex. We define the *extension* $\mathcal{D}_X$ of $\mathcal{D}$ to an agreement test on X, as follows:

1. Sample $t \in X(d)$.
2. Query $s_1, s_2, ..., s_q \subseteq t$ according to $\Delta_k(d)$.

[4] i.e. for every permutation $\pi : [d] \to [d]$ it holds that $\mathbb{P}[s_1, s_2, ..., s_q] = \mathbb{P}[\pi(s_1), \pi(s_2), ..., \pi(s_q)]$.

We note that by the symmetry of D the second step doesn't depend on the way we identify the vertices of t with the vertices of $\Delta_k(d)$. Let us give two examples for such tests, that were considered in previous works.

Definition 2.8 (Two-query V-test)**.** Let $d = 2k - \sqrt{k+1} + 1$.

1. Sample some $t \in X(d)$.
2. Sample uniformly $s_1, s_2 \in X(k)$ such that $s_1, s_2 \subseteq t$, conditioned on $|s_1 \cap s_2| = \sqrt{k+1}$.

Definition 2.9 (Three-query Z-test)**.** Let $d = 3k - 2\sqrt{k+1} + 2)$.

1. Sample some $t \in X(d)$.
2. Sample three $s_1, s_2, s_3 \in X(k)$ such that $s_1, s_2, s_3 \subseteq t$, conditioned on $|s_1 \cap s_2|, |s_2 \cap s_3| = \sqrt{k+1}$ and $s_1 \cap s_3 = \emptyset$.

A sound distribution is a distribution that supports an agreement theorem.

Definition 2.10. Let X be a simplicial complex and let $\mathcal{D}$ be an agreement distribution on X. Let $\eta, \varepsilon_0 > 0$ be constants. We say that $\mathcal{D}$ is (η, ε_0)*-sound* if for every ensemble of functions $\mathcal{F}$ such that $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}) = \varepsilon \geqslant \varepsilon_0$, there exists a function $L : X(0) \to \Sigma$, such that

$$\mathop{\mathbb{P}}_{r_1, r_2, \ldots, r_q \sim \mathcal{D}} \left[\forall j \; L|_{r_j} \overset{1-\eta}{\approx} f_{r_j} \text{ and } \forall i, j \; f_{r_i} = f_{r_j} \right] \geqslant \frac{1}{2}\varepsilon.$$

Here $f \overset{1-\eta}{\approx} g$ means that f, g differ on at most a η-fraction of their coordinates, or stated differently $\mathrm{dist}(f, g) \leqslant \eta$.

Here are some examples of such distributions.

Example 2.11.
1. Dinur and Goldenberg showed that the V-test extended to $X = \Delta_k(n)$ is $(\sqrt{k}, k^{-c})$-sound for $d \geqslant k^3$ and $c > 0$ [DG08][5].
2. The 1%-agreement theorem by Impagliazzo, Kabanets and Wigderson, showed that the Z-test extended to $X = \Delta_k(n)$ together with the techniques used in [DG08, Theorem 5.1] show that the Z-test is $(k^{-0.2}, exp(-\Omega(k^{1/2}))$-sound.
3. Dinur and Livni-Navon, together with the techniques used in [DG08, Theorem 5.1] show that the Z-test is $(\lambda, exp(-\Omega(k)))$-sound for every constant $\lambda > 0$ [DL17].

Here $n \gg k$ in all cases. We remark that the exact distributions in the works in the example were not described the way we described them (the conditioning in the second step required the intersection to be at least a certain size, not exactly a certain size). However, when the complex is large enough the TV-distance between the distributions in all previous works, and the distributions we described above, is negligible in the regime of parameters we consider. We therefore ignore this small point.

We also stress that while previous works give very strong agreement results, that apply for $\varepsilon_0 = \exp(-\Omega(k))$, our technique currently yields results on for (η, ε_0) such that $\eta \exp(\mathrm{poly}(1/\varepsilon)) \ll 1$, so in particular, we will always think of $\varepsilon = \log(\mathrm{poly}(1/k))$ in the main theorem.

2.3 Covering maps

In this subsection we give a short introduction to covers and their connection to 1-cohomology. We stress that everything we state in this subsection is well known. For a more in depth discussion, see [Sur84].

Definition 2.12 (Covering map)**.** Let Y, X be simplicial complexes. We say that a map $\rho : Y(0) \to X(0)$ is a covering map if the following holds.

1. ρ is a surjective homomorphism.

[5] In fact, [DG08] considered a V-test with a smaller intersection size, but the same result hold for $\sqrt{k}$ too. See [IKW12] for a proof for the intersection size of $\sqrt{k}$.

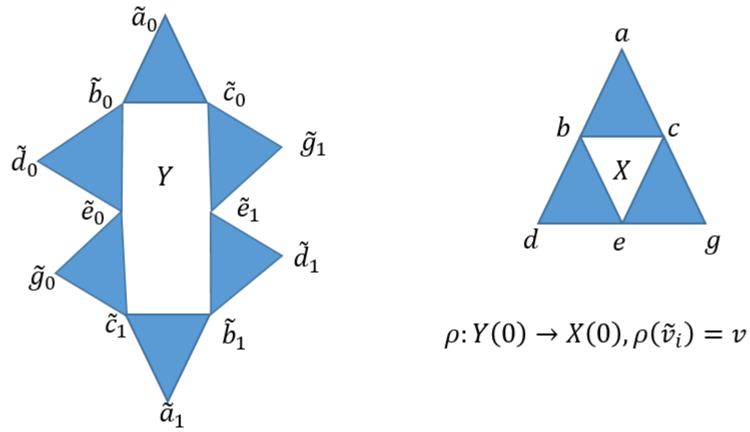

Figure 1: Non trivial connected cover

2. For every $v \in X(0)$, and $(v,i) \in \rho^{-1}(\{v\})$ it holds that $\rho|_{Y_{(v,i)}} : Y_{(v,i)}(0) \to X_v(0)$ is an isomorphism.

We often denote $\rho : Y \to X$. We say that ρ is an ℓ-cover if for every $v \in X(0)$ it holds that $\left|\rho^{-1}(\{v\})\right| = \ell$. If there exists such a covering map $\rho : Y \to X$ we say that Y covers X.

Let us see two examples for covers.

Example 2.13. Let X be any simplicial complex. Let Y be ℓ-disconnected copies of X. Then the projection of every to X, $\rho : Y \to X$, is a covering map. This cover a called the trivial cover.

Example 2.14. Figure 1 contains an example of a $\rho : Y \to X$ such that Y is connected.

Next we develop some basic properties of covers that are both necessary for our result, and give some intuition of the structure of covers.

The first property we show is that a cover $\rho : Y \to X$ induces a permutations π_{uv} for every directed edge $uv \in \tilde{X}(1)$ that "encode" the covers information as in the claim below.

Claim 2.15. Let X be a simplicial complex and let $\rho : Y \to X$ be a cover. Let $vu \in \vec{X}(1)$. Let $\rho^{-1}(v) = \{(v,1),(v,2),\ldots,(v,\ell)\}$ and $\rho^{-1}(u) = \{(u,1),(u,2),\ldots,(u,\ell')\}$. Then $\ell = \ell'$ and there exists $\pi_{uv} : [\ell] \to [\ell]$ such that $\pi_{uv}(i) = j$ if and only if $(v,i),(u,j) \in Y(1)$.

Using these permutations we can show that all covers of a connected simplicial complex are ℓ-covers.

Corollary 2.16. *If X is connected Then any cover of X is an ℓ-cover for some $\ell \in \mathbb{N} \cup \{\infty\}$.*

Proof of Claim 2.15. First we note that for every (v,i) there is a unique (u,j) such that $(v,i)(u,j) \in Y(1)$. This is because $\rho|_{\tilde{X}_{(v,i)}} : X_{(v,i)}(0) \to X_v(0)$ is an isomorphism so there is a single preimage of u in $X_{(v,i)}(0)$. Hence there are functions $\pi_{uv} : [\ell] \to [\ell']$ and $\pi_{uv} : [\ell'] \to [\ell]$ such that $\pi_{uv}(i) = j$ if and only if $(v,i)(u,j) \in Y(1)$. These functions invert each other from their definition, i.e. $\pi_{uv}(i) = j$ if and only if $(v,i)(u,j) \in Y(1)$ if and only if $\pi_{vu}(j) = i$ hence $\ell = \ell'$ and this is the required permutation. $\square$

Proof of Corollary 2.16. Let X be a connected complex and assume that there exists a cover $\rho : Y \to X$ such that Y is not an ℓ-cover for any ℓ. Let ℓ be the number of images of some arbitrary vertex $v \in X(0)$. Let $B = \left\{v' \in X(0) \middle| \left|\rho^{-1}(v')\right| = \ell\right\} \neq \emptyset$. If $X(0) \setminus B$ is not empty then the cut between B and $X(0) \setminus B$ has an edge crossing it. But the number of preimages for both sides of the edge is equal which lead to a contradiction. $\square$

Finally, let us show that in an ℓ-cover, every $s \in X$ has ℓ-inverse images.

Claim 2.17. Let $\rho : Y \to X$ be an ℓ-cover. Then for every non-empty $s \in X$,

$$|\rho^{-1}(s)| = \ell.$$

Proof of Claim 2.17. Let $s \in X$. Let $v \in s$ be any vertex inside s and we write $s = \{v\} \mathbin{\dot\cup} r$. By definition, there are ℓ-vertices $(v,1),(v,2),\dots,(v,\ell)$ such that $\rho((v,i)) = v$. By the local isomorphism property between X_v and every $Y_{(v,i)}$, there are faces $\tilde{r}_i \in Y_{(v,i)}$ such that $\rho(\tilde{r}_i) = r$. The faces $\tilde{s}_i = \tilde{r}_i \mathbin{\dot\cup} \{(v,i)\}$ are ℓ-inverse images of s. This shows that $|\rho^{-1}(s)| \geqslant \ell$. Let us see that every preimage of s it one of these $\tilde{s}_i$. Indeed, let $\hat{s} \in \rho^{-1}(s)$. Let $(v,i) \in \hat{s}$ be the preimage of v in $\hat{s}$. Then $\hat{r} = \hat{s} \setminus \{(v,i)\} \in Y_{(v,i)}$ maps to r. By the fact that $\rho|_{Y_{(v,i)}} : Y_{(v,i)}(0) \to X_v(0)$, this implies that $\hat{r} = \tilde{r}_i$. Thus $\hat{s} = \tilde{s}_i$. □

2.3.1 The induced function

Let $\rho : Y \to X$ be an ℓ-cover. Without loss of generality we identify

$$Y(0) = X(0) \times [\ell] = \{(v,i) \mid \ v \in X(0),\ i = 1,2,\dots,\ell\}$$

where $\rho((v,i)) = v$. We define *the induced function*

$$\psi_\rho : \vec{X}(1) \to Sym(\ell),\ \psi_\rho(vu) = \pi_{vu} \tag{2.3}$$

where $\pi_{uv}(i) = j$ are such that $\{(u,i)(v,j)\} \in Y(1)$ (as in Claim 2.15). The first thing we notice is that this function is asymmetric, i.e., $\psi(uv) = \psi(vu)^{-1}$ for every edge $uv \in \vec{X}(1)$. Let

$$C^1(X, Sym(\ell)) = \left\{ \psi : \vec{X}(1) \to Sym(\ell) \;\middle|\; f \text{ is asymmetric} \right\}$$

be the space of asymmetric functions. These functions are sometimes referred to as non-abelian cochains in the literature. One may suspect if there is a bijection between covers ρ and $\psi \in C^1(X, Sym(\ell))$, but there are many $\psi \in C^1(X, Sym(\ell))$ whose permutations don't give rise to a cover. It turns out that and asymmetric function $\psi \in C^1(X, Sym(\ell))$ corresponds to a cover if and only if for every triangle $uvw \in \vec{X}(2)$ it holds that

$$\psi_\rho(vw) \circ \psi_\rho(uv) = \psi_\rho(uw). \tag{2.4}$$

We denote by $Z^1(X, Sym(\ell)) \subseteq C^1(X, Sym(\ell))$

$$Z^1(X, Sym(\ell)) = \left\{ \psi \in C^1(X, Sym(\ell)) \;\middle|\; \forall uvw \in \vec{X}(2),\ \psi(vw) \circ \psi(uv) = \psi(uw) \right\}.$$

Claim 2.18. Let $\rho : Y \to X$ be an ℓ-cover. Then $\psi_\rho \in Z^1(X, Sym(\ell))$.

Proof of Claim 2.18. Let $\rho : Y \to X$ and fix $uvw \in X(2)$, let us show that $\psi_\rho(vw) \circ \psi_\rho(uv) = \psi_\rho(uw)$. Indeed, by definition $\psi_\rho(xy) = \pi_{xy}$ where $\pi_{xy}(i) = j$ if and only if $\{(x,i),(y,j)\} \in Y(1)$. Indeed, let $\pi_{uv}(i) = j$ and $\pi_{uw}(i) = k$ and we need to show that $\pi_{vw}(j) = k$, which is equivalent to showing that if $\{(v,i),(u,j)\},\{(v,i),(w,k)\} \in Y(1)$ then $(u,j),(w,k) \in Y(1)$. But indeed, as $\rho : Y_{(v,i)}(0) \to X_v(0)$ is an isomorphism, it holds that $\{(u,j),(w,k)\} \in Y_{(v,i)}(1)$ and in particular the edge is in $Y(1)$. □

If $\psi \in Z^1(X, Sym(\ell))$, then we can construct a cover $\rho = \rho_\psi : Y \to X$ such that $\psi_\rho = \psi$ as follows.

Definition 2.19. Let $\psi \in C^1(X, Sym(\ell))$ and denote by $\psi(uv) = \pi_{uv}$. The induced cover is the complex $Y = Y_\psi$ and mapping $\rho_\psi : Y \to X$ defined by

1. $Y(0) = X(0) \times [\ell]$ and $\rho_\psi((v,i)) = v$.
2. $Y(1) = \left\{ \{(v,i),(u,\pi_{vu}(i))\} \;\middle|\; v \in X(0), vu \in \vec{X}(1), i \in [\ell] \right\}$.

3. $\tilde{s} = \{(v_0, i_0), (v_1, i_1), \ldots, (v_m, i_m)\} \in Y(m)$ if and only if $s = \rho(\tilde{s}) = \{v_0, v_1, \ldots, v_m\} \in X(m)$ and for every $v_p, v_q \in s$ it holds that $\pi_{v_p,v_q}(i_p) = i_q$.

Claim 2.20. Let $\psi \in Z^1(X, Sym(\ell))$. Then $\rho_\psi : Y \to X$ is an ℓ-cover of X and $\psi_\rho = \psi$.

Proof of Claim 2.20. It is obvious that ρ is surjective and that it is a homomorphism from the definition. Let $v \in X(0)$ and $(v,i) \in Y(0)$. We show that $\rho|_{Y_{(v,i)}(0)} : Y_{(v,i)}(0) \to X_v(0)$ is an isomorphism. First let us make sure that this is a bijection. By definition, for every edge $\{v,u\} \in X(1)$ there is an edge $\{(v,i),(u,\pi_{vu}(i))\} \in Y(1)$ so $\rho|_{Y_{(v,i)}(0)}$ is surjective. A priori, there could have been more edges $\{(u,k),(v,i)\}$ if $i = \pi_{uv}(k)$, but because ψ is asymmetric, $i = \pi_{uv}(k)$ implies that $\pi_{vu}(i) = k$ so the restriction is also injective.

We continue by showing that this is an isomorphism. First note that this bijection is a homomorphism by definition (there are no faces $\tilde{s} \in Y(m)$ unless their projection $\rho(\tilde{s}) \in X(m)$, so the same holds for the link of (v,i)). Thus it remains to show that for every $s \in X_v(m)$ there is some $\tilde{s} \in X_{(v,i)}(m)$ such that $\rho(\tilde{s}) = s$. By definition of the link, $s \mathbin{\dot\cup} \{v\} = t \in X(m+1)$. We will show that $\tilde{t} = \{(v,i)\} \mathbin{\dot\cup} \{(u,j_u) \mid \ u \in s, j_u = \pi_{vu}(i)\}$ is in $Y(m+1)$, which implies that $\{(u,j_u) \mid \ u \in s, j_u = \pi_{vu}(i)\} = \tilde{s} \in Y_{(v,i)}(m)$ maps to s.

As $\rho(\tilde{t}) = t$ this amounts to showing that all possible edges $xy \in Y(1)$ for $x,y \in \tilde{t}$. One case is when, say, $x = (v,i)$ and $y = (u,j_u)$. In this case, $\{(v,i),(u,\pi_{vu}(i))\} \in Y(1)$ and $\pi_{vu}(i) = j_u$ by definition, hence the $xy \in Y(1)$. The other case is when $x = u_{j_u}$ and $y = w_{j_w}$ for some $u,w \in s$. in this case we note that $\pi_{vu}(i) = j_u, \pi_{vw}(i) = j_w$. By (2.4) applied for $vuw \in X(2)$, it holds that $j_w = \pi_{vw}(i) = \pi_{uw}(\pi_{vu}(i)) = \pi_{uw}(j_u)$. Hence $\{u_{j_u}, w_{j_w}\} \in Y(1)$ and the statement is proven.

The fact that $\psi_\rho = \psi$ follows directly from the definition of the edges in Y. $\square$

2.3.2 Connectivity of covers

A simply connected complex is a connected complex such that every cover looks like Example 2.13, i.e., a bunch of disconnected copies of the original complex.

Definition 2.21 (simply connected complex)**.** Let X be a connected simplicial complex. We say that X is *simply connected* if it is connected and if for every ℓ-cover $\rho : Y \to X$ there exists a partition of Y to ℓ disconnected components $Y = Y_1 \mathbin{\dot\cup} Y_2 \mathbin{\dot\cup} \ldots \mathbin{\dot\cup} Y_\ell$ such that $\rho|_{Y_i} : Y_i \to X$ is an isomorphism.

2.3.3 Further properties of covers

Claim 2.22. Let $\rho : Y \to X$ be an ℓ-cover. Let $A \subseteq X(0)$ and let X' be the induced subcomplex over the vertices of A. Let $Y' = \rho^{-1}(A)$. Then $\rho|_{Y'} : Y' \to X'$ is an ℓ-cover.

Proof of Claim 2.22. Let $\psi \in Z^1(X, Sym(\ell))$ such that $Y = X^\psi$ is the induced cover as in Definition 2.19. Then $\psi|_{X'(1)} \in Z^1(X', Sym(\ell))$ and $\rho|_{Y'}$ is the induced cover of this restriction. $\square$

Claim 2.23. Let X be a clique complex and let $\rho : Y \to X$ be a cover of X. Then Y is a clique complex.

Proof of Claim 2.23. Let $\tilde{s} = \{(v_0,i_0),(v_1,i_1),...,(v_k,i_k)\} \subseteq Y(0)$ be clique. Thus $\rho(\tilde{s}) = s \subseteq X(0)$ is also a clique, and thus $s \in X(k)$, or equivalently $s \setminus \{v_0\} \in X_{v_0}(k-1)$. The link of v_0 in X is isomorphic to the link of (v_0,i_0) via ρ. In particular this implies that $\tilde{s} \setminus \{(v_0,i_0)\} \in Y_{(v_0,i_0)}(k-1)$ which is equivalent to $\tilde{s} \in Y(k)$. $\square$

Claim 2.24. Let X be a λ-one or two sided high dimensional expander. Then any connected cover is a $\frac{\lambda}{1-\lambda}$-on or spectral expander respectively.

The proof of this claim relies on the trickling down theorem by Oppenheim [Opp18].

Theorem 2.25 ([Opp18])**.** *Let X be a* connected *simplicial complex and assume that for any vertex $v \in X(0)$ it holds that X_s is a λ-one or two sided high dimensional expander. Then the underlying graph of X is a $\frac{\lambda}{1-\lambda}$-one or two sided spectral expander (respectively), which implies that X is a $\frac{\lambda}{1-\lambda}$-one or two sided high dimensional expander (respectively).*

Proof of Claim 2.24. Let $\rho : Y \to X$ be a cover such that Y is connected. For every vertex $\tilde{v} \in Y$ such that $\rho(\tilde{v}) = v$, it holds that $Y_{\tilde{v}} \cong X_v$. Thus in particular, if X is a λ-one or two sided high dimensional expander, this implies that $Y_{\tilde{v}}$ is a λ-high dimensional expander. By Theorem 2.25, Y is also a $\frac{\lambda}{1-\lambda}$-one or two sided high dimensional expander. □

2.4 Cosystolic expansion and cover property testing

Recall that $Z^1(X, Sym(\ell)) \subseteq C^1(X, Sym(\ell))$ are all asymmetric functions such that $\psi(uw) = \psi(vw) \circ \psi(uv)$ for every triangle $uvw \in \vec{X}(2)$. For our result we will need simplicial complexes where this relation is locally testable. For this we define for every two function $\psi, \phi : \vec{X}(k) \to Sym(\ell)$ their distance

$$\mathrm{dist}(\psi, \phi) = \mathop{\mathbb{P}}_{s \in \vec{X}(k)} [\psi(s) \neq \phi(s)] . \tag{2.5}$$

We also denote the weight of the function $\mathrm{wt}(\psi) = \mathrm{dist}(\psi, Id)$ (where $Id : \vec{X}(k) \to Sym(\ell)$ assigns every face $s \in \vec{X}(k)$ the identity permutation).

For $\psi \in C^1(X, Sym(\ell))$ we define $\delta\psi : \vec{X}(2) \to Sym(\ell)$ by

$$\delta(\psi) = \psi(wu) \circ \psi(vw) \circ \psi(uv). \tag{2.6}$$

We are ready to define cosystolic expansion.

Definition 2.26. Let X be a d-dimensional simplicial complex for $d \geqslant 2$. Let $\beta > 0$. We say that X is a β-cosystolic expander if for every $\ell \in \mathbb{N}$, and every $\psi \in C^1(X, Sym(\ell))$ there exists some $\phi \in Z^1(X, Sym(\ell))$ such that

$$\beta \, \mathrm{dist}(\psi, \phi) \leqslant \mathrm{wt}(\delta\psi). \tag{2.7}$$

An explanation is in order. We think of the equations $EQ = \left\{ \psi(uw) = \psi(vw) \circ \psi(uv) \;\middle|\; uvw \in \vec{X}(2) \right\}$ as a set of tests and the weight $wt(\delta\psi)$ measures the probability that $\psi(uw) \neq \psi(vw) \circ \psi(uv)$, or in other words, the probability that ψ fails the test. When X is a β-cosystolic expander, this implies that if $\mathrm{wt}(\psi) = \varepsilon$ then there is a function $\phi \in Z^1(X, Sym(\ell))$ that is ε/β close to ψ.

We remark that in other works, many other coefficient groups were used instead of $Sym(\ell)$. For our result this definition is sufficient.

Finally, we say that X is a β-coboundary expander if X is simply connected and X is a β-cosystolic expander. Dinur and Meshulam already observed that cosystolic expansion (and coboundary expansion) is closely in fact equivalent testability of covers, which they call *cover stability* [DM22].

2.4.1 Near-cosystols from flag complexes

The following technical claim will be convenient later. The setup is as follows. Let X be a simplicial complex. Often is the case where for every edge $uv \in X(1)$ we have permutations $\pi_{u,uv}, \pi_{v,uv}$ and we are interested in constructing a co-chain $\psi(uv) = \pi_{u,uv}^{-1} \circ \pi_{v,uv}$. There is a natural consistency property that implies that ψ is (close to) a co-cycle: suppose that for every triangle $t \in X(2)$ and every $v \in t$ or sub-edge of $e \subset t$ there are permutations $\pi_{v,t}, \pi_{e,t}$. Then for a given $t \in X(2)$, if for every $v \in e \subset t$ it holds that

$$\pi_{v,t} = \pi_{e,t} \circ \pi_{v,e}, \tag{2.8}$$

then $\delta\psi(uvw) = Id$. See Figure 2 for an illustration.

To state this more easily, let us introduce the flag complex.

Definition 2.27 (Flag complex)**.** Let X be a simplicial complex. The flag complex is the complex GX whose vertices are the faces of X, and $\{s_0, s_1, ..., s_k\} \in GX(k)$ if $s_0 \subset s_1 \subset ... \subset s_k$ (for some ordering of the faces).

Claim 2.28. Let X be a two dimensional simplicial complex and let $\phi \in C^1(GX, Sym(\ell))$. Let $\psi = \psi_\phi : X(1) \to Sym(\ell)$ be given by $\psi(uv) = \phi(uv, v)\phi(u, uv)$. Then for any triangle $t \in X(2)$, if $\delta\phi(\{v \in e \subset t\}) = Id$ on all flags $\{v \in e \subset t\} \in GX(2)$ that contain t, then $\delta\psi(t) = Id$.

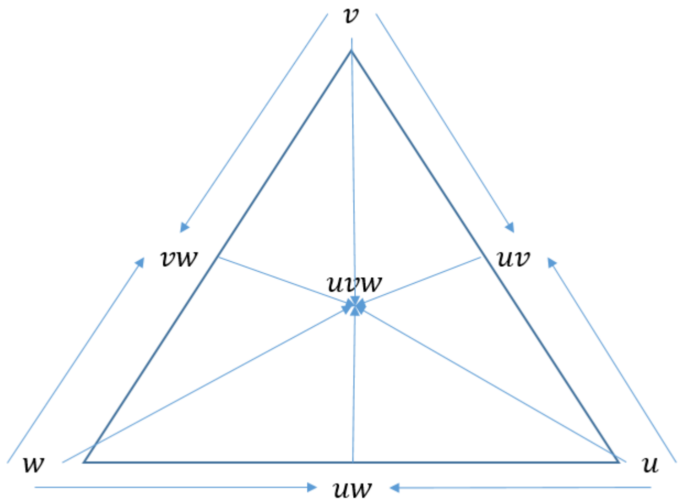

Figure 2: The following diagram should commute for $\delta\psi(uvw) = Id$

Corollary 2.29. *Let X be a two dimensional simplicial complex and let $\phi \in Z^1(GX, Sym(\ell))$. Let $\psi = \psi_\phi : X(1) \to Sym(\ell)$ be given by $\psi(uv) = \phi(uv, v)\phi(u, uv)$. Then $\psi \in Z^1(X, Sym(\ell))$.*

Proof of Claim 2.28. The proof is just a calculation. Let $uvw \in X(2)$ be a triangle. We need to show that $\psi(wv) \circ \psi(uw) = \psi(uv)$. Note that

$$\psi(uv) = \phi(uv, v)\phi(u, uv) = \phi(uv, v)\left[\phi(uvw, uv)\phi(uv, uvw)\right]\phi(u, uv) = \phi(uvw, v)\phi(u, uvw).$$

Where the last equality follows from $\phi \in Z^1(GX, Sym(\ell))$ so $\phi(uv, uvw)\phi(u, uv) = \phi(u, uvw)$ (we are using $\delta\phi(u, uv, uvw) = Id$) and $\phi(uv, v)\phi(uvw, uv) = \phi(uvw, v)$ (we are using $\delta\phi(uvw, uv, v) = Id$). Similarly, we have that

$$\psi(uw) = \phi(uvw, w)\phi(u, uvw); \ \psi(wv) = \phi(uvw, v)\phi(w, uvw). \tag{2.9}$$

Thus,

$$\begin{aligned}
\psi(uv) &= \phi(uvw, v)\phi(u, uvw) \\
&= \phi(uvw, v)\left[\phi(w, uvw)\phi(uvw, w)\right]\phi(u, uvw) \\
&= \left[\phi(uvw, v)\phi(w, uvw)\right] \cdot \left[\phi(uvw, w)\phi(u, uvw)\right] \\
&= \psi(wv)\psi(uw).
\end{aligned}$$

□

2.5 The faces complex

Definition 2.30. Let X be a d-dimensional simplicial complex. Let $d_1 \leqslant d$. We denote by $FX = F(X, d_1)$ the simplicial complex whose vertices are $FX(0) = X(d_1)$ and whose faces are all $\left\{\{s_0, s_1, ..., s_j\} \,\middle|\, s_0 \mathbin{\dot\cup} s_1 \mathbin{\dot\cup} \cdots \mathbin{\dot\cup} s_j \in X\right\}$.

It is easy to verify that this complex is $\left(\lfloor \frac{d+1}{d_1+1} \rfloor - 1\right)$-dimensional and that if X is a clique complex then so is FX. We omit d_1 from the notation when it is clear.

For a face $s \in X$ we write $FX_s = F(X_s)$. In case $s \in X(d_1)$ then FX_s is actually the link of s in FX, i.e. $F(X_s) = (FX)_s$, so this notation is not overloaded. We will be interested in FX_r for faces r not necessarily in $X(d_1)$. These are sub-complexes of FX that are not links per se.

Definition 2.31 (Well connected)**.** Let X be a d-dimensional simplicial complex. Let $d_1 \leqslant d+2$. We say that X is d_1*-well-connected* if for every $r \in X^{\leqslant d_1}$ it holds that $F(X_r, d_1) = FX_r$ is connected. Moreover, if $r \in X(0)$ then we require that FX_r is *simply connected*. When d_1 is clear from context we omit it and say that X is *well connected*.

2.6 Sampling in HDXs

Theorem 2.32 ([DH23])**.** *Let $k \leqslant d_1 \leqslant d$. Let X be one of the following:*

1. *A complete complex over d vertices.*
2. *A 2^{-7d}-two sided high dimensional expander.*
3. *A 2^{-7d}-one sided high dimensional expander that is a d-skeleton of a d'-partite complex for $d' \geqslant d^2$.*

Let $\delta > 0$ and let $f : X(k) \to [0,1]$. Let

$$B(f) = \left\{ t \in X(d_1) \,\middle|\, \left| \mathop{\mathbb{E}}_{r \in X(k), r \subseteq t} [f(r)] - \mathop{\mathbb{E}}_{r \in X(k)} [f] \right| > \delta \right\}.$$

Then

$$\mathop{\mathbb{P}}_{t \in X(d_1)} [B(f)] \leqslant \exp(-\operatorname{poly}(\delta)\frac{d_1}{k}).$$

In particular, if $A \subseteq X(k)$ and

$$B(A) = \left\{ t \in X(d_1) \,\middle|\, \left| \mathop{\mathbb{P}}_{r \in X(k)} [A \mid r \subseteq t] - \mathop{\mathbb{P}}_{r \in X(k)} [A] \right| > \delta \right\},$$

then

$$\mathop{\mathbb{P}}_{t \in X(d_1)} [B(A)] \leqslant \exp(-\operatorname{poly}(\delta)\frac{d_1}{k}).$$

The "in particular" part follows from the case for $f = \mathbf{1}_A$ the indicator of A.

Claim 2.33. Let $k \leqslant d_1 \leqslant d$ and $q \in \mathbb{N}$. Let $d_1 \geqslant qk$ and let $\delta > 0$ be some constant. Let X be a d-dimensional simplicial complex that is either a 2^{-7d}-two sided spectral expander, or a 2^{-7d}-one sided spectral expander that is a skeleton of a D partite complex for $D \geqslant d^2$. Let $\mathcal{D}$ be an agreement distribution over $\Delta_{qk+q}(k)$. and let $E \subseteq \operatorname{supp} \mathcal{D}_X$. Then

$$\mathop{\mathbb{P}}_{s \in X(d_1)} \left[\left| \mathop{\mathbb{P}}_{\{r_i\} \sim \mathcal{D}} [E \mid \{r_i\} \subseteq s] - \mathop{\mathbb{P}}_{\{r_i\} \sim D} [E] \right| > \delta \right] \leqslant \exp(-\operatorname{poly}(\delta)\frac{d_1}{k}).$$

Proof of Claim 2.33. Let $f : X(qk+q-1) \to [0,1]$ be $f(a) = \mathbb{P}_{\{r_i\} \subseteq a} [E]$. Then $\mathbb{E}[f] = \mathbb{P}[E]$ and $\mathbb{E}_{a \subseteq s} [f(a)] = \mathbb{P}_{\{r_i\} \sim \mathcal{D}} [E \mid \{r_i\} \subseteq s]$. Theorem 2.32 gives us the claim. ◻

2.7 Suitable complexes

Definition 2.34 (Suitable complex)**.** Let d, k be integers, and let $\alpha > 0$. Let X be a d-dimensional simplicial complex. We say that X is (d, k, α)*-suitable* if it has the following properties:

1. There exists some integer d_1 with the following properties:
   (a) $k^3 \leqslant d_1 \leqslant d \exp(-\alpha \frac{d_1}{k})$.
   (b) The faces complex $F(X, d_1)$ is a $\exp(-\alpha \frac{d_1}{k})$-cosystolic expander.
   (c) X is d_1-well connected as in Definition 2.31.
2. One of the following expansion properties hold:
   (a) X is a 2^{-7d}-two sided local spectral expander.

(b) X is a 2^{-7d}-one sided local spectral expander. In addition, X is a skeleton of a partite simplicial complex of dimension $\geqslant d^2$.

3. X is a clique complex.

These properties are abstract and a priori it is not clear whether there is a complex that satisfies them, especially the property of the face complex being a coboundary expander. However, the following two families of complexes satisfy all of these properties.

Example 2.35.

1. The $SL_d(\mathbb{F}_q)$-spherical building (see Definition 2.6) for sufficiently large prime power q and dimension d.
2. [LSV05a] complexes, that are connected simplicial complexes whose links look like the spherical buildings in the first item.

Theorem 2.36 ([DD23b])**.** *Let d, k, α be such that $\alpha \in (0,1)$ and $d \geqslant \exp(\mathrm{poly}(k,\alpha))$. Then there exists a constant $q_0 = q_0(k,\alpha)$ such that for every prime power $q > q_0$:*

1. *The d-skeleton of the $SL_{d^2}(\mathbb{F}_q)$-spherical building is (d,k,α)-suitable.*
2. *d-dimensional skeletons of the D-dimensional [LSV05a] complexes, such that the link of every vertex is isomorphic to $SL_{D-1}(\mathbb{F}_q)$-spherical building are (d,k,α)-suitable, provided that $D \geqslant \exp(O(d))$.*
3. *d-dimensional skeletons of D-partite [LSV05a] complexes, such that the link of every vertex is isomorphic to the $SL_{D-1}(\mathbb{F}_q)$-spherical building is (d,k,α)-suitable, provided that $D \geqslant d^2$.*

This paper is focused on the 1%-agreement theorem; the proof of the above theorem is in a companion paper [DD23b].

3 Agreement Theorem

In this section we describe and prove our agreement theorem. Recall that we defined *suitable complexes* in Definition 2.34.

Theorem 3.1 (Main)**.** *For every $\varepsilon_0 > 0$, $p > 0$, and $k \geqslant \exp(\mathrm{poly}(1/\varepsilon_0))$ there exists $\alpha = \mathrm{poly}(\varepsilon)$ and $d \in \mathbb{N}$ such that the following holds. Let $\mathcal{D}$ be an agreement distribution on $\Delta_k(pk)$ such that its extension to $\Delta_k(m)$ is (η,ε)-sound for every $m \geqslant k^3$. Let X be a (d,k,α)-suitable complex and let $\mathcal{F} = \{f_r : r \to \Sigma \mid \ r \in X(k)\}$ be an ensemble of functions for some finite alphabet Σ. If*

$$Agree_{\mathcal{D}_X}(\mathcal{F}) = \varepsilon \geqslant \varepsilon_0$$

then there exists a simplicial $\mathrm{poly}(1/\varepsilon)$-cover $\rho : Y \to X$ and a global function $G : Y(0) \to \Sigma$ such that

$$\mathop{\mathbb{P}}_{r \in Y(k)}\left[f_{\rho(r)} \circ \rho \overset{1-\gamma}{\approx} G|_r\right] = poly(\varepsilon). \tag{3.1}$$

where $\gamma = \exp(\mathrm{poly}(1/\varepsilon))\eta$.

This theorem gives us Theorem 1.3 and Theorem 1.5 as corollaries. We derive these corollaries from Theorem 3.1 in the end of this section.

Corollary 3.2 (Theorem 1.3, formal)**.** *For every $\varepsilon_0 > 0$ and $k \geqslant \exp(\mathrm{poly}(1/\varepsilon_0))$ there exists $d_0 \geqslant k$ and integer q_0 such that the following holds. Let X be a d-dimensional spherical building with a sufficiently large field size $q \geqslant q_0$ and dimension $d \geqslant d_0$. Let $\mathcal{D}$ be either the V-test or the Z-test for sets of dimension k. Then for every $\mathcal{F} = \{f_r : r \to \Sigma \mid \ r \in X(k)\}$ such that $Agree_{\mathcal{D}_X}(\mathcal{F}) = \varepsilon \geqslant \varepsilon_0$, there exists $\ell = \mathrm{poly}(1/\varepsilon)$ and a list of functions $G_1, G_2, \dots, G_\ell : Y(0) \to \Sigma$ such that for every G_i,*

$$\mathop{\mathbb{P}}_{r \in X(k)}\left[f_r \overset{1-\gamma}{\approx} G_i|_r\right] = poly(\varepsilon). \tag{3.2}$$

where $\gamma = \exp(\mathrm{poly}(1/\varepsilon))\,\mathrm{poly}(1/k)$.

Moreover, the list is exhaustive in the following sense

$$\mathop{\mathbb{P}}_{\{r_i\}\sim\mathcal{D}}\left[\{f_{r_i}\} \text{ passes the agreement test but for every } G_m \text{ there is at least one } f_{r_i} \text{ s.t. } f_{r_i} \overset{1-\gamma}{\not\approx} G_m|_{r_i}\right] \leqslant \varepsilon^2. \tag{3.3}$$

Corollary 3.3 (Theorem 1.5, formal)**.** *For every* $\varepsilon_0 > 0$ *and* $k \geqslant \exp(\mathrm{poly}(1/\varepsilon_0))$*, there exists* $d_0 \geqslant k$ *and integer* q_0 *such that the following holds. Let* X *be a simplicial complex whose links are (isomorphic to) the* $SL_d(\mathbb{F}_q)$*-spherical building with a sufficiently large field size* $q \geqslant q_0$ *and dimension* $d \geqslant d_0$*. Let* $\mathcal{D}$ *be either the* V*-test or the* Z*-test for sets of dimension* k*. Then for every* $\mathcal{F} = \{f_r : r \to \Sigma \mid \ r \in X(k)\}$ *such that* $Agree_{\mathcal{D}_X}(\mathcal{F}) = \varepsilon \geqslant \varepsilon_0$*, there exists* $\ell = \mathrm{poly}(1/\varepsilon)$*, and an* ℓ*-cover* $\rho : Y \to X$ *and a global function* $G : Y(0) \to \Sigma$ *such that*

$$\mathop{\mathbb{P}}_{r\in Y(k)}\left[f_{\rho(r)} \circ \rho \overset{1-\gamma}{\approx} G|_r\right] = poly(\varepsilon).$$

where $\gamma = \exp(\mathrm{poly}(1/\varepsilon))\,\mathrm{poly}(1/k)$.

Note that this corollary applies for the [LSV05a] complexes in particular (but also slightly more generally).

Proof of the agreement theorem

In this sub-section we prove Theorem 3.1. The proof will have an overview style, i.e. we will give a high-level proof, relying on a few claims and lemmas that will only be stated at this point. Then we will prove all outstanding claims in the following sections, formally completing the proof. For the rest of the section, we fix X to be a suitable complex and some ensemble of functions $\mathcal{F}$ such that $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}) = \varepsilon$.

3.1 From 1% agreement to 99.9% list-agreement

Our starting point is the following list-decoding lemma which we prove in Section 4.

Lemma 3.4. *Let* $X, \mathcal{D}$ *and* $\mathcal{F}$ *be as in Theorem 3.1. There exists* $\ell = \mathrm{poly}(1/\varepsilon)$ *such that for every* $s \in X(d_1) \cup X(2d_1+1) \cup X(3d_1+2)$ *there exists a tuple of functions*

$$\overline{L}_s = (L_s^1, \ldots, L_s^\ell), \qquad L_s^i : s \to \Sigma$$

such that the following holds.

1. *All but* $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ *fraction of faces in* $X(d)$ *are good, where a face is good if for all* $i = 1, \ldots, \ell$*,*

$$\mathop{\mathbb{P}}_{r\in X(k), r\subset s}\left[L_s^i|_r \overset{1-\gamma}{\approx} f_r\right] \geqslant \mathrm{poly}(\varepsilon)$$

   for $\gamma = \eta \exp(\mathrm{poly}(1/\varepsilon))$.

2. *For every pair* $s \subset t$*,* $s, t \in X(d_1) \cup X(2d_1+1) \cup X(3d_1+2)$ *there is a permutation* $\pi_{s,t} : [\ell] \to [\ell]$ *such that*

$$\mathop{\mathbb{P}}_{s,t}\left[\forall i = 1, 2, ..., \ell, j = \pi_{s,t}(i)\ L_s^i \overset{1-\gamma}{\approx} L_t^j|_s\right] \geqslant 1 - exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k}).$$

3. *For a random triple* $s, u, t \in X(d_1) \cup X(2d_1+1) \cup X(3d_1+2)$ *such that* $s \subset t \subset u$ *it holds that*

$$\mathop{\mathbb{P}}_{s\subset t\subset u}[\pi_{s,u} = \pi_{s,t} \circ \pi_{t,u}] \geqslant 1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k}).$$

3.2 The flags complex FX and an almost-cocycle

Recall Definition 2.30. The d_1-face complex FX is the complex whose vertices are $X(d_1)$ and whose faces are all $\{s_1, s_2, ..., s_m\}$ such that $s_1 \uplus s_2 \uplus \cdots \uplus s_m \in X$. Using Lemma 3.4 we construct an almost-cycle $\psi \in C^1(FX, Sym(\ell))$ as follows.

For each edge (s_1, s_2) of FX, $\psi((s_1, s_2)) = \pi_{s_2,t}^{-1} \circ \pi_{s_1,t}$ where $t = s_1 \uplus s_2$.

Lemma 3.5.

$$\mathrm{wt}(\delta(\psi)) = \mathop{\mathbb{P}}_{(s_1,s_2,s_3)\in FX(2)} [\psi(s_1, s_2) \circ \psi(s_2, s_3) \neq \psi(s_1, s_3)] = exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k}). \tag{3.4}$$

Proof of Lemma 3.5. We wish to use Claim 2.28 for the flag complex of the face complex. Let Z be the flag complex of FX. We define $\phi : Z(1) \to Sym(\ell)$ as follows. Let $\{\{s_1\}, \{s_1, s_2\}, \{s_1, s_2, s_3\}\}$ be a triangle in the flag complex. Let $u = s_1 \uplus s_2$ and let $t = s_1 \uplus s_2 \uplus s_3$. Then

$$\phi(\{s_1\}, \{s_1, s_2\}) = \pi_{s_1,u}, \ \phi(\{s_1\}, \{s_1, s_2, s_3\}) = \pi_{s_1,t}, \ \phi(\{s_1, s_2\}, \{s_1, s_2, s_3\}) = \pi_{u,t}.$$

Where the permutations are the ones given in Lemma 3.4. By the third item of Lemma 3.4, it holds that with probability $\mathrm{wt}(\delta\phi) = \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ and by Claim 2.28 the same holds for $\delta(\psi)$. □

Using Lemma 3.5 and $\beta = \exp(-\alpha\frac{d_1}{k})$-cosystolic expansion of FX we have that

Corollary 3.6. *There exists a cocycle $\psi' \in Z^1(X, Sym(\ell))$ such that* $\mathrm{dist}(\psi', \psi) = \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k} + \alpha\frac{d_1}{k})$, *namely,*

$$\mathop{\mathbb{P}}_{s_1,s_2}\left[\psi'(s_1, s_2) = \pi_{s_2,t}^{-1} \circ \pi_{s_1,t}\right] \geqslant 1 - \exp(-\operatorname{poly}(\varepsilon)d_1/k). \quad \square$$

Here is where we require that β is large enough relative $\exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$. By Definition 2.19, ψ' gives rise to an ℓ-cover $\rho_{\psi'} : \widetilde{FX} \to FX$.

3.3 Construction of a cover $Y \twoheadrightarrow X$

Recall that our goal is to construct a cover $Y \twoheadrightarrow X$ (together with a global function on the vertices of Y such that (3.1) holds). So far we managed to construct a cover $\nu = \nu_{\psi'} : \widetilde{FX} \to FX$. It seems natural to expect that every such cover $\widetilde{FX}$ is (isomorphic to) a face complex FY of some complex Y. When X is a well-connected clique complex, this is indeed the case.

Lemma 3.7. *Let $\nu : \widetilde{FX} \to FX$ be an ℓ-covering map. Then there exists an ℓ-covering map $\rho : Y \to X$ such that FY and $\widetilde{FX}$ are isomorphic.*

We denote by $\iota : FY \xrightarrow{\sim} \widetilde{FX}$ the isomorphism. We prove this lemma in Section 5.

3.4 The global function G

Given Lemma 3.7, we can define an ensemble of functions for FY to be $h_{\tilde{s}} : \tilde{s} \to \Sigma$, $h_{\tilde{s}} = L_s^j \circ \rho$ such that $\iota(\tilde{s}) = (s, j)$ is the identification promised in Lemma 3.7. We define a majority (plurality) function $G : Y(0) \to \Sigma$ by $G(v) = plurality_{s \ni v} h_s(v)$. We prove that

Claim 3.8. There exists a universal constant $c > 0$ such that $\mathbb{P}_{\tilde{s}\in Y(d_1)}\left[h_{\tilde{s}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{s}}\right] \geqslant 1 - \exp(-\operatorname{poly}(\varepsilon)\left(\frac{d_1}{k}\right)^c)$.

This claim is proven in Section 6. Combining all the pieces we can finally prove the theorem. Consider the probability $\mathbb{P}_{\tilde{r}\in Y(k)}\left[f_r \circ \rho \overset{1-O(\eta)}{\not\approx} G|_{\tilde{r}}\right]$. Sampling $\tilde{r} \in Y(k)$ is the marginal of sampling a pair $(\tilde{r} \subseteq \tilde{s})$ in Y such that $\tilde{s} \in Y(d_1)$. Denote by $\iota(\tilde{s}) = (s, j)$. If $f_r \overset{1-O(\gamma)}{\approx} L_s^j|_r$, and $h_{\tilde{s}}|_{\tilde{r}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{r}}$ then

$$f_r \circ \rho \overset{1-O(\gamma)}{\approx} L_s^j|_r \circ \rho = h_{\tilde{s}}|_{\tilde{r}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{r}}.$$

Thus we define the "bad events"

1. E_1 the event that $f_r \overset{1-\gamma}{\not\approx} L_s^j|_r$.
2. E_2 the event that $h_{\tilde{s}}|_{\tilde{r}} \overset{1-O(\gamma)}{\not\approx} G|_{\tilde{r}}$.

and we have that

$$\mathop{\mathbb{P}}_{\tilde{r}\in Y(k)}\left[f_r \circ \rho \overset{1-O(\gamma)}{\not\approx} G|_{\tilde{r}}\right] \leqslant \mathbb{P}\left[E_1\right] + \mathbb{P}\left[E_2\right].$$

By Lemma 3.4 $\mathbb{P}\left[E_1\right] \leqslant 1 - \mathrm{poly}(\varepsilon)$. Thus to show the theorem it is enough to bound $\mathbb{P}\left[E_2\right] \leqslant \exp(-\mathrm{poly}(\gamma)k) + \exp(-\mathrm{poly}(\varepsilon)\left(\frac{d_1}{k}\right)^c)$ for the $c > 0$ in Claim 3.8 (k and d_1 are chosen large enough such that this is much smaller than the upper bound on $\mathbb{P}\left[E_1\right]$). Indeed,

$$\mathbb{P}\left[E_2\right] \leqslant \mathop{\mathbb{P}}_{\tilde{s},\tilde{r}}\left[h_{\tilde{s}} \overset{1-O(\gamma)}{\not\approx} G|_{\tilde{s}}\right] + \mathop{\mathbb{P}}_{\tilde{s},\tilde{r}}\left[h_{\tilde{s}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{s}} \text{ and } h_{\tilde{s}}|_{\tilde{r}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{r}}\right].$$

The first item in this sum is bounded by $\exp(-\mathrm{poly}(\varepsilon)\left(\frac{d_1}{k}\right)^c)$ by Claim 3.8. The second is bounded by $\exp(-\mathrm{poly}(\gamma)k)$ by a standard Chernoff argument.

Thus

$$\mathop{\mathbb{P}}_{\tilde{r}\in Y(k)}\left[f_r \circ \rho \overset{1-O(\eta)}{\approx} G|_{\tilde{r}}\right] \geqslant \mathrm{poly}(\varepsilon) - \exp(-\mathrm{poly}(\varepsilon)\left(\frac{d_1}{k}\right)^c) - \exp(-\mathrm{poly}(\gamma)k) = \mathrm{poly}(\varepsilon)$$

and Theorem 3.1 follows.

3.5 Deriving Corollary 3.2 and Corollary 3.3

Proof of Corollary 3.2. Let us first show that there exists *one* global function $G : X(0) \to \Sigma$ such that (3.2) holds for G.

Fix ε_0 and let d, k, α be as in Theorem 3.1 (for agreement distributions for either 2 or 3 sets). By Theorem 2.36 there is a q_0 such that for $q \geqslant q_0$, X is suitable. As seen in Example 2.11, both the V-test and the Z-test are (ε_1, η)-sound for $\varepsilon_1, \eta = \mathrm{poly}(1/k)$. We take k to be large enough so that $\varepsilon_0 \geqslant \varepsilon_1$ and so that $\gamma = \exp(1/\mathrm{poly}(\varepsilon_0)\eta < 1$. By Theorem 3.1, there exists a cover $\rho : Y \to X$ and a global function such that $G' : Y(0) \to \Sigma$ agrees with the ensemble as in (3.1).

It is well known that the spherical building is simply connected (see e.g. [DD23a] for a proof of this fact). By simple connectivity, every cover is trivial, so a cover is in fact composed of many disconnected components, each isomorphic to X. Hence we can think of the global function $G' : Y(0) \to \Sigma$ as a list of functions $G_1, G_2, \ldots, G_\ell : X(0) \to \Sigma$, each is a restriction of G' to one of the components isomorphic to X. From this we get that there is at least function $G = G_i$ such that (3.1) holds conditioned on sampling $\tilde{r}$ in G_i's component.

Turning to (3.3), by [DG08, Theorem 5.1], if

$$\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}) = \varepsilon \qquad \Rightarrow \qquad \exists G : X(0) \to \Sigma \mathop{\mathbb{P}}_{r\in X(k)}\left[f_r \overset{1-\gamma}{\approx} G|_r\right] \geqslant \mathrm{poly}(\varepsilon)$$

holds, then there is an (inefficient) algorithm that outputs a list of functions $G_1, G_2, \ldots, G_\ell$ for $\ell = \mathrm{poly}(1/\varepsilon)$ such that (3.3) holds. □

Proof of Corollary 3.3. Fix ε_0 and let d, k, α be as in Theorem 3.1 (for agreement distributions for either 2 or 3 sets). By Theorem 2.36 there is a q_0 such that for $q \geqslant q_0$, X is suitable. As seen in Example 2.11, both the V-test and the Z-test are (ε_1, η)-sound for $\varepsilon_1, \eta = \mathrm{poly}(1/k)$. We take k to be large enough so that $\varepsilon_0 \geqslant \varepsilon_1$ and so that $\gamma = \exp(1/\mathrm{poly}(\varepsilon_0))\eta < 1$. The corollary follows directly from Theorem 3.1. □

4 From one percent agreement to list agreement

In this section we prove Lemma 3.4.

Lemma (Restatement of Lemma 3.4)**.** *Let $X, \mathcal{D}$ and $\mathcal{F}$ be as in Theorem 3.1. There exists $\ell = \mathrm{poly}(1/\varepsilon)$ such that for every $s \in X(d_1) \cup X(2d_1+1) \cup X(3d_1+2)$ there exists a tuple of functions*

$$\overline{L}_s = (L_s^1, \dots, L_s^\ell), \qquad L_s^i : s \to \Sigma$$

such that the following holds.

1. *All but $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ fraction of faces in $X(d)$ are good, where a face is good if for all $i = 1, \dots, \ell$,*
$$\mathop{\mathbb{P}}_{r \in X(k), r \subset s}\left[L_s^i|_r \overset{1-\gamma}{\approx} f_r\right] \geqslant \mathrm{poly}(\varepsilon)$$
for $\gamma = \eta \exp(\mathrm{poly}(1/\varepsilon))$.

2. *For every pair $s \subset t$, $s, t \in X(d_1) \cup X(2d_1+1) \cup X(3d_1+2)$ there is a permutation $\pi_{s,t} : [\ell] \to [\ell]$ such that*
$$\mathop{\mathbb{P}}_{s,t}\left[\forall i = 1, 2, ..., \ell, j = \pi_{s,t}(i)\ L_s^i \overset{1-\gamma}{\approx} L_t^j|_s\right] \geqslant 1 - exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k}).$$

3. *For a random triple $s, u, t \in X(d_1) \cup X(2d_1+1) \cup X(3d_1+2)$ such that $s \subset t \subset u$ it holds that*
$$\mathop{\mathbb{P}}_{s \subset t \subset u}[\pi_{s,u} = \pi_{s,t} \circ \pi_{t,u}] \geqslant 1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k}).$$

Fix X to be a suitable simplicial complex and fix $\mathcal{F} = \{f_r : r \to \Sigma \mid r \in X(k)\}$ to be an ensemble of functions on $X(k)$. The distribution $\mathcal{D}$ is (ε_0, η)-sound. We also assume that $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F})^M = \varepsilon^M \geqslant \varepsilon_0$ for some constant integer $M > 0$ (which we do not explicitly calculate).

For a function $g : X(0) \to \Sigma$ We denote by $\mathrm{supp}_\delta(g) = \left\{r \in X(k) \,\middle|\, f_r \overset{1-\delta}{\approx} g\right\}$ and call this set the δ-support of g. We also denote by $A_\delta(g) = \{\{r_1, r_2\} \mid r_1, r_2 \in \mathrm{supp}_\delta(g) \text{ and } f_{r_1} = f_{r_2}\}$ the "agreeing edge set of g" and by $\alpha_\delta(g) = \mathbb{P}_{\{r_1,r_2\}\sim D}[A_\delta(g)]$. Here D is the agreement distribution. When the ensemble of functions is not clear from context we denote $\mathrm{supp}_\delta^{\mathcal{F}}(g), A_\delta^{\mathcal{F}}(g)$ etc.

For a face $t \in X$, we denote by $\mathcal{F}_t = \{f_r \in \mathcal{F} \mid r \subseteq t\}$. The set of functions whose agreeing set of edges is large is denoted by $\overline{L}_{\tau,\delta}(t) = \left\{g : t \to \Sigma \,\middle|\, \alpha_\delta^{\mathcal{F}_t}(g) \geqslant \tau\right\}$. In a lot of the following section we will consider partial functions $g : t \to \Sigma$ (instead of all $X(0)$). In this case when we write $\mathrm{supp}_\delta(g), \alpha_\delta(g)$ or $A_\delta(g)$ we will mean the support or agreement of g on its domain with respect to $\mathcal{F}_t$.

4.1 Overview - constructing the permutations in Lemma 3.4

The idea is to use the the agreement theorem at hand to create lists of functions $\overline{L}(s) = \{L_s^1, L_s^2, \dots, L_s^\ell\} \subseteq \overline{L}_{\tau,\delta}(s)$ for every face s. Then one needs to show that there is a matching between most lists $\overline{L}(s), \overline{L}(t)$ such that $s \subseteq t$. The matching we construct is so that $\pi_{s,t}(i) = j$ if and only if $L_t^i|_s \overset{1-\gamma}{\approx} L_t^j$ for some small $\gamma > 0$.

There are three degrees of freedom here, namely the choice of τ, δ and γ, and even after choosing them, it is not clear a priori which functions should appear in the list. We need the size of the lists to be $\mathrm{poly}(\frac{1}{\varepsilon})$, so we cannot take all functions in $\overline{L}^0_{\tau,\delta}(s)$. This gives rise to many problematic issues. The first two things we need to consider are:

1. If $f_r \overset{1-\gamma_1}{\approx} g|_r$ and $g|_r \overset{1-\gamma_2}{\approx} g'|_r$ then $f_r \overset{1-\gamma_1-\gamma_2}{\approx} g'|_r$. Hence, if two functions $g, g' : t \to \Sigma$ are close (in Hamming distance), then their support will have a lot of intersection. In particular, if $L_t^i \in \overline{L}(t)$ we should exclude from the list a small Hamming ball around L_t^i. Otherwise, if $L_t^i|_s$ is γ-close to both $L_s^j, L_s^{j'}$ we will have a difficulty of choosing whether $\pi_{s,t}(i) = j$ or $\pi_{s,t}(i) = j'$ in a way that will ensure that $\pi_{s,u} = \pi_{s,t} \circ \pi_{t,u}$ for most $s \subseteq t \subseteq u$.

2. We want the lists to be exhaustive. That is, if $g : s \to \Sigma$ has that $f_r \approx g|_r$ for $\geqslant \tau = \text{poly}(\varepsilon)$ fraction of the $r \in X(k)$ inside s, then either $g \in \overline{L}(s)$ or $g' \in L$ for some $g' \approx g$. Otherwise, this could also lead to a difference in $\overline{L}(s), \overline{L}(t)$'s sizes, since for instance, it may be that $g \in \overline{L}(t)$ but its restriction $g|_s$ is not close to any element in the list of $\overline{L}(s)$.

Towards dealing with these issues we define *density* and *separation* of a set.

Definition 4.1. Let $\overline{L}_1, \overline{L}_2 \subseteq \{g : s \to \Sigma\}$. Let $\gamma > 0$.

1. We say that $\overline{L}_1$ is γ-separated if for every two $g_1, g_2 \in \overline{L}_1$ it holds that $\text{dist}(g_1, g_2) > \gamma$.
2. We say that $\overline{L}_1$ is γ-dense in $\overline{L}_2$ if for every $h \in \overline{L}_2$ there exists $g \in \overline{L}_1$ such that $\text{dist}(g, h) \leqslant \gamma$ (or as we denoted $g \overset{1-\gamma}{\approx} h$).

Thus, for every s we will need to find a list $\overline{L}(s)$ that is both separated (to overcome the first issue) and dense inside $\overline{L}_{\tau,\delta}(s)$ (to overcome the second issue). In fact, a method by [Din+18a] allows us to find lists that are γ-dense but 10γ-separated (for a smartly chosen γ). This is the content of Claim 4.8. We will need this property in order to match between lists. If $\overline{L}(s)$ is γ-dense in $\overline{L}_{\tau,\delta}(s)$ and $\overline{L}(t)|_s = \left\{ L_t^i|_s \;\middle|\; L_t^i \in \overline{L}(t) \right\}$ is contained in $\overline{L}_{\tau,\delta}(s)$ then for every $L_t^i \in \overline{L}(t)$ there exists some $L_s^j \in \overline{L}(s)$ such that $L_s^j \overset{1-\gamma}{\approx} L_t^i|_s$.[6]

The 10γ-separation will promise that there is only a single such L_s^j for every L_t^i, because if $L_s^j, L_s^{j'}$ are both γ-close to L_t^i then they are 2γ-close to one another, which is a contradiction to 10γ-separation. If we have such lists this implies that $\pi_{s,t}(i) = j \Leftrightarrow L_t^i|_s \overset{1-\gamma}{\approx} L_s^j$ is a well defined function.

By making sure that $\overline{L}(t)$ is also 10γ-separated, we get that this function $\pi_{s,t}$ is injective for most $s \subseteq t$. If $L_t^i, L_t^{i'}$ are 10γ-separated, then for most $s \subseteq t$, it will hold that $L_t^i|_s, L_t^{i'}|_s$ will still be more than 2γ-far. Hence, both $L_t^i|_s, L_t^{i'}|_s$ cannot be γ-close to the same L_s^j by the triangle inequality. Finally, we show in the proof that the separation of all lists will ensure that their size stays $\text{poly}(\frac{1}{\varepsilon})$. For more details on the necessary condition for constructing these injections, see Claim 4.5 and its proof.

Up until now we have explained in high level how to construct injective functions from $\overline{L}(t)$ to $\overline{L}(s)$ for $t \supseteq s$. Surjectivity requires more care because $\pi_{s,t}$ may still not be surjective even if $\overline{L}(s), \overline{L}(t)$ are dense and separated. It could be that there is a "new" function $L_s^j \in \overline{L}_{\tau,\delta}(s)$ that is not close to *any* function in the $L_t^i|_s$. Another way of saying this is that $\overline{L}(t)|_s$ may not be dense in $\overline{L}_{\tau,\delta}(s)$.

To overcome this we will first show a structural property. We show that for some fixed t, if $\pi_{s,t}$ is not surjective for a non-negligible fraction of the $s \subseteq t$, this implies that there is some $\delta' = O(\delta)$, and a function $g : t \to \Sigma$ such that $\alpha_{\delta'}(g) \geqslant \tau - \text{poly}(\varepsilon)$, that is δ'-far from all functions in $\overline{L}_{\tau,\delta}(t)$. Showing this is where the previous agreement theorems come into play; if it is true that $\overline{L}_{\tau,\delta}(t)$ is not dense for many of the $s \subseteq t$, this means that even after rerandomizing f_r for $r \in \bigcup_{g \in \overline{L}_{\tau,\delta}(t)}$, the ensemble will pass the agreement test with non-negligible probability. By the agreement soundness guarantee, there is a new function g such that $A_\delta(g) \geqslant \tau^3$. We then show that by scaling δ by a constant factor, it also holds that $A_{\delta'}(g) \geqslant \tau - \text{poly}(\varepsilon)$ (this happens because g needs to agree with the functions f_r for $r \notin \bigcup_{g \in \overline{L}_{\tau,\delta}(t)}$, since these are the functions that were not rerandomized, although this needs to be meticulously argued).

On the other hand (keeping t fixed), we will show that there exists δ, τ such that there are no new functions like this: We take any sequence $\{(\tau_i, \delta_i)\}_{i=1}^m$ where $\tau_1 = \varepsilon^2, \tau_{i+1} = \tau_i - \varepsilon^{50}$ and $\delta_{i+1} = 20\delta_i$. Suppose that for every such pair there is a "new" g such that $\alpha_{\delta_{i+1}}(g) \geqslant \tau_{i+1} \geqslant \varepsilon^{10}$ that is far from all functions in $\overline{L}_{\tau_i,\delta_i}(t)$. This will imply that $\mathbb{P}\left[A_{\delta_i}(g_i) \setminus \bigcup_{j=1}^{i-1} \alpha_{\delta_j}(g)\right] \geqslant \varepsilon^{10}$. This is because if g_i, g_j are far away from one another then most h_r that are close to g_i will be far from g_j. Thus

$$m\varepsilon^{10} \leqslant \sum_{i=1}^{m} \mathbb{P}_{r_1,r_2}\left[A_{\delta_i}(g_i) \setminus \bigcup_{j=1}^{i-1} A_{\delta_j}(g_j)\right] = \mathbb{P}_{r_1,r_2}\left[\bigcup_{i=1}^{m} A_{\delta_i}(g_i)\right] \leqslant 1.$$

[6] There is a subtlety here that even if $\alpha_\delta(L_t^i|_s) \geqslant \tau$, it could be that $\alpha_\delta(L_t^i) < \tau$. We will explain in the actual proof how to overcome this using sampling arguments, but for this overview let us just ignore this issue and assume that $\alpha_\delta(L_t^i|_s) = \alpha_\delta(L_t^i)$.

We conclude that $m \leqslant \varepsilon^{-10}$, i.e. that in every taking a long enough such sequence will result in some $\{(\tau_i, \delta_i)\}$ where this doesn't happen, which implies that $\overline{L}_{\tau_i,\delta_i}(t)$ is dense inside most $\overline{L}_{\tau_i,\delta_i}(s)$. This argument is depicted in Claim 4.7.

We use the pigeonhole principle to find a pair (τ_i, δ_i) and $M_{(\tau_i,\delta_i)} \subseteq X(d)$ of relative size $\mathrm{poly}(\varepsilon)$, such that for every $t \in M_{(\tau_i,\delta_i)}$ and for all but a negligible fraction of $s \subseteq t$ it holds that $\overline{L}_{\tau_i,\delta_i}(t)|_s$ is dense inside $\overline{L}_{\tau_{i+1},\delta_{i+1}}(s)$. Finally, using this set, we show that if $t \in M_{(\tau_i,\delta_i)}$ then most $s \subseteq t$ also have the same property (i.e. that $\overline{L}_{\tau_i,\delta_i}(s)|_u$ is dense inside most $\overline{L}_{\tau_{i+1},\delta_{i+1}}(u)$ for most $u \subseteq s$). Using sampling, we conclude that this property propagates, i.e. that for a small enough dimension $d' \ll d$ this density property holds for most $t' \in X(d')$. This argument is in Proposition 4.4.

Thus to conclude, we find our lists $\overline{L}(s)$ by first finding the aforementioned (τ_i, δ_i). Then we find lists $\overline{L}(s)$ that are γ-dense and 10γ-separated inside $\overline{L}_{\tau_{i+1},\delta_{i+1}}(s)$ and define the permutation $\pi_{s,t}(i) = j$ if and only if $L_t^i|_s \overset{1-\gamma}{\approx} L_s^j$. These are well defined for most $s \subseteq t$, and we use these for showing Lemma 3.4.

4.2 Proof of Lemma 3.4

This definition will be useful for the rest of this section.

Definition 4.2. Let $s \subseteq t$, $\delta > 0$ and L be a list of functions on t. We say that s *samples L well* if for every $g_1, g_2 \in L$ it holds that $|\mathrm{dist}(g_1|_s, g_2|_s) - \mathrm{dist}(g_1, g_2)| \leqslant \varepsilon^{100}$. We say that s *samples L's agreement* well if s samples L well and in addition, for every $g \in L$ it holds that $|\mathbb{P}[A_\delta(g)] - \mathbb{P}[A_\delta(g|_s)]| \leqslant \varepsilon^{100}$.

We also define the requirements we need from the lists $\overline{L}(s)$.

Definition 4.3. Let $d_1 \leqslant d_2 \leqslant d_3 \leqslant d$ be dimensions. Let $\tau, \delta, \gamma > 0$ be constants and let ℓ be integers. Let $M_i = M(\tau, \delta, \gamma, \ell) \subseteq X(d_i)$ be all the $t \in X(d_i)$ such that there exists some $\overline{L}(t) \subseteq \overline{L}_{\tau,\delta}(t)$ such that:

1. $\overline{L}(t)$ is 9γ-separated.
2. $\overline{L}(t)$ is γ-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(t)$.
3. $|\overline{L}(t)| = \ell$.
4. For $i = 2, 3$, and $i' < i$. Then the following set $P_t(i')$ has probability $\mathbb{P}_{s\subseteq t, s\in X(i')}[P_t(i')] \geqslant 1 - \exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k}))$. $P_t(i')$ is all the $s \subseteq t, s \in X(i')$ such that s samples $\overline{L}(t)$'s agreement well.

The proof of Lemma 3.4 relies on the following proposition.

Proposition 4.4. *There exists parameters $\tau = \mathrm{poly}(\varepsilon), \delta, \gamma \leqslant \eta \exp(\mathrm{poly}(1/\varepsilon))$ and $\ell = \mathrm{poly}(1/\varepsilon)$ such that the following holds. Let $d_1 \leqslant d_2 \leqslant d_3 \leqslant d$ be such that $\frac{d_3}{d} = \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$. Then $\mathbb{P}[M_i] \geqslant 1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$.*

We will also use this claim,

Claim 4.5. Let $L_1 = \{g_1, g_2, \dots, g_{\ell_1}\}, L_2 = \{h_1, h_2, \dots, h_{\ell_2}\}$ be such that L_1 is γ-dense in L_2 and such both L_1 and L_2 are 2γ-separated. Then there exists an injective function $\pi : [\ell_1] \to [\ell_2]$ such that $\pi(j) = k$ if and only if $g_j \overset{1-\gamma}{\approx} h_k$.

Proof of Lemma 3.4. Given this proposition, for every $s \in M_i$ we take such a list $\overline{L}(s) = \{L_s^1, L_s^2, \dots, L_s^\ell\}$ (for $s \notin M_i$ we take arbitrary lists). For $s \subseteq t$ such that $t \in M_i, s \in P_t(i') \cap M_{i'}$ we define $\pi_{s,t} : [\ell] \to [\ell]$ by

$$\pi_{s,t}(j) = k \Leftrightarrow L_t^j|_s \overset{1-\gamma}{\approx} L_s^k. \tag{4.1}$$

If either $t \notin M_i$ or $s \notin P_t(i') \cap M_{i'}$ we take $\pi_{s,t} = Id$ as an arbitrary choice. We show that the assumptions in Claim 4.5 hold, which implies that this function is well defined. First we note that if s samples $\overline{L}(t)$'s agreement well then $\overline{L}(t)|_s \subseteq \overline{L}_{\tau-\varepsilon^{100},\delta}(s)$ and $\overline{L}(t)|_u$ is 8γ-separated. Moreover if $\overline{L}(s)$ is γ-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(s)$, and $\overline{L}(t)|_s \subseteq \overline{L}_{\tau-\varepsilon^{100},\delta}(s)$ then in particular $\overline{L}(s)$ is γ-dense in $\overline{L}(t)|_s$. Finally, by definition

$\overline{L}(t)$ is also 9γ-separated. Hence Claim 4.5 implies that this is a well defined and injective function. As both $\overline{L}(t)|_s$ and $\overline{L}(t)$ have the same size ℓ, this is indeed a permutation.

Finally, we will prove that

$$\underset{s\subseteq t\subseteq u, s\in X(d_1), t\in X(d_2), u\in X(d_3)}{\mathbb{P}}[\pi_{s,t}\circ\pi_{t,u}=\pi_{s,u}] \geqslant 1-\exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k}).$$

Note that the following events are all occur with probability $1-\exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k})$:

1. $u\in M_1, s\in M_2, t\in M_3$.
2. $t\in P_u(i_2), s\in P_t(i_1)\cap P_u(i_1)$.
3. For every j and k,
$$\left|\mathrm{dist}(L_t^j|_s, L_u^k|_s)-\mathrm{dist}(L_t^j, L_u^k|_t)\right| \leqslant \gamma.$$

The first two items are by Proposition 4.4. The last item is via this claim, proven later on (we will need this claim also for later):

Claim 4.6. Fix t, L and δ. Then the fraction of $s\subseteq t, s\in X(d_1)$ that *don't* sample L's agreement well is at most $|L|^2\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$.

Indeed, if t samples $\overline{L}(u)$ well and s samples $\overline{L}(u)|_t\cup\overline{L}(t)$ well, then distances between members of $\overline{L}(u)|_t\cup\overline{L}(t)$ are preserved and the third item holds (and both these events occur with probability $1-\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k}))$.

Thus to prove the theorem we assume $s\subseteq t\subseteq u$ are such that all these events occur and show that $\pi_{s,t}\circ\pi_{t,u}=\pi_{s,u}$. Let $k=\pi_{s,t}\circ\pi_{t,u}(j)$ and $k'=\pi_{s,u}(j)$. Then it holds that $L_u^j|_s \overset{1-\gamma}{\approx} L_s^{k'}$, and that $L_u^j|_t \overset{1-\gamma}{\approx} L_t^{\pi_{t,u}(j)}$ and $L_t^{\pi_{t,u}(j)}|_s \overset{1-\gamma}{\approx} L_s^k$. By the third item, distances are approximately preserved, so $L_u^j|_s \overset{1-2\gamma}{\approx} L_t^{\pi_{u,t}(j)}|_s$. Hence by the triangle inequality it holds that $L_u^j|_s \overset{1-3\gamma}{\approx} L_s^k$, which implies that $L_s^k \overset{1-4\gamma}{\approx} L_s^{k'}$. By 10γ-separation of $\overline{L}(u)$ it holds that $k=k'$. □

The proof of Claim 4.5 is direct.

Proof of Claim 4.5. By density of L_2 in L_1, it holds that for every g_j there exists some h_k such that $g_j \overset{1-\gamma}{\approx} h_k$. Moreover, there is only one such k: assume that for some j there are k,k' such that $g_j \overset{1-\gamma}{\approx} h_k$ and $g_j \overset{1-\gamma}{\approx} h_{k'}$. Then by the triangle inequality $h_k \overset{1-2\gamma}{\approx} h_{k'}$, which by 2γ-separation implies that $k=k'$. Thus it is a well defined function.

The same argument shows that this is an injection. Indeed, Let $k=\pi(j)=\pi(j')$. That is, $h_j \overset{1-\gamma}{\approx} h_k \overset{1-\gamma}{\approx} h_{j'}$. Then $h_j \overset{1-2\gamma}{\approx} h_{j'}$. by 2γ-separation of $\overline{L}(s)|_u$, this implies that $j=j'$, i.e. the function is an injection. □

So is the proof of Claim 4.6.

Proof of Claim 4.6. Let us start with sampling well. There are $\leqslant \binom{|L|}{2}$ pairs $g,g'\in L$ so by a union bound it is enough to argue that for at most $\exp(-\mathrm{poly}(\varepsilon)d_1)$ of the $s\subseteq t$ it holds that $|\mathrm{dist}(g,g')-\mathrm{dist}(g|_s,g'|_s)|>\varepsilon^{100}$. Let $A\subseteq t$ be the set of $v\in t$ such that $g(v)\neq g'(v)$. By Theorem 2.32 the fraction of $s\subseteq t$ such that $|\mathbb{P}[A]-\mathbb{P}_{v\subseteq s}[A]|>\varepsilon^{100}$ is $\exp(-\mathrm{poly}(\varepsilon)d_1)$. On the other hand, $\mathbb{P}[A]=\mathrm{dist}(g,g')$ and $\mathbb{P}_{v\subseteq s}[A]=\mathrm{dist}(g|_s,g'|_s)$ giving us the that at most $\binom{|L|}{2}\exp(-\mathrm{poly}(\varepsilon)d_1)$ don't sample L well.

Now let us make sure that most $s\subseteq t$ also sample L's agreement well. Indeed, let $g\in L$. By Claim 2.33 there is at most $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ of the $s\subseteq t$ such that

$$|A_\delta(g)-A_\delta(g|_s)|>\varepsilon^{100}.$$

Another union bound gives us us the claim. □

The proof of Proposition 4.4 will follow from these assertions. We prove them after the proof of proposition.

Claim 4.7. Let $d_1 \leqslant d$ and fix $t \in X(d)$ such that $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}_t) \geqslant \varepsilon - \varepsilon^{100}$. There exists $m \leqslant \frac{1}{\varepsilon^{10}}$ such that the following holds for $\tau_m = \varepsilon^2(1-\varepsilon^{50})^i$ and $\delta_m = 3\eta 20^m$. There exists $L \subseteq \overline{L}_{\tau_m,\delta_m}(t)$ such that for $1 - \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ of the $u \subseteq t, u \in X(d_1)$ it holds that $L|_u$ is $20\delta_m$-dense in $\overline{L}_{\tau_{m+1},\delta_m}(u)$.

Claim 4.8. Let $\delta, \tau > 0$ and let $t \in X(d)$. Then there exists some $\gamma = \frac{24}{23}24^i\delta$ for $i = 1, 2, \ldots, \varepsilon^{-10}$ and some $L \subseteq \overline{L}_{\tau,\delta}(t)$ such that L is γ-dense, $|L| \leqslant \frac{1}{\tau^3}$ and 23γ-separated.

Proof of Proposition 4.4. By Claim 2.33 there are $1 - \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ $t \in X(d)$ such that $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}_t) \geqslant \varepsilon - \varepsilon^{100}$ (we use Claim 2.33 on the set of $\{r_i\}$ such that $\{f_{r_i}\}$ pass the agreement test). By Claim 4.7, for every such $t \in X(d)$ there exists

$$\tau = \varepsilon^2(1-\varepsilon^{50})^i, \delta = 3\eta 20^i$$

(for $i \in \{1, 2, \ldots, \frac{1}{\varepsilon^{10}}\}$), and a list $L'(t) \subseteq \overline{L}_{\tau,\delta}(t)$ as in Claim 4.7.

Moreover, by Claim 4.8 there exists some

$$\gamma' \in [\frac{24}{23}24\delta, \frac{24}{23}24^{\frac{1}{\varepsilon^{10}}}\delta]$$

and $\overline{L}(t) \subseteq \overline{L}_{\tau,\delta}(t)$ such that $\overline{L}(t)$ has length $\leqslant \operatorname{poly}(1/\varepsilon)$, $\overline{L}(t)$ is γ'-dense in $\overline{L}_{\tau,\delta}(t)$ and $23\gamma'$-separated.

Thus by the pigeonhole principle, there exists $\tau, \delta, \gamma', \ell$ and a set $M_d = M_d(\tau, \delta, \gamma', \ell) \subseteq X(d)$ of probability $\mathbb{P}[M_d] \geqslant \operatorname{poly}(\varepsilon)$, such that for every $t \in X(d)$ there exists $L'(t) \subseteq \overline{L}_{\tau,\delta}(t)$ as in Claim 4.8 and $\overline{L}(t) \subseteq \overline{L}_{\tau,\delta}(t)$ of size $\left|\overline{L}(t)\right| = \ell$, $\overline{L}(t)$ is γ-dense in $\overline{L}_{\tau,\delta}(t)$ and 23γ-separated.

Let us define $M_i \subseteq X(d_i)$ be the set of $s \in X(d)$ such that there exists some $t \in M_d, t \supset s$ such that:

1. $s \subseteq t$ samples $L'(t) \cup \overline{L}(t)$ well.
2. $L'(t)$ is 20δ-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(s)$.
3. If $i = 2, 3$ and $i' < i$, then $\mathbb{P}_{u \subseteq s}[M_{i'}] \leqslant \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ (This definition is recursive but M_i is always well defined).

We show that for every $s \in M_i$ the proposition holds for τ, δ, ℓ and $\gamma = 2\gamma'$. We take $\overline{L}(s) = \overline{L}(t)|_s$ for some arbitrary $t \in M_d$ that contains s and such that the items above holds for this t. By the fact that $\overline{L}(t)$ is sampled well we have that $\overline{L}(s) \subseteq \overline{L}_{\tau-\varepsilon^{100},\delta}(t)$. Moreover if $\overline{L}(t)$ is $23\gamma' = 11.5\gamma$-separated and in particular it holds that $11\gamma'$-separated (and in particular, all restrictions indeed result in distinct functions, i.e. $\left|\overline{L}(s)\right| = \left|\overline{L}(t)|_s\right| = \ell$).

Next we show density in $\overline{L}_{\tau-\varepsilon^{100},\delta}(s)$. Note that $L'(t)|_s$ is 20δ-dense in $\overline{L}_{\tau-\varepsilon,\delta}(s)$, and $\overline{L}(t)$ is γ'-dense in $L'(t)$ (not to say that this is a subset of $L'(t)$, just that for every $g' \in L'(t)$ there exists some $g \in \overline{L}(t)$ such that $g' \overset{1-\gamma'}{\approx} g$). Thus by the fact that s samples $L'(t) \cup \overline{L}(t)$ well, this implies that $\overline{L}(t)$ is $\gamma' + \varepsilon^{100}$-dense in $L'(t)$. Hence, it holds that $\overline{L}(t)$ is $\gamma' + \varepsilon^{100} + 20\delta \leqslant \gamma$-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(s)$.

Finally, we show that when $i = 2, 3$ and $i' \leqslant i$, it holds that $P_s(i')$ from the definition of the proposition has size $1 - \exp(-\Omega(\operatorname{poly}(\varepsilon)\frac{d_1}{k}))$. Let us go one by one:

1. For every t, the fraction of $u \subseteq s$ that don't sample $\overline{L}(s)$ well is $\exp(-\Omega(\operatorname{poly}(\varepsilon)\frac{d_1}{k}))$ by Claim 4.6.
2. The probability that $u \subseteq s$ is in $M_{i'}$ is $\exp(-\Omega(\operatorname{poly}(\varepsilon)\frac{d_1}{k}))$. Let us show that in this case $\overline{L}(t)|_s$ is 2γ-dense. $u \in M_{i'}$, therefore there is some $\overline{L}(u)$ of size ℓ that is γ-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(u)$, and 11γ-separated (by what we already proved above). Thus if we show that for every $g \in \overline{L}(u)$ there is some $g' \in \overline{L}(t)|_s$ such that $g \overset{1-\gamma}{\approx} g'$ it will follow that $\overline{L}(t)|_s$ is 2γ-dense. By Claim 4.5, the fact that $\overline{L}(u)$ is dense in $\overline{L}(t)|_s$ and the fact that both are 2γ-separated, there is an injective function $\phi : \overline{L}(t)|_s \to \overline{L}(u)$ where $\phi(g') = g$ if and only if $g \overset{1-\gamma}{\approx} g'$. As both lists have the same size, this shows that this is a surjection, i.e. that for every $g \in \overline{L}(u)$ there is some $g' \in \overline{L}(t)|_s$.

To conclude, let us show that $\mathbb{P}[M_i] \geqslant 1 - \exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k})$. Let us begin with M_1. Let B_1 be the fraction of s's such that $\mathbb{P}_{t \supseteq s}[M_d] \leqslant \mathrm{poly}(\varepsilon)$. Its fraction is $\mathbb{P}[B_1] = O(\frac{d_1}{d\,\mathrm{poly}(\varepsilon)}) = \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ by Corollary 2.3. Let $B_2 \subseteq X(d_i)$ be the event that more than $\frac{1}{3}$ of the $t \supseteq s, t \in M_d$ have:

1. Either s doesn't sample $\overline{L}(t) \cup L'(t)$ well.
2. $L'(t)$ is 20δ-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(s)$.

On the one hand, by Claim 4.6 and Claim 4.7 there are at most $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ such pairs. On the other hand, every $s \in B_2 \setminus B_1$ contributes $\mathrm{poly}(\varepsilon)$ such pairs. By Markov's inequality, $\mathbb{P}[B_2 \setminus B_1] = \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$. Hence, if $s \notin B_1 \cup B_2 = B_1 \cup (B_2 \setminus B_1)$, then $s \in M_1$ so $\mathbb{P}[M_1] = 1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$.

Continuing with M_2. The same arguments for M_1 show that $1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ of the $s \in X(d_2)$ sample $\overline{L}(t) \cup L'(t)$ well and that $L'(t)$ is 20δ-dense in $\overline{L}_{\tau-\varepsilon^{100},\delta}(s)$. So we concentrate on the last property. The probability that $s \in X(d_2)$ contains more than
$probM_1{}^{1/2}$-fraction of $u \notin M_1$ is at most $(\mathbb{P}[M_1])^{1/2}$ (by Markov's inequality), and this is also $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$, albeit with a slightly worse $\mathrm{poly}(\varepsilon)$ than before, since since $\mathbb{P}[M_1] = \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$. Hence $\mathbb{P}[M_2] = 1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$. The proof for M_3 is similar. □

Proof of Claim 4.8. Consider the following method, resembling [Din+18a]. Let $L_1 \subseteq \overline{L}_{\tau,\delta}(t)$ be a maximal 24δ-separated list. Consider the following chain $L_1 \supseteq L_2 \supseteq L_3 \dots$ where every L_i is a maximal $24^i\delta$-separated set inside L_{i-1}. Obviously, the sizes of these sets are monotonically decreasing (but they are always non-empty), therefore there must be some j such that $L_j = L_{j+1}$. Take as $L = L_i$ for i being the first such j.

Recall that $\gamma_i = \frac{24}{23}24^i$. By definition, L_{i+1} is $23\gamma_i = 24^{i+1}\delta$-separated. Let us see that is it also $26 \cdot 24^i\delta$-dense in $\overline{L}_{\tau,\delta}(t)$. First note that every L_j is $24^i\delta$-dense in L_{j-1} (if it weren't, we could have added to L_j another $g \in L_{j-1} \setminus L_j$ that is far from the current L_j, contradicting maximality). Thus for every $g \in \overline{L}_{\tau,\delta}(t)$ there is a sequence $g = g_0, g_1, g_2, \dots, g_i$ where $g_j \in L_j$ for $j > 0$, such that $\mathrm{dist}(g_{j-1}, g_j) \leqslant 24^j\delta$. Thus concluding that

$$\mathrm{dist}(g, g_i) \leqslant \sum_{j=1}^{i} 24^j = \frac{24}{23}(24^i - 1)\delta \leqslant \frac{24}{23} \cdot 24^i = \gamma_i.$$

Finally, we must show that $i \leqslant \frac{1}{\tau^3}$. If we show that $|L_1| \leqslant \frac{1}{\tau^3}$ this will follow, since $|L_i| \geqslant 1$ and for $j \leqslant i$, $|L_j| \leqslant |L_{j-1}| - 1$. We distinguish this part of the proof to a separate claim since we will need it in also later on.

Claim 4.9. let $\tau \gg \exp(-\Omega(\delta^2 k))$ and let δ and $\gamma \geqslant 3\delta$. Let $L \subseteq \overline{L}_{\tau,\delta}(t)$ be a γ-separated set. Then $|L| \leqslant \frac{1}{\tau^3}$. □

Proof of Claim 4.9. Assume towards contradiction that $|L_1| > \frac{1}{\tau^3}$ and take some subset $L' \subseteq L_1$ of size $\frac{1}{\tau^3} + 1$. As it is at least 3δ-separated, for every distinct $g, g' \in L'$, the fraction of $r \in X(k)$ such that $\mathrm{dist}(g|_r, g'|_r) \leqslant 2\delta$ (i.e. $g|_r \overset{1-2\delta}{\approx} g'|_r$) is at most $\exp(-\Omega(\delta^2 k)) \leqslant \tau^{10}$. This shows that there are almost no intersection between the supports of g and g', since $\mathrm{supp}_\delta(g) \cap \mathrm{supp}_\delta(g') \subseteq \{r \in X(k) \mid \mathrm{dist}(g|_r, g'|_r) \leqslant 2\delta\}$. That is, $\mathbb{P}[\mathrm{supp}_\delta(g) \cap \mathrm{supp}_\delta(g')] \leqslant \exp(-\Omega(\delta^2 k)) \leqslant \tau^{10}$.

Denote by

$$D(g) = A_\delta(g) \setminus \bigcup_{g' \in L', g' \neq g} A_\delta(g').$$

Then in particular

$$\mathbb{P}[D(g)] \geqslant \mathbb{P}[A_\delta(g)] - \sum_{g' \in L', g' \neq g} \mathbb{P}\left[A_\delta(g) \cap A_\delta(g')\right] \tag{4.2}$$

$$\geqslant \mathbb{P}[A_\eta(g)] - \sum_{g' \in L', g' \neq g} 2\,\mathbb{P}\left[\mathrm{supp}_\delta(g) \cap \mathrm{supp}_\delta(g')\right] \tag{4.3}$$

$$\geqslant \tau - 2\tau^6. \tag{4.4}$$

On the other hand, the sets $\{D(g)\}_{g\in L'}$ are mutually disjoint, namely $1 \geqslant \sum_{g\in L'} \mathbb{P}\left[D(g)\right] \geqslant \frac{\tau-2\tau^6}{\tau^3+1} > 1$ which is a contradiction. $\square$

4.3 Proof of Claim 4.7

We start with a weaker claim than Claim 4.7, saying that $\overline{L}_{\tau^3,\delta}(t)$ restricted to $s \subseteq t$ is dense in $\overline{L}_{\tau,\delta}(s)$ for most s. This claim is where we use the agreement soundness of the original distribution.

Claim 4.10. Fix t,τ, δ and $\gamma \geqslant 3\delta$. Let $L \subseteq \overline{L}_{\tau^3,\delta}(t)$ be a γ-dense and γ-separated list. Then

$$\mathop{\mathbb{P}}_{s\subseteq t, s\in X(d_1)}\left[L|_s \text{ is } (\gamma+4\delta)\text{-dense in } \overline{L}_{\tau,\delta}(s)\right] \geqslant 1-\exp(-\Omega(\frac{d_1}{k})).$$

Proof of Claim 4.10. The claim will follow from these assertions.

1. If L is γ-dense and has size $\mathrm{poly}(1/\varepsilon)$ in $\overline{L}_{\tau^3,\delta}(t)$ then

$$\mathop{\mathbb{P}}_{r_1,r_2\subseteq t, r_1,r_2\sim D}\left[f_{r_1} = f_{r_2} \text{ and } r_1,r_2 \notin \bigcup_{g\in L} \mathrm{supp}_{\gamma+2\delta}(g)\right] \leqslant 2\tau^{1.5}.$$

2. This quantity is sampled well. That is, let $P = \left\{(r_1,r_2) \;\middle|\; f_{r_1} = f_{r_2} \text{ and } r_1,r_2 \notin \bigcup_{g\in L}\mathrm{supp}_{\gamma+2\delta}(g)\right\}$. Then for $1-\exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k}))$ of the $s\subseteq t$ it holds that

$$\left|\mathop{\mathbb{P}}_{r_1,r_2\subseteq s}\left[P\right] - \mathbb{P}\left[P\right]\right| \leqslant \varepsilon^{100}.$$

Let us explain why. If there exists some $h \in \overline{L}_{\tau,\delta}(s)$ such that $\mathrm{dist}(h,g) \geqslant \gamma+4\delta$ for all $g \in L|_s$, then $A_\delta(h) \geqslant \tau$ and moreover it holds that $\mathbb{P}\left[\mathrm{supp}_\delta(h)\cap \mathrm{supp}_{\gamma+2\delta}(g)\right] = \exp(-\mathrm{poly}(\varepsilon)k) \ll \mathrm{poly}(\varepsilon)$ (since whenever $f_r \overset{1-\delta}{\approx} h|_r$ and $f_r \overset{1-\gamma-2\delta}{\approx} g|_r$ then it holds that $\mathrm{dist}(g|_r, h|_r) \leqslant \gamma+3\delta$ which happens with $\exp(-\mathrm{poly}(\varepsilon)k)$ probability at most). By Claim 4.9, $|L| = \mathrm{poly}(1/\varepsilon)$ so all these intersections are negligible. In particular, more than (say) half of the edges $(r_1,r_2) \in A_\delta(h)$ have that $r_1,r_2 \notin \bigcup_{g\in L}\mathrm{supp}_{\gamma+3\delta+\varepsilon^{100}}(g)$. Thus $\mathbb{P}_{r_1,r_2\subseteq s, r_1,r_2\sim D}\left[P\right] \geqslant \frac{1}{2}A_\delta(h) \geqslant \frac{1}{2}\tau$. On the other hand, by the first item $\mathbb{P}\left[P\right] \leqslant 2\tau^{1.5}$. By the second item $\mathbb{P}_{r_1,r_2\subseteq s}\left[P\right] \leqslant \mathbb{P}\left[P\right] + \varepsilon^{100} \leqslant 3\tau^{1.5}$ with probability as high as $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ and hence this cannot occur for more than $\exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$.

The second item follows directly from Claim 2.33. The main effort is to show the first item. Assume that it doesn't hold. Define $\mathcal{F}'_t$ by erasing f'_r for every $r \in \bigcup_{g\in L}\mathrm{supp}_{\gamma+3\delta+\varepsilon^{100}}(g)$ (and taking $f'_r = f_r$ for the rest of the faces)[7]. Then it still holds that

$$\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}'_t) \geqslant 2\tau^{1.5}.$$

By the agreement soundness (assuming that $\tau^{1.5} \geqslant \varepsilon_0$), there is some $h \in \overline{L}_{\tau^3,\eta}(t)$ such that $A^{\mathcal{F}'}_\eta(h) \geqslant \tau^3$. On the other hand, there must by some $g \in L$ such that $\mathrm{dist}(h,g) \leqslant \gamma$. This implies that $\mathbb{P}\left[\mathrm{supp}^{\mathcal{F}'}_\eta(h) \setminus \mathrm{supp}_{\gamma+2\delta}(g)\right] \leqslant \exp(-\mathrm{poly}(\delta^2 k)) \ll \tau^3$. This is a contradiction since $A^{\mathcal{F}'}_\eta(h) \subseteq \mathrm{supp}^{\mathcal{F}'}_\eta(h) \setminus \mathrm{supp}_{3\delta+\gamma}(g)$ (as we erasing g's support). $\square$

Proof of Claim 4.7. Recall that $\tau_m = \varepsilon^2(1-\varepsilon^{10})^m$ and that $\delta_m = 3\eta\cdot 20^m$. Fix some t. For every τ_m,δ_m we choose some (arbitrary) $L^{aux}_{m+1}(t) \subseteq \overline{L}_{\tau^3_{m+1},\delta_m}(t)$ that is 3δ-dense and 3δ-separated. We will sequentially construct lists $L_m = \{g_1,g_2,\dots,g_m\} \subseteq \overline{L}_{\tau_m,\delta_m}(t)$, for $m = 1,2,\dots$ until L_m satisfies the claim. We will

[7] Formally, when we "erase" a function, we actually set the f'_r to be such that it agrees with no other functions. This can be done easily by extending the alphabet Σ to a large enough set.

define a potential function $p(m)$ that takes values in $[0,1]$ and show that whenever L_m doesn't satisfy the properties we need then $p(m) \geqslant m \cdot \mathrm{poly}(\varepsilon)$, thus concluding that L_m will satisfy the claim for some $m \leqslant \mathrm{poly}(1/\varepsilon)$. The potential function is

$$p(m) = \mathop{\mathbb{P}}_{r_1,r_2 \subseteq t, r_1,r_2 \sim D}\left[\bigcup_{g_j \in L_m} A_{9\delta_{j-1}}(g_j)\right].$$

For $m = 0$ we take $L_0 = \emptyset \subseteq \overline{L}_{\tau_0,\delta_0}(t)$. Let us describe how to construct L_m using L_{m-1}, provided that L_{m-1} is not $\delta_m = 20\delta_{m-1}$-dense in $\overline{L}_{\tau_{m+1},\delta_m}(u)$ for $1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ of the $u \in X(d_1)$:

1. Find some $u \subseteq t$ such that:
   - L_{m-1} is not $20\delta_{m-1}$-dense in $\overline{L}_{\tau_{m+1},\delta_m}(u)$.
   - u samples $L^{aux}_{m+1} \cup L_{m-1}$ well.
   - L^{aux}_{m+1} is $7\delta_m$-dense in $\overline{L}_{\tau_{m+1},\delta_m}(u)$.
   - For every $g \in L^{aux}_{m+1}(t)$ it holds that $\left|A^{\mathcal{F}_t}_{9\delta_{m-1}}(g) - A^{\mathcal{F}_u}_{9\delta_{m-1}}(g|_u)\right| \leqslant \varepsilon^{100}$.[8]
2. Find some $h \in \overline{L}_{\tau_{m+1},\delta_m}(u)$ that is $20\delta_{m-1}$-far from every element in L_{m-1}. There exists such an h since L_{m-1} is not $20\delta_{m-1}$-dense in $\overline{L}_{\tau_{m+1},\delta_m}(u)$.
3. Take some $g = g_m \in L^{aux}_{m+1}$ such that $\mathrm{dist}(g|_u, h_u) \leqslant 7\delta_{m-1}$. Set $L_m = L_{m-1} \cup \{g_m\}$. This step is possible since $L^{aux}_{m+1}(t)|_u$ is $7\delta_m$-dense inside $\overline{L}_{\tau_{m+1},\delta_{m+1}}(u)$.

Observe that because $\mathrm{dist}(g_m|_u, h) \leqslant 7\delta_{m-1}$ then the fraction of $r \in X(k)$ such that $\mathrm{dist}(g_m|_r, h|_r) > 8\delta_{m-1}$ is $\exp(-\Omega(\delta^2 k)) \ll \varepsilon^{100}$. In addition, when $f_r \overset{1-\delta_m}{\approx} h_r$ and $h_r \overset{1-8\delta_{m-1}}{\approx} g_m|_r$ then $h_r \overset{1-9\delta_{m-1}}{\approx} g_m$. Hence it holds that $\mathbb{P}\left[A_{9\delta_{m-1}}(g_m|_u)\right] \geqslant \mathbb{P}\left[A_{\delta_{m-1}}(h)\right] - \varepsilon^{100}$. By assumption on u it holds that $\mathbb{P}\left[A_{9\delta_{m-1}}(g_m)\right] \geqslant \mathbb{P}\left[A_{9\delta_{m-1}}(g_m|_u)\right] - \varepsilon^{100}$ so

$$\mathbb{P}\left[A_{9\delta_{m-1}}(g_m)\right] \geqslant \mathbb{P}\left[A_{9\delta_{m-1}}(h)\right] - 2\varepsilon^{100} \geqslant \tau_m - 2\varepsilon^{100} \geqslant \tau_{m+1}.$$

That is, $g_m \in \overline{L}_{\tau_{m+1},\delta_{m+1}}(t)$ (and thus every $L_m \subseteq \overline{L}_{\tau_{m+1},\delta_{m+1}}(t)$).

Moreover, note that for every $j \leqslant m$ it holds that $\mathrm{dist}(g_m|_u, g_j|_u) \geqslant \mathrm{dist}(h, g_j|_u) - \mathrm{dist}(g_m|_u, h) \geqslant 13\delta_m$. The face u samples $L^{aux}_{m+1} \cup L_{m-1}$ well, hence $\mathrm{dist}(g_m, g_j) \geqslant \mathrm{dist}(g_m|_u, g_j|_u) - \delta_{m-1} \geqslant 12\delta_{m-1}$. This is much larger than $9\delta_{m-1} + \delta_{j-1}$. Hence (as before) it holds that $\mathbb{P}\left[\mathrm{supp}_{9\delta_{j-1}}(g_j) \cap \mathrm{supp}_{9\delta_{m-1}}(g_m)\right] \ll \varepsilon^{100}$. This implies that $\mathbb{P}\left[A_{9\delta_{j-1}}(g_j) \setminus \bigcup_{j' \neq j} A_{9\delta_{j'-1}}(g_{j'})\right] \geqslant \frac{1}{2}\mathbb{P}\left[A_{9\delta_{j-1}}(g_j)\right]$.

Thus it holds that while $m \leqslant \varepsilon^{-30}$, then $p(m) \geqslant \frac{1}{2}\sum_{j=1}^{m} \mathbb{P}\left[A_{9\delta_{j-1}}(g_j)\right] \geqslant \frac{m}{2}\varepsilon^{10}$. For $m > 2\varepsilon^{10}$ this is greater than 1 hence the process must stop beforehand. $\square$

5 Constructing a cover

In this subsection we prove Lemma 3.7. We do so in two parts: first we construct a simplicial complex Y and prove that it covers X. In the second part we explicitly construct the isomorphism $\iota : FY \to \widetilde{FX}$. In addition to proving Lemma 3.7, we will use the explicit definition of ι for proving Claim 3.8 later on.

Let X be a d-dimensional clique complex. Let d_1 be such that $3d_1 \leqslant d - 2$. Let $FX = F(X, d_1)$ be as in Definition 2.30 and assume that X is well connected as in Definition 2.31. Fix $\nu : \widetilde{FX} \to FX$ to be an ℓ-cover. In this section, we will show how to construct an ℓ-cover of $\rho : Y \to X$ such that $\widetilde{FX} \cong FY$.

The main idea will be to encode the vertices $v \in X(0)$ by the links $FX_v \subseteq FX$ (these are the $FX_v = F(X_v, d_1)$ as in Definition 2.30 which are contained as sub complexes inside FX). We observe that this

[8] The fraction of $u \subseteq t$ such that these items do not hold is $\exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k}))$ by Claim 2.33 and Claim 4.10. Hence if L_m is not $20\delta_m$-dense for for $1 - \exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k}))$ we can find a $u \subseteq t$ such that this holds.

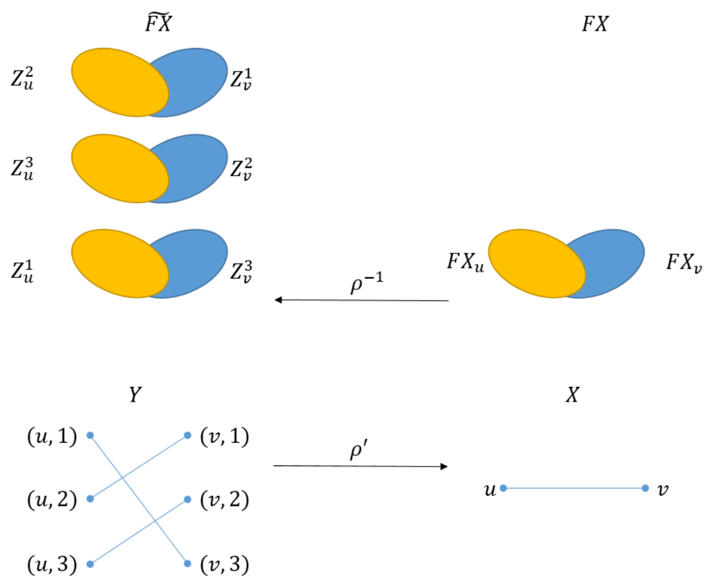

Figure 3: Constructing a cover of X from a cover of FX

encoding has the property that if $v \sim u$ then $FX_v \cap FX_u = FX_{uv}$. We will use these non-empty intersections to define the edges in the cover. For every $r \in X$, let $Z_r = \nu^{-1}(FX_r) \subset \widetilde{FX}$. Using the well-connectedness of X, we will show that

Claim 5.1.

1. For every non empty $r \in X^{\leqslant d_1}$, Z_r decomposes to exactly ℓ connected-components[9] $Z_r^1, Z_r^2, ..., Z_r^\ell \subseteq \widetilde{FX}$, such that ν is an isomorphism between each component and FX_r, $\nu|_{Z_r^i} : Z_r^i \xrightarrow{\sim} FX_r$.

2. For every non empty $r, s \in X^{\leqslant d_1}$ such that $r \subseteq s$ there is a permutation $\pi_{r,s} : [\ell] \to [\ell]$ such that $Z_s^i \subseteq Z_r^j$ if and only if $j = \pi_{r,s}(i)$.

We prove this claim at the end of this subsection. We define the cover Y of X as follows. Let $\psi \in C^1(X, Sym(\ell))$ be

$$\psi(uv) = \pi_{u,uv}^{-1} \circ \pi_{v,uv} \tag{5.1}$$

where $\pi_{u,uv}, \pi_{v,uv}$ are the permutations promised by the second item of Claim 5.1. Let $Y = X^\psi$, i.e. Y is the clique complex such that $Y(0) = X \times [\ell]$ and $\{(u,i),(v,j)\} \in Y(1)$ if $uv \in X(1)$ and $j = \psi(uv).i$. Then

Lemma 5.2. *Y is a cover of X. Moreover, it holds that $(v,i) \sim (u,j)$ if and only if $vu \in X(1)$ and there exists some k such that $Z_v^i \cap Z_u^j \supseteq Z_{vu}^k$.*

For the proof of Lemma 5.2 we use the fact that the $\pi_{r,s}$ in Claim 5.1 give us a cocycle on the flag complex of FX (see below). For this we recall the following corollary:

Corollary (Restatement of Corollary 2.29)**.** *Let X be a two dimensional simplicial complex and let $\phi \in Z^1(GX, Sym(\ell))$. Let $\psi = \psi_\phi : X(1) \to Sym(\ell)$ be given by $\psi(uv) = \phi(uv,v)\phi(u,uv)$. Then $\psi \in Z^1(X, Sym(\ell))$.*

In this notation $\phi(u,uv) = \pi_{u,uv}$ and thus $\psi(uv) = \pi_{u,uv}^{-1} \circ \pi_{v,uv}$.

[9]When we say "connected components" we mean that the induced complex Z_r has these ℓ connected components. Of course it may hold that there are paths from Z_i to Z_j in $\widetilde{FX}$, but they must go through some vertex $v \in \widetilde{FX} \setminus Z_r$.

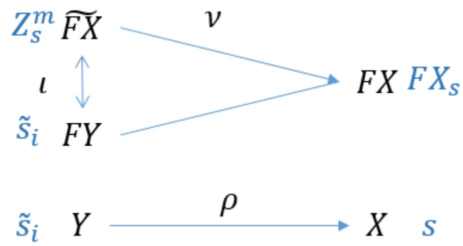

Figure 4: Commutative diagram

Proof of Lemma 5.2. Let us show that Y is a cover. This is equivalent to showing that ψ is a cocycle. Let $X^{\leqslant 2}$ be the two skeleton of X and let G be its flag-complex as in Definition 2.27. Let $\phi : GX(1) \to Sym(\ell)$ be given by $\phi(r,s) = \pi_{r,s}$ for every $r \subseteq s$ (where $\pi_{r,s}$ are the permutations in Claim 5.1). If we show that $\phi \in Z^1(GX, Sym(\ell))$ then by Corollary 2.29 ψ is also a cocycle. This amounts to show that for every $uvw \in X$ it holds that $\pi_{uv,uvw} \circ \pi_{u,uv} = \pi_{u,uvw}$.

Indeed, fix i and let $k = \pi_{u,uvw}(i)$, and $k' = \pi_{uv,uvw} \circ \pi_{u,uv}(i)$. By definition this implies that $Z_{uvw}^k \subseteq Z_u^i$ and that $Z_{uvw}^{k'} \subseteq Z_{uv}^{\pi_{u,uv}(i)} \subseteq Z_u^i$. The second item of Claim 5.1 says that $Z_{uvw}^{k'} \subseteq Z_u^i$ implies that $k' = \pi_{u,uvw}(i) = k$.

Next, for the "moreover" statement, note that $(v,i) \sim (u,j)$ if and only if there is some k such that $\pi_{v,uv}(i) = k$ and such that $\pi_{u,uv}(j) = k$. This occurs if and only if $Z_{uv}^k \subset Z_u^i \cap Z_v^j$. □

Proof of Claim 5.1. Let us start with the first item. We begin by explaining why it is enough to prove it assuming r is a vertex. Suppose that for every vertex $v \in X(0)$ it holds that Z_v decomposes to ℓ connected-components. Let $r \in X^{\leqslant d_1}$ and let $v \in r$ be an arbitrarily chosen vertex inside r. Take Z_r^i to be the subcomplex induced by $Z_r^i(0) := Z_v^i(0) \cap \nu^{-1}(FX_r(0))$. As every Z_v^i is isomorphic to FX_v then in particular, the induced sub-complex Z_r^i is isomorphic to FX_r. By well connectedness of X, FX_r is connected and thus we get that these are ℓ connected-components (they cannot be connected to one another because each lies in a different Z_v^i).

Now we show the first item for every vertex $v \in X(0)$. Towards this end, we note that the restriction $\nu|_{Z_v} : Z_v \to FX_v$ is also a covering map (for a proof of this statement, see Claim 2.22). From the well connectivity of X, FX_v is simply connected, so Z_v must decompose to ℓ disjoint connected-components.

We move to the second item. Fix $r \subseteq s$. We need to show that for every $i \in [\ell]$ there is a unique j such that $Z_s^j \subseteq Z_r^i$, so let us fix some $i \in [\ell]$. By definition FX_s, and $\nu|_{Z_r^i}$ is an isomorphism between Z_r^i and FX_r. In particular, it follows from the definition of an isomorphism that $\nu^{-1}(FX_s) \cap Z_r^i$ is isomorphic (by $\nu|_{Z_r^i}$) to $FX_s \subset FX_r$, and thus is one of the connected components $Z_s^j \subseteq \nu^{-1}(FX_s)$ promised by the first item we already proved. This is the index j we needed. Note that it is unique since $Z_s^j = \nu^{-1}(FX_s) \cap Z_r^i$ (so there are no more faces in $\nu^{-1}(FX_s)$ and Z_r^i that are not already in Z_s^j). □

5.1 The isomorphism between FY to $\widetilde{FX}$

We continue with the notation in the previous subsection. In this subsection we describe the isomorphism $\iota : FY \to \widetilde{FX}$ explicitly. See Figure 4 for the relations between X, Y, FX, FY and $\widetilde{FX}$.

The idea is simple. Given some $\tilde{s} = \{(v_0, i_0), (v_1, i_1), ..., (v_{d_1}, i_{d_1})\} \in FY(0)$, such that $\rho(\tilde{s}) = s = \{v_0, v_1, \dots, v_{d_1}\}$, we need to specify which $j \in [\ell]$ is such that $\iota(\tilde{s}) = (s,j)$. Using arguments similar to Lemma 5.2, we will show that the face $\tilde{s}$ is such that $\bigcap_{m=0}^{d_1} Z_{v_m}^{i_m}$ contains a "copy" Z_s^j of FX_s. We will define $\iota(\tilde{s}) = (s,m)$ which is connected to this copy. See Figure 5 for a pictorial description of ι. We first claim that ι is well defined.

Claim 5.3. Let $\tilde{s} \in FY(0)$, $\tilde{s} = \{(v_0, i_0), (v_1, i_1), ..., (v_{d_1}, i_{d_1})\}$ such that $\rho(\tilde{s}) = s$. Then there is a unique $j \in [\ell]$ such that $Z_s^j \subseteq \bigcap_{m=0}^{d_1} Z_{v_m}^{i_m}$.

We prove this claim below. Assuming it is true, let us show that this is actually an isomorphism.

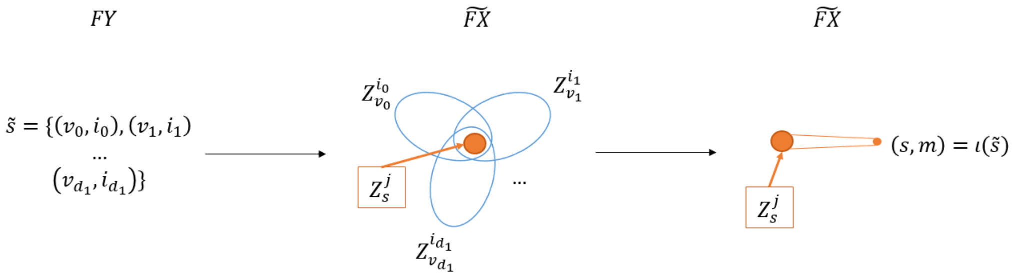

Figure 5: From $\tilde{s}$ to $\iota(\tilde{s}) = (s, m)$

Lemma 5.4. *$\iota : FY \to \widetilde{FX}$ is an isomorphism.*

Proof of Lemma 5.4. Let us start by explaining why this is a bijection. First note that for every $s \in FX(0)$, ι maps the ℓ-elements of $\rho^{-1}(s) \subseteq FY(0)$ to the ℓ-elements of $\nu^{-1}(s) \subseteq \widetilde{FX}$. Thus if we show that ι is injective, it will also follow that it is a bijection. Let $\tilde{s}_1, \tilde{s}_2 \in \rho^{-1}(s)$ be distinct elements and we denote $\iota(\tilde{s}_i) = (s, m_i)$. We'll show that $m_1 \neq m_2$. Take some vertex $v \in X(0)$ in $s \in FX(0)$, that is $v \in s = \rho(\tilde{s}_i)$. In particular there are two distinct vertices $(v, i_1), (v, i_2) \in Y(0)$ such that $(v, i_1) \in \tilde{s}_1$ and $(v, i_2) \in \tilde{s}_2$. Note that $\iota(\tilde{s}_i)$ is the vertex (s, m_i) that is connected to the $Z_s^{j_i}$ that sits in the intersection $\bigcap_{(u,k)\in\tilde{s}_i} Z_u^k$. In particular $Z_s^{j_1} \subseteq Z_v^{i_1}$ and $Z_s^{j_2} \subseteq Z_v^{i_2}$, thus $j_1 \neq j_2$, i.e. (s, m_1) and (s, m_2) are connected to different copies of the link of s. In particular this means that $m_1 \neq m_2$ and the map is injective.

It remains to prove that this is an isomorphism of simplicial complexes. First, we note that both FY and $\widetilde{FX}$ are clique complexes:

1. X is a clique complex, hence by Claim 2.23 its cover Y is also a clique complex. The face complex FY of a clique complex is also a clique complex.

2. FX is a clique complex since it is a face complex of a clique complex. By Claim 2.23 the cover $\widetilde{FX}$ is also a clique complex.

Thus it is enough to show that the bijection ι is a graph isomorphism between $FY^{\leqslant 1}$ to $(\widetilde{FX})^{\leqslant 1}$.

Let us begin by showing that ι is a graph homomorphism. Let $\{\tilde{s}_1, \tilde{s}_2\} \in FY(1)$ be an edge. Let $\iota(\tilde{s}_1) = (s_1, m_1), \iota(\tilde{s}_2) = (s_2, m_2)$. We need to show that $\{(s_1, j_1), (s_2, j_2)\} \in \widetilde{FX}(1)$. Since $\{\tilde{s}_1, \tilde{s}_2\} \in FY(1)$, then in particular they are contained in some triangle $\{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3\} \in FY(2)$($Y$ is pure so FY is also pure), and denote by $\nu(\tilde{s}_3) = s_3$. Let $Z_{s_1}^{j_1}, Z_{s_2}^{j_2}$ be the copies of FX_{s_1}, FX_{s_2} such that $(s_1, j_1), (s_2, j_2)$ are connected to $(s_1, m_1), (s_2, m_2)$ respectively (these copies are uniquely determined by the definition of ι). If we show that the same copy of s_3 is in the intersection of $Z_{s_1}^{j_1}, Z_{s_1}^{j_2} \subseteq \widetilde{FX}$, namely that there is some $(s_3, k) \in Z_{s_1}^{j_1} \cap Z_{s_2}^{j_2}$ this will imply that $\{(s_1, j_1), (s_2, j_2)\} \in \widetilde{FX}(1)$ (because this implies that both $(s_1, j_1), (s_2, j_2)$ are in the link of $(s_3, k) \in \widetilde{FX}$. This link is isomorphic by ν to the link of $s_3 \in FX$. There we know that $\{s_1, s_2\}$ is an edge, so in particular $\{(s_1, j_1), (s_2, j_2)\} \in \widetilde{FX}(1)$).

Take two vertices $(v_1, k_1), (v_2, k_2) \in Y$ such that $(v_i, k_i) \in \tilde{s}_i$. Because $\{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3\} \in FY(2)$ implies that $\tilde{s}_1 \mathbin{\dot\cup} \tilde{s}_2 \mathbin{\dot\cup} \tilde{s}_3 \in Y$. This means that

1. $\{(v_1, k_1), (v_2, k_2)\} \in Y(1)$, and that

2. $s_3 \in FX_{v_1 v_2}$.

By Lemma 5.2 the fact that $\{(v_1, k_1), (v_2, k_2)\} \in Y(1)$ implies that there is a copy $Z_{v_1,v_2}^p \subseteq Z_{v_1}^{k_1} \cap Z_{v_1}^{k_2}$. The fact that $s_3 \in FX_{v_1 v_2}$ implies that there is some $(s_3, k) \in Z_{v_1,v_2}^p$ and in particular $(s_3, k) \in Z_{v_1}^{k_1}, Z_{v_1}^{k_2}$. We now show that this (s_3, k) is the vertex we are looking for. Let $(s_3, k') \in Z_{s_1}^{j_1}$ be the copy of s_3 in $Z_{s_1}^{j_1}$. by definition of ι, it holds that $Z_{s_1}^{j_1} \subseteq Z_{v_1}^{k_1}$. As $Z_{v_1}^{k_1}$ is isomorphic by ν to FX_{v_1} and both $(s_3, k), (s_3, k') \in Z_{v_1}^{k_1}$ are sent to s_3 by ν, this implies that $(s_3, k) = (s_3, k')$. The same argument applies also to $Z_{s_2}^{j_2}$. We've proven there is an $(s_3, k) \in Z_{s_1}^{k_1} \cap Z_{s_3}^{k_2}$ and this shows that this is a grpah homomorphism.

To be convinced that this is also an isomorphism, note that both $\widetilde{FX}$ and FY are ℓ-covers of FX. In particular, the degree of every $\tilde{s} \in FY$ is equal to the degree of $\iota(\tilde{s}) \in \widetilde{FX}$ (both are equal to the degree of $\rho(s)$ in FX). Any bijective graph isomorphism that preserves degrees must be an isomorphism. $\square$

Proof of Claim 5.3. Let j be such that $Z_s^j \subseteq Z_{v0}^{i_0}$ (there is a unique such j by the second item of Claim 5.1). We need to show that $Z_s^j \subseteq Z_{v_m}^{i_m}$ for all $m = 1, 2, ..., d_1$. Indeed, fix m and let k be such that $Z_{v_0 v_m}^k \subseteq Z_{v0}^{i_0} \cap Z_{v_m}^{i_m}$ (there is such a k by Lemma 5.2). Then by the fact that $FX_s \subseteq FX_{v_0 v_m} \subseteq FX_{v_0}$ and $FX_s \subseteq FX_{v_0 v_m} \subseteq FX_{v_m}$, it holds by the isomorphism that ν induces that $Z_s^j \subseteq Z_{v_0 v_m}^k \subseteq Z_{v0}^{i_0}$, which in turn implies that $Z_s^j \subseteq Z_{v_0 v_m}^k \subseteq Z_{v_m}^{i_m}$. In particular $Z_s^j \subseteq Z_{v_m}^{i_m}$. This shows existence of Z_s^j. For uniqueness we note that there is a unique j such that $Z_s^j \subseteq Z_{v0}^{i_0}$, and $\bigcap_{m=0}^{d_1} Z_{v_m}^{i_m} \subseteq Z_{v0}^{i_0}$ so there must be only a single such j. $\square$

6 A global function on Y

Let us rephrase Corollary 3.6.

Lemma 6.1. *There exists a cocycle $\psi \in Z^1(X, Sym(\ell))$ such that $\mathbb{P}_{s_1,s_2,t=s_1 \cup s_2}\left[\psi(s_1,s_2) = \pi_{s_2,t}^{-1} \circ \pi_{s_1,t_1}\right] = 1 - \exp(-\Omega(d_1/k))$, where $\pi_{s_i,t}$ are the permutations over the list of functions of s_i and t respectively as in Lemma 3.4.*

Let us denote the ℓ-cover that this cocycle induces by $\rho : Y \to X$ such that we have an isomorphism $\iota : FY \to \widetilde{FX}$. Recall that we defined, in Section 3.4, functions $h_{\tilde{s}} : \tilde{s} \to \Sigma$ by

$$h_{\tilde{s}}((v,i)) = L_s^j(\rho(v,i)) = L_s^j(v) \tag{6.1}$$

where $\iota(\tilde{s}) = (s,j)$. Using these functions we defined $G : Y(0) \to \Sigma$ via $G(v,i) =$ *plurality* $\{h_{\tilde{s}}((v,i)) \mid (v,i) \in \tilde{s}\}$, i.e. the most popular assignment of (v,i) from all the $h_{\tilde{s}}$ where $\tilde{s} \ni (v,i)$.

The final component we need for proving our main theorem is that most $h_{\tilde{s}}$ agree with G on most vertices.

Claim (Restatement of Claim 3.8). There exists a universal constant $c > 0$ such that $\mathbb{P}_{\tilde{s} \in Y(d_1)}\left[h_{\tilde{s}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{s}}\right] \geqslant 1 - \exp(-\operatorname{poly}(\varepsilon)\left(\frac{d_1}{k}\right)^c)$.

For the proof of the lemma, we need the following agreement theorem taken from [DH23].

Theorem 6.2 ([DH23])**.** *There exists a universal constant $c > 0$ such that the following holds. Let Y be a $\lambda = 2^{-7d_1}$-two sided high dimensional expander, or 2^{-7d_1}-one sided high dimensional expander that is a skeleton of a q-partite complex for $q \geqslant d_1^2$. Let DU_n be the $(d_1, \frac{3}{2}d_1)$-non-lazy-up-down walk in Y as in Definition 2.5. Let $\mathcal{H} = \{h_{\tilde{s}}\}_{\tilde{s} \in Y(d_1)}$ be an ensemble of functions such that*

$$\mathbb{P}_{\tilde{s}_1,\tilde{s}_2 \sim DU_n}\left[h_{\tilde{s}_1}|_{\tilde{s}_1 \cap \tilde{s}_2} \overset{1-\eta}{\approx} h_{\tilde{s}_2}|_{\tilde{s}_1 \cap \tilde{s}_2}\right] \geqslant 1 - \gamma,$$

then

$$\mathbb{P}_{\tilde{s} \in Y(d_1)}\left[G|_{\tilde{s}} \overset{1-4(\eta+\gamma)}{\approx} h_{\tilde{s}}\right] \geqslant 1 - \exp(-\Omega(d_1^c)) - \gamma^c.$$

Proof of Claim 3.8. We want to use Theorem 6.2, and thus we need to show that

$$\mathbb{P}_{\tilde{s}_1,\tilde{s}_2 \sim DU_n}\left[h_{\tilde{s}_1}|_{\tilde{s}_1 \cap \tilde{s}_2} \overset{1-O(\eta)}{\approx} h_{\tilde{s}_2}|_{\tilde{s}_1 \cap \tilde{s}_2}\right] \geqslant 1 - \exp(-\Omega(d_1/k)),$$

Let us define the distribution $(\tilde{s}_1, \tilde{s}_2, \tilde{s}_3) \sim P$ where:

1. $\tilde{s_1}$ and $\tilde{s_2}$ are sampled via the non-lazy-up-down-walk. That is, $\tilde{s_1} \cup \tilde{s}_2 \in X(\frac{3}{2}d_1)$.
2. $\tilde{s}_3 \mathbin{\dot\cup} (\tilde{s_1} \cup \tilde{s}_2) \in X$, i.e. $\tilde{s}_3$ is sampled via a swap walk step from $(\tilde{s_1} \cup \tilde{s}_2)$.

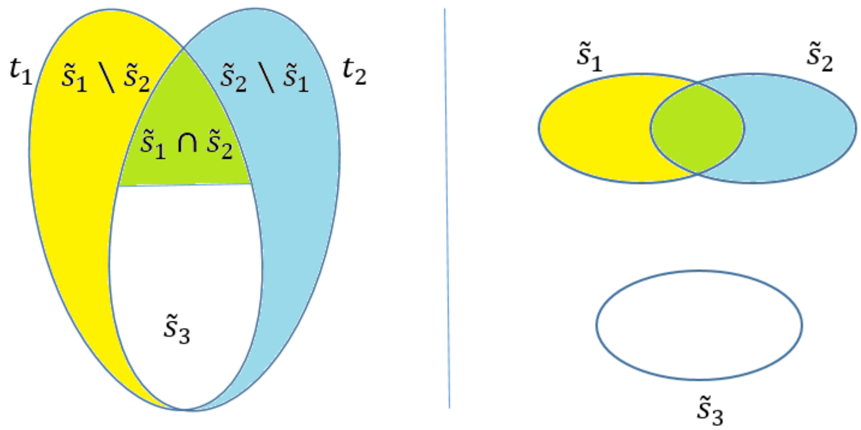

Figure 6: The sets sampled by P. Note that $\tilde{s_1}$ is the yellow and green parts, $\tilde{s_1}$ is the blue and green parts, and the $\tilde{t_i}$'s are the s_i's together with $\tilde{s_3}$ (the white part which is disjoint from $\tilde{s_1}$ and $\tilde{s_2}$).

Let us denote by $\tilde{t_i} = \tilde{s_i} \cup \tilde{s_3}$ for $i = 1, 2$, and let us denote $\tilde{r} = \tilde{s}_1 \cap \tilde{s}_2$. See Figure 6 for an illustration of the sets sampled. We also use the convention that $\rho(\tilde{s}_i) = s_i, \rho(\tilde{t_i}) = t_i, \rho(\tilde{r}) = r$ (hopefully, it will also be clear from context which faces belong to Y and which belong to X).

Let j_i be the index that has that $L_{s_i}^{j_i} \circ \rho = h_{\tilde{s_i}}$ for $i = 1, 2, 3.$ and let $j_i' = \pi_{s_i,t_i}(j_i)$. We define the following two "bad" events

1. $E = \{L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'} \text{ or } L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'} \text{ for } i = 1 \text{ or } i = 2\}$.
2. $F = \{L_{t_1}^{j_1'}|_{s_3} \overset{1-2\eta}{\approx} L_{t_2}^{j_2'}|_{s_3} \text{ and } L_{t_1}^{j_1'}|_r \overset{1-9\eta}{\not\approx} L_{t_2}^{j_2'}|_r\}$.

Let us spell out these events. The event E considers the lists of s_1, s_2, s_3 and t_1, t_2. The first statement in the event, i.e. that $L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}$, roughly states that at least one of the $i = 1, 2$, the lists of s_i and t_i are not compatible (at a specific index j_i). The second statement, namely that $L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}$ is the event that for at least one of the $i = 1, 2$, the entry of j_3 in s_3's list, is not compatible with $j_i' = \pi_{s_i,t}(j_i)$. We will see that with very high probability $j_3 = \pi_{s_3,t}(j_i')$ so this part also essentially says that the one of the lists of t_i and s_3 are not compatible with one another.

The event F says that "even though" $L_{t_1}^{j_1'}$ and $L_{t_2}^{j_2'}$ agree on most vertices of s_3, they do not agree on most vertices of r, the other part of their intersection.

We claim that both events happen with only small probability (and prove this later on in this section):

Claim 6.3. $\mathbb{P}[E] \leqslant \exp(-\Omega(d_1/k))$.

Claim 6.4. $\mathbb{P}[F] \leqslant \exp(-\Omega(d_1))$.

Let us explain why

$$\mathbb{P}\left[h_{\tilde{s_1}} \overset{1-13\eta}{\not\approx} h_{\tilde{s_2}}\right] \leqslant \mathbb{P}[E] + \mathbb{P}[F] = \exp(-\Omega(d_1/k)), \tag{6.2}$$

or in other words, why when $\neg E$ and $\neg F$ occur, then $h_{\tilde{s_1}} \overset{1-13\eta}{\approx} h_{\tilde{s_2}}$.

Indeed, if E doesn't occur then for $i = 1, 2$, $L_{s_i}^{j_i} \overset{1-\eta}{\approx} L_{t_i}^{j_i'}|_{s_i}$, and thus it holds that $L_{s_i}^{j_i}|_r \overset{1-2\eta}{\approx} L_{t_i}^{j_i'}|_r$ (since r is at $\frac{1}{2}$ the size of $\tilde{s_i}$ for both $i = 1, 2$). For similar reasons, $L_{t_1}^{j_1'}|_{s_3} \overset{1-\eta}{\approx} L_{s_3}^{j_3} \overset{1-\eta}{\approx} L_{t_2}^{j_2'}|_{s_3}$ and hence $L_{t_1}^{j_1'}|_{s_3} \overset{1-2\eta}{\approx} L_{t_2}^{j_2'}|_{s_3}$. By $\neg F$ this implies that $L_{t_1}^{j_1'}|_r \overset{1-9\eta}{\approx} L_{t_2}^{j_2'}|_r$. In total, it holds that

$$h_{\tilde{s_1}}|_{\tilde{r}} = (L_{s_1}^{j_1} \circ \rho)|_{\tilde{r}} \overset{1-2\eta}{\approx} (L_{t_1}^{j_1'} \circ \rho)|_{\tilde{r}} \overset{1-9\eta}{\approx} (L_{t_2}^{j_2'} \circ \rho)|_{\tilde{r}} \overset{1-2\eta}{\approx} L_{s_1}^{j_2'} \circ \rho|_{\tilde{r}} = h_{\tilde{s_2}}|_{\tilde{r}},$$

so $h_{\tilde{s_1}}|_{\tilde{r}} \overset{1-13\eta}{\approx} h_{\tilde{s_2}}|_{\tilde{r}}$. The claim follows from Theorem 6.2. □

Proof of Claim 6.3. Let us begin with bounding the probability that $L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j'_i}$ for some $i = 1, 2$. We note that s_i, t_i are chosen according to the usual joint distribution of $(s \subset t)$ in X. Thus the probability that $\mathbb{P}\left[L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j'_i}\right]$ for both $i = 1, 2$ is at most

$$1 - \mathop{\mathbb{P}}_{t \in X(\frac{3}{2}d_1), s \subset t, s \in X(d_1)} \left[\forall j = 1, 2, ..., \ell,\ L_s^j \overset{1-\eta}{\approx} L_t^{\pi_{s,t}(j)}|_s\right] \leqslant \exp(-\Omega(d_1/k))$$

by Lemma 3.4. When this occurs then $L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j'_i}$ for j_i and in particular for j_i and $\pi_{s_i,t_i}(j_i) = j'_i$.

We move towards bounding the probability that $L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{j'_i}$ for some $i = 1, 2$. The pair s_3, t_i is also distributed according to the joint distribution of $(s \subset t)$ in X. Hence it holds that $\mathbb{P}\left[L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{\pi_{s_3,t_i}(j_3)}\right] = \exp(-\Omega(d_1/k))$.

Moreover, by Lemma 6.1, with probability $1 - \exp(-\Omega(d_1/k))$ it holds that $\psi(s_i, s_3) = \pi_{\tilde{s_3},t_i}^{-1} \circ \pi_{\tilde{s_i},t_i}$ (where ψ is the cocycle that defined the cover ρ). Recall that $\iota(\tilde{s_i}) = (s_i, j_i)$ where $\iota : FY \to \widetilde{FX}$ and $\tilde{s_i}, \tilde{s_3} \in FY(0)$ are neighbors in FY which implies that for $i = 1, 2$, $(s_i, j_i) \sim (s_3, \psi(s_i, s_3).j_i)$ are neighbors in $\widetilde{FX}$. Thus with probability $1 - \exp(-\Omega(d_1/k))$,

$$j_3 = \pi_{s_3,t_i}^{-1} \circ \pi_{s_i,t_i}(j_i)$$

for $i = 1, 2$. When this occurs, then

$$\pi_{s_3,t_i}(j_3) = \pi_{s_i,t_i}(j_i) = j'_i,$$

and as above $\mathbb{P}\left[L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{j'_i}\right] \leqslant \exp(-\Omega(d_1/k))$.

□

The proof of Claim 6.4 is just by basic probabalistic arguments, and doesn't rely on any of the machinery we developed.

Proof of Claim 6.4. Fix any t_1, t_2, and denote $t' = t_1 \cap t_2$. Fix any $f = L_{t_1}^{j'_1}|_{t'}, g = L_{t_2}^{j'_2}|_{t'} : t' \to \Sigma$. The choice of $r, s_3 \subseteq t'$ is uniform among all pairs such that $r \cup s_3 = t'$. We will show that

$$\mathop{\mathbb{P}}_{\tilde{s_3}, r}\left[f|_{s_3} \overset{1-2\eta}{\approx} g|_{s_3} \text{ and } f|_r \overset{1-9\eta}{\not\approx} g|_r\right] \leqslant \exp(-\Omega(\eta^2|s|))$$

which proves the claim since $|s| = \Omega(d_1)$ and η is a constant.

Indeed, if $f|_r \overset{1-9\eta}{\not\approx} g|_r$, then $f \overset{1-3\eta}{\not\approx} g$ (since r is a third of the size of t'). In this case, the probability of choosing $s_3 \subseteq t'$ of size greater or equal to half of t' such that $L_{t_1}^{j'_1}|_{s_3} \overset{1-2\eta}{\approx} L_{t_2}^{j'_2}|_{s_3}$ is at most $\exp(-\Omega(\eta^2 d_1))$ by a standard Chernoff bound. □

7 The Counterexample (proof of Lemma 1.2)

In this section we elaborate on the counterexample we presented in the introduction. The setup is the following. X is a simplicial complex that admits a connected 2-cover $\rho : Y \to X$. We also assume that X is a $\lambda = 2^{-7k-1}$ two-sided high dimensional expander. Such complexes were constructed in [LSV05a] (in the referenced paper they only give a bound on one sided expansion, but see also [DK17] for obtaining two-sided expansion). Let $(r, r') \sim \mathcal{D}$ be any distribution over k-faces of X such that the marginal of $\mathcal{D}$ is the distribution over k-faces in X.

Lemma (Restatement of Lemma 1.2)**.** *Let $\delta < \frac{1}{2}$, $k \in \mathbb{N}$, and let $\lambda \leqslant \exp(-7k)$. Let X be a k-dimensional λ-two-sided high dimensional expander, and assume X has a connected 2-cover. Then there exists an ensemble of functions $\mathcal{F} = \{f_r : r \to \{0,1\} \mid r \in X(k)\}$ such that $Agree(\mathcal{F}) \geqslant \frac{1}{2}$, and yet for every $G : X(0) \to \{0,1\}$ it holds that*

$$\mathbb{P}_{r \in X(k)}\left[f_r \overset{1-\delta}{\approx} G|_r\right] \leqslant \exp(-\Omega_\delta(k)).$$

Proof of Lemma 1.2. Without loss of generality we denote $Y(0) = \{(v,0),(v,1) \mid v \in X(0)\}$, the covering map is $\rho(v,i) = v$.

By Claim 2.17 every face $r \in X(k)$ has exactly two preimages under ρ, which we denote $\tilde{r}_1, \tilde{r}_2 \in Y(k)$.

We will sample a random ensemble $\mathcal{F}$, and show that at least one random ensemble has agreement at least $\frac{1}{2}$ by showing that $\mathbb{E}_{\mathcal{F}}[\mathrm{Agree}_{\mathcal{D}}(\mathcal{F})] \geqslant \frac{1}{2}$.

The randomized ensemble is constructed as follows. For every $r \in X(k)$ we select one of the two preimage $\tilde{r} \in Y(k)$ of r, independently and uniformly at random. Then for every $v \in r$,

$$f_r(v) = \begin{cases} 0 & (v,0) \in \tilde{r} \\ 1 & (v,1) \in \tilde{r} \end{cases}.$$

In the introduction we stated that f_r comes from a function $H|_{\tilde{r}}$. This function is $H : Y(0) \to \{0,1\}$, $H(v,i) = i$. Let us verify that $\mathbb{E}_{\mathcal{F}}[\mathrm{Agree}_{\mathcal{D}}(\mathcal{F})] \geqslant \frac{1}{2}$. The expectation is over the choice of preimages $\tilde{r}$ for every $r \in X(k)$.

$$\mathbb{E}_{\mathcal{F}}[\mathrm{Agree}_{\mathcal{D}}(\mathcal{F})] = \mathbb{E}_{\mathcal{F}}\left[\mathbb{E}_{r,r' \sim D}\left[\mathbf{1}_{f_r = f_{r'}}\right]\right] = \mathbb{E}_{r,r' \sim D}\left[\mathbb{E}_{\mathcal{F}}\left[\mathbf{1}_{f_r = f_{r'}}\right]\right].$$

Hence it is enough to show that for every $r, r' \sim D$, $\mathbb{E}_{\mathcal{F}}\left[\mathbf{1}_{f_r = f_{r'}}\right] \geqslant \frac{1}{2}$, or equivalently that $\mathbb{P}_{\mathcal{F}}[f_r = f_{r'}] \geqslant \frac{1}{2}$.

Fix $r, r' \sim D$ an edge in the agreement distribution. If r, r' have an empty intersection or $r = r'$ this is trivial, so let us assume this is not the case. Let $\tilde{r}, \tilde{r}$ be the preimages of r, r' randomly chosen, respectively. Let $v \in r \cap r'$ and suppose without loss of generality that $(v,0) \in \tilde{r}$ be the preimage of v inside the preimage of r. As $\tilde{r}$ is chosen independently, with probability $\frac{1}{2}$ it holds that $(v,0) \in \tilde{r}$. In this case, because $\rho|_{X_{(v,0)}} : X_{(v,0)}(0) \to X_v(0)$ is an isomorphism it holds that for every $u \in r \cap r'$, there is a preimage (u,i) such that $(u,i) \in \tilde{r} \cap \tilde{r}'$. In particular $f_r(u) = f_{r'}(u)$ by definition. Thus indeed $\mathbb{P}_{\mathcal{F}}[f_r = f_{r'}] \geqslant \frac{1}{2}$.

Thus we take $\mathcal{F}$ be any such assignment where $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}) \geqslant \frac{1}{2}$ (which exists since the expectation over the agreement is at least $\frac{1}{2}$). The first item in the lemma holds for this ensemble.

We conclude by showing that for *any* ensemble constructed as above, and any $G : X(0) \to \{0,1\}$ it holds that

$$\mathbb{P}_{r \in X(k)}\left[f_r \overset{1-\delta}{\approx} G|_r\right] = \exp(-\Omega_\delta(k)).$$

Fix a global function $G : X(0) \to \{0,1\}$. Recall that $H : Y \to \{0,1\}$ is $H(v,i) = i$. Let $\tilde{G} : Y(0) \to \{0,1\}$ be $\tilde{G}((v,i)) = G(v)$. Note that if $G|_r \overset{1-\delta}{\approx} f_r$ and f_r answers according to the preimage $\tilde{r}$, then $H|_{\tilde{r}} \overset{1-\delta}{\approx} \tilde{G}|_{\tilde{r}}$. This is because if $G(v) = f_r(v) = j$ and f_r answers according to $\tilde{r}$, this means that $(v,j) \in \tilde{r}$, and hence $\tilde{G}(v,j) = G(v) = f_r(v) = H(v,j)$. We conclude that

$$\mathbb{P}_{r \in X(k)}\left[G|_r \overset{1-\delta}{\approx} f_r\right] \leqslant 2 \mathbb{P}_{\tilde{r} \in Y(k)}\left[\tilde{G}|_{\tilde{r}} \overset{1-\delta}{\approx} H|_{\tilde{r}}\right], \tag{7.1}$$

since the choice of $\tilde{r} \in Y(k)$ is done by first chosing $r \in X(k)$ and then chosing one preimage uniformly at random. Thus we will argue that $\mathbb{P}_{\tilde{r} \in Y(k)}\left[\tilde{G}|_{\tilde{r}} \overset{1-\delta}{\approx} H|_{\tilde{r}}\right] = \exp(-\Omega_\delta(k))$.

We observe that $\mathrm{dist}(H, \tilde{G}) = \frac{1}{2}$ because for every $v \in X(0)$, $H, \tilde{G}$ agree on exactly one of $(v,0),(v,1)$. By Claim 2.24 it holds that Y is also a 2^{-7k}-two sided spectral expander.

Let $A = \left\{(v,i) \in Y(0) \mid H(v,i) \neq \tilde{G}(v,i)\right\}$. Then $\mathbb{P}[A] = \mathrm{dist}(H, \tilde{G}) = \frac{1}{2}$, and a face $\tilde{r}$ is such that $\tilde{G}|_{\tilde{r}} \overset{1-\delta}{\approx} H|_{\tilde{r}}$ if and only if $\mathbb{P}_{(v,i) \in Y(0)}[A \mid (v,i) \in \tilde{r}] < \delta$.

Thus by Theorem 2.32, it holds that

$$\mathop{\mathbb{P}}_{\tilde{r}\in Y(k)}\left[\tilde{G}|_{\tilde{r}} \overset{1-\delta}{\approx} H|_{\tilde{s}}\right] = \exp(-\Omega(\delta^2 k))$$

and by (7.1),

$$\mathop{\mathbb{P}}_{r\in X(k)}\left[G|_r \overset{1-\delta}{\approx} g_r\right] = \exp(-\Omega_\delta(k)).$$

□

Remark 7.1. A similar argument can generalized for complexes X with connected ℓ-covers and an alphabet of $\Sigma = [\ell]$. We can also consider test distributions that samples q k-sets instead of 2, provided that every $\{r_i\} \subseteq \operatorname{supp} \mathcal{D}$ are contained in a simply connected sub complex of X.

In this case we can show that there exists an ensemble with agreement at least $\frac{1}{\ell^{q-1}}$, but

$$\mathop{\mathbb{P}}_{r\in X(k)}\left[G|_r \overset{1-\delta}{\approx} g_r\right] = \exp(-\Omega_\delta(k))$$

will hold for any $G : X(0) \to \Sigma$ and any $\delta < \frac{\ell-1}{\ell}$.

References

[Sur84] David Surowski. "Covers of simplicial complexes and applications to geometry." In: *Geometriae Dedicata* 16 (Apr. 1984), pp. 35–62. DOI: 10.1007/BF00147420.

[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.

[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.

[RS97] R. Raz and S. Safra. "A Sub-Constant Error-Probability Low-Degree Test, and a Sub-Constant Error-Probability PCP Characterization of NP." In: *Proc. 28th ACM Symp. on Theory of Computing*. 1997, pp. 475–484.

[Raz98] Ran Raz. "A Parallel Repetition Theorem." In: *SIAM Journal on Computing* 27.3 (June 1998), pp. 763–803.

[Bog05] Anrej Bogdanov. "Gap Amplification Fails Below 1/2." Comment on ECCC TR05-046, can be found at `http://eccc.uni-trier.de/eccc-reports/2005/TR05-046/commt01.pdf`. 2005.

[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.

[DR06] Irit Dinur and Omer Reingold. "Assignment Testers: Towards a Combinatorial Proof of the PCP Theorem." In: *SIAM J. Comput.* 36.4 (2006), pp. 975–1024. DOI: 10.1137/S0097539705446962. URL: https://doi.org/10.1137/S0097539705446962.

[DG08] Irit Dinur and Elazar Goldenberg. "Locally testing direct products in the high error range." In: FOCS. 2008.

[Imp+08] Russel Impagliazzo, Ragesh Jaiswal, Valentine Kabanets, and Avi Wigderson. "Uniform Direct Product Theorems: Simplified, Optimized, and Derandomized." In: *Proc. 40th ACM Symp. on Theory of Computing*. 2008, pp. 579–588.

[MR10] Dana Moshkovitz and Ran Raz. “Two-query PCP with subconstant error.” In: *J. ACM* 57.5 (2010).

[DM11] Irit Dinur and Or Meir. “Derandomized Parallel Repetition via Structured PCPs.” In: *Comput. Complex.* 20.2 (2011), pp. 207–327. DOI: 10.1007/s00037-011-0013-5. URL: https://doi.org/10.1007/s00037-011-0013-5.

[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.

[DS14a] Irit Dinur and David Steurer. “Analytical approach to parallel repetition.” In: *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*. 2014, pp. 624–633.

[DS14b] Irit Dinur and David Steurer. “Direct Product Testing.” In: *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*. IEEE Computer Society, 2014, pp. 188–196. DOI: 10.1109/CCC.2014.27. URL: https://doi.org/10.1109/CCC.2014.27.

[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.

[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)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2017.

[Dik+18] Yotam Dikstein, Irit Dinur, Yuval Filmus, and Prahladh Harsha. “Boolean function analysis on high-dimensional expanders.” In: *Proc.* 20*th International Workshop on Randomization and Computation (RANDOM)*. Vol. 116. Princeton, NJ: RANDOM/APPROX, 2018, 37:1–37:21. DOI: 10.4230/LIPIcs.APPROX/RANDOM.2018.37. eprint: 1804.08155.

[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. “Towards a proof of the 2-to-1 games conjecture?” In: *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*. Ed. by Ilias Diakonikolas, David Kempe, and Monika Henzinger. ACM, 2018, pp. 376–389. DOI: 10.1145/3188745.3188804. URL: https://doi.org/10.1145/3188745.3188804.

[KMS18] Subhash Khot, Dor Minzer, and Muli Safra. “Pseudorandom Sets in Grassmann Graph Have Near-Perfect Expansion.” In: *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*. Ed. by Mikkel Thorup. IEEE Computer Society, 2018, pp. 592–601. DOI: 10.1109/FOCS.2018.00062. URL: https://doi.org/10.1109/FOCS.2018.00062.

[Opp18] Izhar Oppenheim. “Local Spectral Expansion Approach to High Dimensional Expanders Part I: Descent of Spectral Gaps.” In: *Discrete Comput. Geom.* 59.2 (2018), pp. 293–330. DOI: 10.1007/s00454-017-9948-x. eprint: 1709.04431.

[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.

[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.

[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." In: *ACM-SIAM Symposium on Discrete Algorithms (SODA19)*. 2019.

[GK20] Elazar Goldenberg and Karthik C. S. "Toward a General Direct Product Testing Theorem." In: *ACM Trans. Comput. Theory* 12.1 (2020), 7:1–7:18. DOI: 10.1145/3369939. URL: https://doi.org/10.1145/3369939.

[KO20] Tali Kaufman and Izhar Oppenheim. "High order random walks: Beyond spectral gap." In: *Combinatorica* 40 (2020), pp. 245–281.

[DM22] Irit Dinur and Roy Meshulam. "Near coverings and cosystolic expansion." In: *Archiv der Mathematik* 118.5 (May 2022), pp. 549–561. ISSN: 1420-8938. DOI: 10.1007/s00013-022-01720-6. URL: https://doi.org/10.1007/s00013-022-01720-6.

[GK22] Roy Gotlib and Tali Kaufman. *List Agreement Expansion from Coboundary Expansion*. 2022. eprint: arXiv:2210.15714.

[BM23] Mitali Bafna and Dor Minzer. "Characterizing Direct Product Testing via Coboundary Expansion." 2023.

[DD23a] Yotam Dikstein and Irit Dinur. "Coboundary and cosystolic expansion without dependence on dimension or degree." In: *CoRR* abs/2304.01608 (2023). DOI: 10.48550/arXiv.2304.01608. arXiv: 2304.01608. URL: https://doi.org/10.48550/arXiv.2304.01608.

[DD23b] Yotam Dikstein and Irit Dinur. "Expansion of high dimensional face complexes." Unpublished manuscript. 2023.

[DH23] Yotam Dikstein and Max Hopkins. "Reverse hypercontractivity and Chernoff bounds for high dimensional expanders." Unpublished manuscript. 2023.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il