ECCC

# Agreement theorems for high dimensional expanders in the low acceptance regime: the role of covers

Yotam Dikstein* and Irit Dinur†

April 12, 2024

## Abstract

Let $X$ be a family of $k$-element subsets of $[n]$ and let $\{f_s : s \to \Sigma \mid s \in X\}$ be an ensemble of local functions, each defined over a subset $s \subset [n]$. Is there a global function $G : [n] \to \Sigma$ such that $f_s = G|_s$ for all $s \in X$ ? An agreement test is a randomized property tester for this question. One such test is the V-test, that chooses a random pair of sets $s_1, s_2 \in X$ with prescribed intersection size and accepts if $f_{s_1}, f_{s_2}$ agree on the elements in $s_1 \cap s_2$.

The low acceptance (or 1%) regime is concerned with the situation that the test succeeds with low but non-negligible probability $\mathrm{Agree}(\{f_s\}) \geqslant \varepsilon > 0$. A "classical" low acceptance agreement theorem says

$$\mathrm{Agree}(\{f_s\}) > \varepsilon \quad \Longrightarrow \quad \exists G : [n] \to \Sigma, \quad \mathbb{P}_s[f_s \overset{0.99}{\approx} G|_s] \geqslant \mathrm{poly}(\varepsilon). \tag{$*$}$$

Such statements are motivated by PCP questions. The case $X = \binom{[n]}{k}$ is well-studied and known as "direct product testing", which is related to the parallel repetition theorem. Finding sparser families $X$ that satisfy ($*$) is known as derandomized direct product testing. Prior to this work, the sparsest family satisfying ($*$) had $|X| \approx n^{25}$, and we show $X$ with $|X| \approx n^2$.

We study the general behavior of high dimensional expanders with respect to agreement tests in the low acceptance regime. High dimensional expanders, even very sparse ones with $|X| = O(n)$, are known to satisfy the high acceptance variant (where $\varepsilon = 1 - o(1)$). It has been an open challenge to analyze the low acceptance regime. Surprisingly, topological covers of $X$ play an important role. We show that:

1. If $X$ has no connected covers, then ($*$) holds, provided that $X$ satisfies an additional expansion property, called swap cosystolic expansion.

2. If $X$ *has* a connected cover, then ($*$) fails.

3. If $X$ has a connected cover (and swap-cosystolic-expansion), we replace ($*$) by a statement that takes covers into account:

$$\mathrm{Agree}(\{f_s\}) > \varepsilon \Longrightarrow \quad \exists \text{ cover } \rho : Y \twoheadrightarrow X, \text{ and } G : Y(0) \to \Sigma, \text{ such that} \tag{$**$}$$

$$\mathbb{P}_{\tilde{s} \twoheadrightarrow s}[f_s \overset{0.99}{\approx} G|_{\tilde{s}}] \geqslant \mathrm{poly}(\varepsilon),$$

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

The property of swap-cosystolic-expansion holds for quotients of the Bruhat Tits buildings. As a corollary we derive ($*$) for $X$ being a spherical building, yielding a derandomized family with $|X| \approx n^2$. We also derive ($**$) for LSV complexes $X$, for which $|X| = O(n)$.

# Contents

# 1   Introduction

Let $G : [n] \to \Sigma$ be a function. We often specify $G$ as a length $n$ string over an alphabet $\Sigma$. Alternatively, $G$ can be specified 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 a table with $|X|$ entries, where the $s$-th entry specifies the local function $G|_s$. Formally, we have 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 $f_{s_1}(v) = f_{s_2}(v)$ for all $v \in s_1 \cap s_2$. 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%" or high acceptance regime which is natural in the world of property testing, and the "1%" or low acceptance regime which is the regime of interest in the world of PCPs.

– *The "99%"* ***High A****cceptance regime.* We are given an ensemble $\{f_s\}_{s \in X}$ that passes the agreement test with probability $1 - \varepsilon$, for small $\varepsilon > 0$. We want to conclude that there is a function $G : [n] \to \Sigma$ 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

$$\mathrm{Agree}(\{f_s\}) > 1 - \varepsilon \quad \implies \quad \exists G : [n] \to \Sigma, \quad \mathbb{P}\left[f_s = G|_s\right] \geqslant 1 - O(\varepsilon). \qquad (HA)$$

– *The "1%"* ***Low A****cceptance regime.* Here we assume that the test succeeds with much smaller probability $\varepsilon > 0$ for a small positive constant $\varepsilon$, say $\varepsilon = 0.01$, or even $\varepsilon = o_k(1)$. The hope is that even this modest success probability 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 $\varepsilon$ translates to a larger gap between completeness and soundness of the PCP. We say that $X$ satisfies ($LA$) if for any ensemble $\{f_s\}_{s \in X}$, ($LA$) holds, namely, for every ensemble $\{f_s\}_{s \in X}$,

$$\mathrm{Agree}(\{f_s\}) > \varepsilon \quad \implies \quad \exists G : [n] \to \Sigma, \quad \mathbb{P}_s[f_s \overset{0.99}{\approx} G|_s] \geqslant \mathrm{poly}(\varepsilon). \qquad (LA)$$

In words, ($LA$) says that $\varepsilon$ agreement implies global structure. Of course, to be completely formal we must equip ($LA$) with parameters. This is omitted in the introduction and is done in Definition 2.10.

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

Proving that ($LA$) holds is quite non-trivial even when $X = \binom{[n]}{k}$, where we think of $k$ as a sufficiently large constant while $n \to \infty$. This is known as the (symmetrized version of the) "direct product testing question". 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 ($LA$). Moreover, the tradeoffs between $k, \varepsilon$ and $1 - \zeta$,

---

[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 we use are precisely defined in Section 2.2.

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

as well as the significant difference between two and three queries, have all been relatively well studied, see further discussion below.

Goldreich and Safra [GS97] and later Impagliazzo et. al [IKW12] studied derandomized direct products. Namely, finding smaller families $X \subset \binom{[n]}{k}$, such that $|X| \ll n^k$ yet $X$ still satisfies ($HA$) or ($LA$). The motivation for this question is that the number of subsets in $X$ controls the efficiency of the encoding. Impagliazzo et. al [IKW12] showed that for every $\varepsilon > 0$ there exists $k$ and a sequence of families $\{X_n\}$ that satisfy ($LA$), where $|X_n| \leqslant O(n^c)$ for $c \geqslant 25$ an absolute constant independent of $k, \varepsilon$. It is natural to ask about the limits of such derandomization. Can the same hold true when $|X_n| = O(n)$?

The recent "discovery" of bounded-degree high dimensional expanders by the theoretical computer science community has lead to a hope that these might provide a path towards such results. 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 ($HA$). This gave the first sequence $\{X_n\}$ of linear-size families satisfying ($HA$).

In this paper we address the same question for ($LA$). Are there families $X_n$ that satisfy ($LA$) and whose size is linear in $n$? This has been mentioned as an open question in [DK17]. We conjecture the following:

**Conjecture 1.1.** *For every $\varepsilon > 0$ there exists an integer $k(\varepsilon)$, such that for all $k > k(\varepsilon)$, there is a sequence $\{X_n\}_{n \to \infty}$ of families of $k$-element subsets of $n$ that satisfy ($LA$) and such that $|X_n| \leqslant O_{k,\varepsilon}(n)$.*

It turns out that as far as ($LA$) is concerned, the story is more complicated. Our first result is that spectral high dimensional expansion is not enough. We show that many complexes $X$ that satisfy ($HA$), fail to satisfy ($LA$) in quite a strong way. In fact, as long as a complex has a connected topological cover, it fails to satisfy ($LA$). We remark that many known sparse high dimensional expanders fall under this category as the constructions of both [LSV05b] and [KO18] give a sequence of high dimensional expanders $\{X^{(i)}\}_{i \to \infty}$ where $X^{(i)}$ is covered by $X^{(i+1)}$.

**Lemma 1.2.** *Let $\zeta < \frac{1}{2}$, $k \in \mathbb{N}$, and let $\lambda \leqslant \frac{0.99}{k}$. Let $X$ be a $k$-dimensional $\lambda$-two-sided high dimensional expander, and assume $X$ has a connected $\ell$-cover for some $\ell > 1$ (e.g. $\ell = 2$). Then there exists an ensemble of functions $\mathcal{F} = \{f_r : r \to \{0,1\} \mid r \in X(k)\}$ such that $\text{Agree}(\mathcal{F}) \geqslant \frac{1}{\ell}$, and yet for every $G : X(0) \to \{0,1\}$ it holds that*

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

(We are using standard notation where $X(0)$ are the vertices of $X$, and $X(k)$ are the $k$-faces of $X$). This lemma reveals connected 2-covers (and more generally connected $\ell$-covers) to be an obstruction to 1% agreement theorems. The notion of a connected cover is a basic notion in topology, generalizing the more familiar lift of a graph to higher dimensions. In a nutshell, a complex $Y$ is a $\ell$-cover of a complex $X$ whenever there is a $\ell$-to-1 homomorphism $\rho : Y \twoheadrightarrow X$ such that the preimage of every face of $X$ is $\ell$ pairwise disjoint faces in $Y$, (see Section 2 for details).

Lemma 1.2 shows that high dimensional expanders $X$ that have connected covers necessarily fail to satisfy ($LA$). In contrast, previous works [DG08; IKW12; DS14b; DL17] show that the complete complex $X = \binom{[n]}{k}$, which is known to be a superb high dimensional expander, does satisfy ($LA$). This seeming contradiction is resolved by observing that the complete complex *has no connected covers*. This suggests that sparser high dimensional expanders that are cover-free might satisfy ($LA$), as we show further below in Corollary 1.6.

Our main result is concerned with the more general case. We introduce a new variant of (LA) that *is* satisfied by high dimensional expanders, as long as we extend the definition of a global function to allow covers. We essentially show that the obstruction uncovered in Lemma 1.2 is the only possible obstruction. Our theorem has an additional condition, however, which is that $X$ is also required to have swap-coboundary-expansion (see Definition 2.33). In a companion paper we prove that many familiar complexes, including spherical buildings as well as the Ramanujan complexes of [LSV05a], have this property (see Theorem 1.4 below).

**Theorem 1.3** (Main Theorem, see Theorem 3.1 for a formal statement)**.** *Let $k \in \mathbb{N}$, and let $\varepsilon > \Omega(1/\log k)$. Let $d > k$ be sufficiently large and let $X$ be a $d$-dimensional high dimensional expander with sufficiently good*

*swap-cosystolic-expansion.* Let $\{f_s : s \to \Sigma \mid s \in X(k)\}$ be an ensemble of local functions on $X(k)$.

$$Agree(\{f_s\}) > \varepsilon \quad \Longrightarrow \quad \exists Y \xrightarrow{\rho} X, \exists G : Y(0) \to \Sigma, \quad \mathbb{P}_s[f_s \text{ is explained by } G] \geqslant \mathrm{poly}(\varepsilon). \qquad (CLA)$$

*where $\rho : Y \to X$ is a $\ell = \mathrm{poly}(1/\varepsilon)$ covering map.*

Here the phrase "$f_s$ is explained by $G$" informally means that $G|_{s'} \approx f_s$ for some $s' \in Y$ that covers $s$. More formally, the covering map $\rho$ gives a bijection from $s' \in Y(k)$ to $s \in X(k)$, and by $G|_{s'} \approx f_s$ we mean that for (almost) every $v \in s'$, $G(v) = f_s(\rho(v))$.

This theorem should be viewed as a new type of structure theorem, replacing ($LA$) with its cover version, ($CLA$): Every $\varepsilon$-agreeing ensemble on $X$ comes from a global function $G$, perhaps not on the vertices of $X$, but on the vertices of a cover $Y$ of $X$, which is not much larger than $X$.

One might wonder whether any progress has been made by replacing the quest for complexes that satisfy 1% agreement theorems with complexes having swap-coboundary-expansion. In a companion paper [DD23b] we show that there are important families of complexes that have swap-coboundary-expansion.

**Theorem 1.4** (Main theorem in [DD23b]). *Let $X$ be a $d$-dimensional simplicial whose links are spherical buildings for a sufficiently large prime $q$. Then for every $d_1 \leqslant d$, $X$ is a $\exp(-\sqrt{d_1})$-swap cosystolic expander.*

The bounded-degree high dimensional expanders of [LSV05b; LSV05a] satisfy the requirement of Theorem 1.4 and therefore also of Theorem 1.3. Thus, we get the following lift-decoding theorem.

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

*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 $G : Y(0) \to \Sigma$, such that*

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

As a special case of this corollary, we get a new 1% agreement theorem for spherical buildings. These are relatively sparse simplicial complexes that are known to have no connected covers. The only possible $\ell$-cover of such a complex $X$ consists of $\ell$ 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 $\ell$ global functions on $X(0)$.

**Corollary 1.6** (Agreement for spherical buildings). *Let $\varepsilon > 0$, and let $k > \exp(\mathrm{poly}(1/\varepsilon))$ be an integer. Let $d > k$ be sufficiently large and let $X$ be a $d$-dimensional spherical building that is a $\frac{1}{d^2}$ high dimensional expander. Then $X$ satisfies ($LA$).*

This puts spherical buildings as the current best known derandomized direct product encoding that satisfies ($LA$), improving the size of the family from $|X_n| \leqslant O(n^{25})$ in [IKW12] to $|X_n| = O(n^2)$.

Finally, we revisit Conjecture 1.1. There is still some hope for a positive answer via high dimensional expanders. Even when $X$ has connected covers, there is a quantitative aspect connecting the agreement parameter $\varepsilon$ and the size $\ell$ of the cover, which may allow us to move from ($CLA$) to ($LA$). I.e., if $X$ has the additional property that all of its $\ell$-covers for $\ell < \mathrm{poly}(1/\varepsilon)$ are necessarily disconnected, then we immediately get conclusion a la ($LA$). This motivates the study and construction of complexes that are free of connected $\ell$-covers for *small* $\ell$. A construction of such a family of complexes would give a linear-size family of complexes supporting ($LA$).

## Covers and Agreement

Let us describe how topological covers give rise to the counter-example driving Lemma 1.2.

1. Let $X$ be a simplicial complex, and let $Y \xrightarrow{\rho} X$ be a $\ell$-cover with $\ell = 2$. Explicitly, $Y$ is a simplicial complex and we have a 2-to-1 map $\rho : Y(0) \to X(0)$ such that the image of each $Y$-face is a $X$-face, and the preimage of each $X$-face is two disjoint $Y$-faces.

2. Fix a global function $H : Y(0) \to \{0, 1\}$ with the property that every pair of vertices $(v, 0), (v, 1) \in Y(0)$ that map to the same $v \in X(0)$, are given distinct values $H(v, 0) \neq H(v, 1)$.
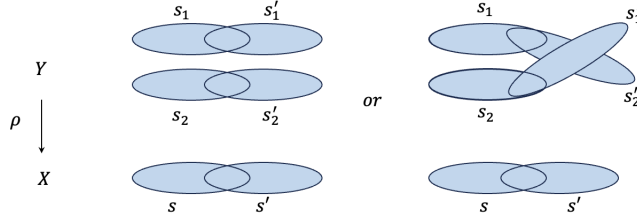
Figure 1: Possible ways two intersecting sets $s, s' \in X$ are covered by $Y$.

3. Every face $s \in X(k)$ is covered by two faces $s_1, s_2 \in Y(k)$, i.e. $\rho(s_i) = s$. Thus, we naturally get two possible assignments to $f_s$ that biject down from either $H|_{s_1}$ or $H|_{s_2}$. By construction, these two assignments are negations of one another.

   Moreover, suppose, as in Figure 1, that $s, s' \in X(k)$ intersect, and let $s_1, s_2 \in Y(k)$ cover $s$ and $s'_1, s'_2 \in Y(k)$ cover $s'$. Then $s_1$ intersects exactly one of $s'_1, s'_2$, and similarly for $s_2$. Namely, the intersection pattern gives a bijection between $\{s_1, s_2\}$ and $\{s'_1, s'_2\}$.

4. Our final ensemble for $X$ is obtained by choosing, for each $s$ independently at random, one of the two possible assignments. For any fixed pair of intersecting sets $s, s' \in X(k)$, the probability that $f_s$ agrees with $f_{s'}$ is exactly $1/2$. Thus, in total, the average agreement probability of the entire ensemble $\{f_s\}$ is about $1/2$. Moreover, sampling arguments (relying on the high dimensional expansion of $X, Y$) will show that no global function $G : X(0) \to \{0, 1\}$ can agree with more than $\exp(-\Omega(k))$ of the subsets.

What this construction shows is that whenever a complex has a connected cover, it cannot satisfy ($LA$), but it could still possibly satisfy ($CLA$). More details can be found in Section 7.1.

## Proof of main theorem - Overview

We next turn to describing the technical ingredients that go into the proof of our main theorem, Theorem 1.3, which essentially reverses the construction described just above. 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.

### Step 1 - Constructing a local list for every face

Fix $k \ll d_1 \ll d$. Here we treat every $s \in X(d_1)$ as a ground set whose set of $k$-element subsets form a copy of the complete complex on $d_1 + 1$ vertices. 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 $\mathrm{Agree}(\{f_s\}) > \varepsilon$, 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 $\varepsilon$ for nearly all sets $s \in X(d_1)$. So for each of these local ensembles, we can use the previous agreement theorem from [IKW12]. We obtain a global function $g_s$ for each $s$ that agrees with about $\varepsilon$ of the local functions in $\{f_r\}_{r \subseteq s}$.

An important complication is that there could be several equally valid choices for $g_s$, each supported on a different $\varepsilon$ fraction of the local functions $\{f_r\}_{r \subset s}$. Therefore, we must take into account an entire **list** of functions for $s$, which we denote by $\bar{L}(s)$.

Similarly to all $s \in X(d_1)$, we find such lists for faces $t \in X(2d_1 + 1)$ and $u \in X(3d_1 + 2)$.

4

**Step 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 $\ell$ so that $|\overline{L}(s)| = \ell$ for nearly all $s$. Moreover, denoting $\overline{L}(s) = \{L_s^1, L_s^2, \ldots, L_s^\ell\}$, for most pairs $s \subseteq t$, 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.

**Step 3 - Unique games on the *faces* complex**

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

**Definition 1.7** (The Faces Complex)**.** Let $X$ be a $d$-dimensional simplicial complex. Let $d_1 \leqslant d$. We denote by $\mathrm{F}^{d_1}X$ the simplicial complex whose vertices are $\mathrm{F}^{d_1}X(0) = X(d_1)$ and whose faces are all $\left\{ \{s_0, s_1, ..., s_j\} \,\middle|\, s_0 \uplus s_1 \uplus \cdots \uplus s_j \in X \right\}$.

The 1-skeleton of $\mathrm{F}^{d_1}X$ is a graph whose vertices are the $d_1$-sets. Each vertex $s_1 \in \mathrm{F}^{d_1}X$ 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 $\psi(s_1, s_2) = \pi_{t,s_2}^{-1} \circ \pi_{t,s_1}$ where $t = s_1 \uplus s_2$. This can be viewed as a unique games instance over $\mathrm{F}^{d_1}X$. We show that for most triangles $s_1 s_2 s_3 \in \mathrm{F}^{d_1}X$, there are no local contradictions. Namely, if we denote by $\psi_{s_i, s_j}$ the permutation from the list of $s_i$ to the list of $s_j$, then

$$\psi(s_2, s_3) \circ \psi(s_1, s_2) = \psi(s_1, s_3) \tag{1.1}$$

nearly always. In topological language this collection of bijections gives us a 1-cochain in $C^1(\mathrm{F}^{d_1}, Sym(\ell))$ where $Sym(\ell)$ is the groups of permutations over $\ell$ elements. The consistency implies that this 1-cochain is nearly a cocycle.

**Step 4 - Swap cosystolic expansion**

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 (and this is precisely our definition of swap cosystolic expansion, see Definition 2.33).

It remains an open question to characterize complexes with sufficiently strong swap cosystolic expansion. We don't know if cosystolic expansion implies swap cosystolic expansion. Nevertheless, as stated above, we prove in a companion paper [DD23b] that every high dimensional expander whose links are spherical buildings, has $\mathrm{F}^{d_1}X$ has swap-cosystolic expansion (Theorem 1.4).

**Step 5 - Constructing the cover**

It is well known that any collection of permutations $\{\pi_{uv}\}_{uv \in \mathrm{F}^{d_1}X(1)}$ with *no* local contradictions[3] corresponds to a cover $\rho_{\mathrm{F}^{d_1}X} : \widetilde{\mathrm{F}^{d_1}X} \to \mathrm{F}^{d_1}X$ (see e.g. [Sur84]). Let us describe the underlying graph of the cover. Its vertices are pairs $(s, i)$ for every $s \in \mathrm{F}^{d_1}X(0)$ and $i \in [\ell]$. We connect $(s_1, i)$ to $(s_2, j)$ by an edge in $\widetilde{\mathrm{F}^{d_1}X}$ if $s_1 s_2 \in \mathrm{F}^{d_1}X(1)$ and $\pi_{s_1, s_2}(i) = j$. The fact that (1.1) holds for every triangle is the reason that this underlying graph is indeed a graph skeleton of a cover $\rho_{\mathrm{F}^{d_1}X} : \widetilde{\mathrm{F}^{d_1}X} \to \mathrm{F}^{d_1}X$ (see Definition 2.19 for the formal construction).

The cover we get from this step is a cover of the faces complex $\mathrm{F}^{d_1}X$. The next step is to construct a cover $Y$ of $X$ from the cover $\widetilde{\mathrm{F}^{d_1}X}$ of $\mathrm{F}^{d_1}X$. Moreover, the complex $Y$ is such that its faces complex is isomorphic to $\widetilde{\mathrm{F}^{d_1}X}$, $\iota : \mathrm{F}^{d_1}Y \xrightarrow{\sim} \widetilde{\mathrm{F}^{d_1}X}$. 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.

---

[3]I.e. (1.1) holds for every triangle.

**Step 6 - 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 $\iota$ and the permutations $\pi_{s_1, t}$ above, we show that there is a agreement distribution $\mathcal{D}_Y$ where

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

Then all that remains is to use a known $99\%$-agreement theorem a la ($HA$) 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 $\{f_r\}$.

Note that the quantitative cosystolic expansion needed in Theorem 1.3 is rather modest; it is allowed to decay to zero exponentially in $d_1/k$. The local spectral expansion required is inverse polynomial in the dimension $d$.

## Related work

**PCPs and agreement tests**

Agreement tests in the low acceptance "$1\%$" regime appear in several PCP constructions, such as in the plane-versus-plane [RS97] and line-versus-line low degree tests [AS97]. Analysis in the low acceptance regime translates to a PCP with perfect completeness and small soundness $\varepsilon$. The fact that the test makes only two queries is essential for getting a two-query PCP such as label cover, which is an essential feature for downstream applications for hardness of approximation.

The most popular large-gap PCP (also known as label cover) used in hardness of approximation reductions is the 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 a black-box conversion from an agreement-test theorem to a parallel repetition theorem, see also [DS14b; DS14a].

A major motivation for studying derandomized agreement tests is towards construction of more length-efficient PCPs that still have a large gap between the completeness and soundness cases. The length parameter is important for various reasons ranging from practical implementations of PCPs to theoretical implications for hardness of approximation. An agreement test can often be transformed into a PCP construction, although such a transformation is not automatic. Indeed, Impagliazzo et. al [IKW12] constructed a derandomized family $X$ that satisfies ($LA$), but it took more work [DM11] to transform it into an length-efficient PCP theorem with a large gap. 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 $n \cdot \text{poly} \log n$) is an open question. Agreement tests for highly derandomized families $X$ may lead to progress in this direction.

More recently agreement tests appear in the proof of the 2:2 theorem [Din+18b; KMS18], where $X$ is the collection of all $\ell$-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, which was then proven in [KMS18]. Interestingly, the 2:2 agreement test cannot satisfy ($LA$), but rather a weakened version of it, that takes into account "zoom-ins" and "zoom-outs".

Here as well as in other low degree tests the local functions $f_s$ are not arbitrary but rather 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).

**Bogdanov's counter-example**

The gap amplification proof of the PCP theorem [Din07] is based on repeated applications of an agreement test in the high acceptance regime. Once the soundness has reached some constant $1 - \zeta$, it is natural to

wonder if one (or a few) additional steps can lead to arbitrarily small soundness $\varepsilon > 0$. Bogdanov shows [Bog05] an example of a PCP instance where the soundness cannot drop below $1/2$. We revisit his example in Section 7.2 and show that it is a special case of our Lemma 1.2. This suggests, perhaps, high dimensional expanders as candidates for gap amplification with soundness below $1/2$.

## Parameters of Agreement tests

Direct product tests, or agreement tests for the complete complex $\binom{[n]}{k}$, have been relatively well studied. For two queries, ($LA$) holds for $\varepsilon = \text{poly}(1/k)$, and fails when $\varepsilon \ll 1/k$ [DG08]. For three queries, ($LA$) holds for $\varepsilon$ even exponentially small in $k$ [IKW12; DS14b; DL17], and beyond that it necessarily fails. The parameters we get for our main theorem are significantly weaker, with $\varepsilon = \Omega(1/\log k)$; and further investigation of this tradeoff is left for future work.

## List-agreement and cosystolic expansion

The idea behind the last step in our proof of Theorem 1.3 is close in spirit to the list decoding argument in [GK22]. Their work builds 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.

## Comparison with [BM23]

Independent work by Bafna and Minzer [BM23] also gave a theorem characterizing low acceptance agreement in high dimensional expanders.

In their work they give a necessary and sufficient condition for a complex satisfying ($LA$), that is similar to the swap-cosystolic expansion condition. Let us explain the condition by explaining the "sufficient" part of their proof. The first part of their proof resembles steps $1-3$ of ours. Both works construct lists of functions on faces (using previous works on agreement tests) together with permutations on edges of the faces complex.

The two proofs diverge in step four. Roughly, in their work, the condition they impose on $X$ is that the faces complex is a swap-coboundary expander with respect to permutations *that also come with lists of functions as in step* 3 (see [BM23, Definitions 1.7, 1.8]). Let us call this condition ($\star$). Using ($\star$), they are able to construct local functions with 99% agreement on $X$. Then they patch them up to a global function in a way similar to our step 6.

In our proof, step 4 uses swap-cosystolic expansion to do almost the same thing - but instead of constructing local functions on $X$ we do so on a cover $Y$ of $X$. In the special case where $X$ is a swap-*coboundary* expander (which is strictly stronger than being a swap-*cosystolic* expander), the only possible covers of $X$ are disjoint copies of $X$. That is, in the swap-coboundary expander case, our proof gives the same construction as theirs and shows ($LA$). We also mention, that in retrospect, our proof works *as is*, if we replace our assumptions by ($\star$).

Our proof applies to a more general situation. It is common for a high dimensional expander to have covers (and not be a swap-coboundary expander) especially as known constructions of sparse high dimensional expanders are attained via families of complexes that cover one another. In this case, we manage to derive a structural characterization that is necessarily weaker than ($LA$) yet still encapsulates the global structure hiding behind functions that pass the agreement test, i.e. ($CLA$).

In addition, the work of [BM23] shows that ($\star$) is a necessary condition for satisfying ($LA$). They show that if an ensemble fails ($\star$), it must give rise to a counter example.

Thus, their condition ($\star$) is a *characterization* of complexes satisfying ($LA$). In contrast, our condition of swap-cosystolic expansion is not equivalent, and even swap-coboundary expansion seems to be stronger than ($LA$), so in any case not a characterization. Nevertheless, our condition applies to well-known constructions,

allowing us to derive concrete corollaries (Corollary 1.5 and Corollary 1.6) from our abstract theorem (for these derivations we rely on Theorem 1.4 from the companion paper).

In terms of concrete examples of complexes that fail ($LA$), both works point to the same issue. Their work also observes that the [LSV05a] complexes sometimes fail to satisfy ($LA$). The counterexample they give is in fact the same counterexample we give in Section 7.1 - they use a cocycle that comes from a cover to produce the counterexample.

We remark that the work of [BM23] raises an interesting open question: Does swap-coboundary expansion follow from their (a priori) weaker variant of ($\star$)?

## 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 pure $d$-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 $X(0) = A_0 \uplus A_1 \uplus \cdots \uplus A_d$ such that for every $s \in X(d)$ and $i \in \{0, \ldots, d\}$ 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 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 in a high dimensional expander

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

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

We denote by $\lambda(X_s)$ 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 (in absolute value) eigenvalue of the adjacency operator of $X_s^{\leqslant 1}$.

**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 $\lambda$-*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 $\lambda$-*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 $X^{\leqslant 1}$ itself 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 $(k,\ell)$-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 by choosing $t \sim X(k)$ and a uniform $s \subseteq t$ of size $\ell + 1$.

**Theorem 2.2** ([KO20])**.** *Let $X$ be a $d$-dimensional $\lambda$-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 $A \subseteq X(k)$, $\ell \leqslant k$ and $\zeta > 0$. Let*

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

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

A related walk is the swap walk. Let $k, \ell, d$ be integers such that $\ell + k \leqslant d - 1$. The $k, \ell$-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 $t \cup s \in X$ (and in particular $t, s$ are disjoint). 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 $u = t \cup s$. 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 $\lambda$-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 $\mathbb{F}_q^d$. Its 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 $\Sigma$ be some fixed alphabet and suppose we have an ensemble of functions $\mathcal{F} = \{f_r : r \to \Sigma \mid r \in X(k)\}$.

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

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

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 $r_1, r_2, ..., r_q \in X(k)$, where the agreement of the ensemble

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

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 $r_1, r_2, ..., r_q \subseteq t$ according to $\Delta_k(d)$.

We note that by the symmetry of $\mathcal{D}$ 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 $\eta, \varepsilon_0, e > 0$ be constants. We say that $\mathcal{D}$ is $(\eta, \varepsilon_0, e)$-*sound* if for every ensemble of functions $\mathcal{F}$ such that $\text{Agree}_{\mathcal{D}}(\mathcal{F}) = \varepsilon \geqslant \varepsilon_0$, there exists a function $G : X(0) \to \Sigma$, such that

$$\underset{r_1, r_2, ..., r_q \sim \mathcal{D}}{\mathbb{P}} \left[ \forall j \ G|_{r_j} \overset{1-\eta}{\approx} f_{r_j} \text{ and } \forall i, j \ f_{r_i} = f_{r_j} \right] \geqslant \frac{1}{2} \varepsilon^e. \tag{2.3}$$

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

We extend this definition to soundness with respect to covers in Definition 2.26, after we introduce simplicial covers formally.

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 $(\sqrt{k}, k^{-c}, 1)$-sound for $d \geqslant k^3$ and $c > 0$ [DG08][5].

---

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

[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 $\sqrt{k}$.
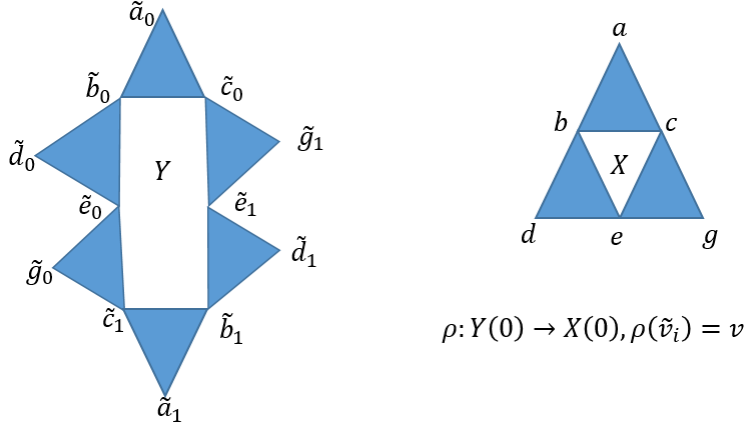
Figure 2: Non trivial connected cover

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 $(k^{-0.2}, exp(-\Omega(k^{1/2}), O(1))$-sound.

3. Dinur and Livni-Navon, together with the techniques used in [DG08, Theorem 5.1] show that the $Z$-test is $(\lambda, exp(-\Omega(k)), O(1))$-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 $(\eta, \varepsilon_0, O(1))$ such that $\eta \exp(\text{poly}(1/\varepsilon)) \ll 1$, so in particular, we will always think of $\varepsilon = 1/\log(k)$ 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. $\rho$ is a surjective homomorphism.

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 $\rho$ is an $\ell$-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 $\ell$-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 2 contains an example of a $\rho : Y \to X$ such that $Y$ is connected.

11

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 $\pi_{uv}$ for every directed edge $uv \in \vec{X}(1)$ 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 $\ell$-covers.

**Corollary 2.16.** *If $X$ is connected then any cover of $X$ is an $\ell$-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. □

*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 $\ell$-cover for any $\ell$. Let $\ell$ be the number of images of some arbitrary vertex $v \in X(0)$. Let $B = \{v' \in X(0) \mid |\rho^{-1}(v')| = \ell\} \neq \emptyset$. 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. □

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

*Claim* 2.17. Let $\rho : Y \to X$ be an $\ell$-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 $s = \{v\} \uplus r$. By definition, there are $\ell$-vertices $(v,1),(v,2),\ldots,(v,\ell)$ 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 $\tilde{s}_i = \tilde{r}_i \uplus \{(v,i)\}$ are $\ell$-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 $\ell$-cover. Without loss of generality we identify

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

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

$$\psi_\rho : \vec{X}(1) \to Sym(\ell), \ \psi_\rho(vu) = \pi_{uv} \tag{2.4}$$

where $\pi_{uv}(i) = j$ are such that $\{(v,i),(u,j)\} \in Y(1)$ (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

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

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 $\rho$ 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

$$\psi_\rho(vw) \circ \psi_\rho(uv) = \psi_\rho(uw). \tag{2.5}$$

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 $\ell$-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 $\psi_\rho(xy) = \pi_{yx}$ where $\pi_{yx}(i) = j$ if and only if $\{(x, i), (y, j)\} \in Y(1)$. Indeed, let $\pi_{uv}(i) = j$ and $\pi_{wv}(i) = k$ and we need to show that $\pi_{wu}(j) = k$, which is equivalent to showing that if $\{(v, i), (u, j)\}, \{(v, i), (w, k)\} \in Y(1)$ then $\{(u, j), (w, k)\} \in Y(1)$. $\rho : Y_{(v,i)}(0) \to X_v(0)$ is an isomorphism, it holds that $\{(u, j), (w, k)\} \in Y_{(v,i)}(1)$ and in particular the edge is in $Y(1)$. $\qed$

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

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

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

*Proof of Claim 2.20.* It is obvious that $\rho$ 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 $\{(v, i), (u, \pi_{uv}(i))\} \in Y(1)$ so $\rho|_{Y_{(v,i)}(0)}$ is surjective. A priori, there could have been more edges $\{(v, i), (u, k)\}$ if $i = \pi_{vu}(k)$, but because $\psi$ is asymmetric, $i = \pi_{vu}(k)$ implies that $\pi_{uv}(i) = k$ 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, $s \cup \{v\} = t \in X(m+1)$. We will show that $\tilde{t} = \{(v, i)\} \cup \{(u, j_u) \mid u \in s, j_u = \pi_{uv}(i)\}$ is in $Y(m+1)$, which implies that $\{(u, j_u) \mid u \in s, j_u = \pi_{uv}(i)\} = \tilde{s} \in Y_{(v,i)}(m)$ 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, $\{(v, i), (u, \pi_{uv}(i))\} \in Y(1)$ and $\pi_{uv}(i) = j_u$ 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 $\pi_{uv}(i) = j_u, \pi_{wv}(i) = j_w$. By (2.5) applied for $vuw \in X(2)$, it holds that $j_w = \pi_{wv}(i) = \pi_{wu}(\pi_{uv}(i)) = \pi_{wu}(j_u)$. 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 $Y$. $\qed$

### 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 $\ell$-cover $\rho : Y \to X$ there exists a partition of $Y$ to $\ell$ disconnected components $Y = Y_1 \uplus Y_2 \uplus \dots \uplus Y_\ell$ 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 $\ell$-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 $\ell$-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 $\rho$. 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 $\lambda$-one or two sided high dimensional expander. Then any connected cover is a $\frac{\lambda}{1-\lambda}$-one or two sided 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 $X_v$ is a $\lambda$-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 $\lambda$-one or two sided high dimensional expander, this implies that $Y_{\tilde{v}}$ is a $\lambda$-high dimensional expander. By Theorem 2.25, $Y$ is also a $\frac{\lambda}{1-\lambda}$-one or two sided high dimensional expander. $\square$

## 2.4 Cover-soundness à la (CLA)

Let us extend the definition of sound agreement tests to accommodate for our new notion of covers.

**Definition 2.26** (Cover sound). Let $X$ be a simplicial complex and let $\mathcal{D}$ be an agreement distribution on $X$. Let $\eta, \varepsilon_0, e > 0$ be constants. We say that $\mathcal{D}$ is $(\eta, \varepsilon_0, e)$-*cover-sound* if for every ensemble of functions $\mathcal{F} = \{f_r : r \to \Sigma \mid r \in X(k)\}$ such that $\mathrm{Agree}_{\mathcal{D}}(\mathcal{F}) = \varepsilon \geqslant \varepsilon_0$, there exists a simplicial $\frac{1}{\varepsilon^e}$-cover $\rho : Y \to X$ and a global function $G : Y(0) \to \Sigma$ such that

$$\mathbb{P}_{r \in Y(k)} \left[ f_{\rho(r)} \circ \rho \overset{1-\eta}{\approx} G|_r \right] \geqslant \frac{1}{2} \varepsilon^e. \tag{2.6}$$

The probability in (2.6) is equivalent to choosing $r \in X(k)$ and then a random preimage $\tilde{r} \in Y(k)$. Therefore this inequality implies that for at least $\frac{1}{2}\varepsilon^e$-fraction of the $f_r$'s, the function agrees with $G|_{\tilde{r}}$ for one of the preimages $\tilde{r} \in \rho^{-1}(r)$ (i.e. using $\rho$ to send vertices from $\tilde{r}$ to $r$).

We note that here we did not require the cover to explain most of the agreement, but this definition actually implies such a statement. See discussion in [DG08] (there it is done for the complete complex, but their techniques are general).

## 2.5 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 $uvw \in \overrightarrow{X}(2)$. For our result we will need simplicial complexes where this relation is locally testable. For this we define for every two function $\psi, \phi : \overrightarrow{X}(k) \to Sym(\ell)$ their distance

$$\mathrm{dist}(\psi, \phi) = \mathbb{P}_{s \in \overrightarrow{X}(k)} [\psi(s) \neq \phi(s)]. \tag{2.7}$$

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

$$\delta(\psi) = \psi(wu) \circ \psi(vw) \circ \psi(uv). \tag{2.8}$$

We are ready to define cosystolic expansion.

**Definition 2.27.** Let $X$ be a $d$-dimensional simplicial complex for $d \geqslant 2$. Let $\beta > 0$. We say that $X$ is a $\beta$-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

$$\beta \, \mathrm{dist}(\psi, \phi) \leqslant \mathrm{wt}(\delta\psi). \tag{2.9}$$

Furthermore, we say that $X$ is a $\beta$-coboundary expander if $X$ is a $\beta$-cosystolic expander, and it is simply connected.

An equivalent way to define coboundary expansion is that for every, $\phi \in Z^1(X, Sym(\ell))$ there is some $h : X(0) \to Sym(\ell)$, such that $\phi(uv) = h(v) \circ h(u)^{-1}$ for all $uv \in \vec{X}(1)$. We call $\phi$'s that have such an $h$ coboundaries, hence the name "coboundary expansion".

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 $\psi$ fails the test. When $X$ is a $\beta$-cosystolic expander, this implies that if $\mathrm{wt}(\psi) = \varepsilon$ then there is a function $\phi \in Z^1(X, Sym(\ell))$ that is $\varepsilon/\beta$ close to $\psi$.

We remark that in other works, many other coefficient groups were used instead of $Sym(\ell)$. For our result this definition is sufficient. Note that some prior works such as [EK16] also require that $\psi \in Z^1 \setminus B^1$ have large support. This separate requirement is not necessary in our work.

Dinur and Meshulam already observed that cosystolic expansion (and coboundary expansion) is equivalent testability of covers, which they call *cover stability* [DM22].

### 2.5.1   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 $\psi(vu) = \pi_{uv,u}^{-1} \circ \pi_{uv,v} = \pi_{u,uv} \circ \pi_{uv,v}$[6]. There is a natural consistency property that implies that $\psi$ is (close to) a cocycle: 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

$$\pi_{t,v} = \pi_{t,e} \circ \pi_{e,v}, \tag{2.10}$$

then $\delta\psi(uvw) = Id$. See Figure 3 for an illustration.

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

**Definition 2.28** (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* 2.29. 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 every triangle $t \in X(2)$, if for all six choices of $v \in e \in t$, $\delta\phi(\{v \in e \subset t\}) = Id$, then $\delta\psi(t) = Id$.

**Corollary 2.30.** *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 $\psi \in Z^1(X, Sym(\ell))$.*

---

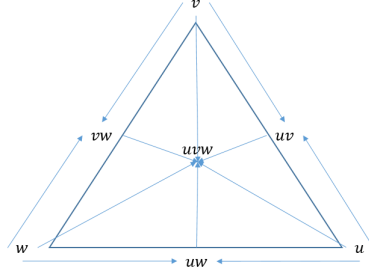[6]As a notational convention we use $\pi_{x,y} = \pi_{y,x}^{-1}$.

Figure 3: The following diagram should commute for $\delta\psi(uvw) = Id$

*Proof of Claim 2.29.* 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 $\delta\phi(u, uv, uvw) = Id$ so $\phi(uv, uvw)\phi(u, uv) = \phi(u, uvw)$ and $\phi(uv, v)\phi(uvw, uv) = \phi(uvw, v)$. Similarly, we have that

$$\psi(uw) = \phi(uvw, w)\phi(u, uvw); \ \psi(wv) = \phi(uvw, v)\phi(w, uvw). \tag{2.11}$$

Thus,

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

$\square$

## 2.6 Swap cosystolic expansion and the Faces complex

For any simplicial complex $X$, we define a related complex called the *faces complex* of $X$, denoted $\mathrm{F}^{d_1}X$. This is a complex whose vertices are $d_1$ faces of $X$ and whose edges connect two vertices whose disjoint union is also a face of $X$. Thus, the underlying graph of $\mathrm{F}^{d_1}X$ is the graph of the swap walk on $X$, see [AJT19; DD19]. This definition already appeared as Definition 1.7 but we restate it again here.

**Definition 2.31** (Faces Complex)**.** Let $X$ be a $d$-dimensional simplicial complex. Let $d_1 \leqslant d$. We denote by $\mathrm{F}^{d_1}X = F(X, d_1)$ the simplicial complex whose vertices are $\mathrm{F}^{d_1}X(0) = X(d_1)$ and whose faces are all $\left\{\{s_0, s_1, ..., s_j\} \,\middle|\, s_0 \uplus s_1 \uplus \cdots \uplus s_j \in X\right\}$.

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 $\mathrm{F}^{d_1}X$. We omit $d_1$ from the notation when it is clear.

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

**Definition 2.32** (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 $\mathrm{F}^{d_1}(X_r) = \mathrm{F}^{d_1}X_r$ is connected. Moreover, if $r \in X(0)$ then we require that $\mathrm{F}^{d_1}X_r$ is *simply connected*. When $d_1$ is clear from context we omit it and say that $X$ is *well connected*.

We say that a complex $X$ has swap cosystolic expansion if $\mathrm{F}^{d_1}X$ is a cosystolic expander. Namely, recalling Definition 2.27,

**Definition 2.33** (Swap cosystolic expansion). Let $d_1 \in \mathbb{N}$ and let $\beta > 0$. A simplicial complex $X$ of dimension $d > d_1$ has $(\beta, d_1)$ swap-cosystolic -expansion if $F^{d_1}X$ is a $\beta$ cosystolic expander. We further say that $X$ has $(\beta, d_1)$ swap-coboundary-expansion if it has $(\beta, d_1)$ swap-cosystolic-expansion and furthermore, $F^{d_1}X$ is simply connected.

## 2.7 Sampling in HDXs

**Theorem 2.34** ([DH23]). *Let $k \leqslant d_1 \leqslant d$. Let $\gamma < 1$ and let $X$ be $\frac{\gamma}{d}$-one sided high dimensional expander. Let $\zeta > 0$ and let $f : X(k) \to [0,1]$. Let*

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

*Then*

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

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

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

*then*

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

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

Theorem 2.34 allows us to get a similar sampling guarantee on sampling of of tuples according to $\mathcal{D}$ distributions extended to high dimensional expanders $X$.

*Claim* 2.35. Let $k \leqslant d_1 \leqslant d$ and $q \in \mathbb{N}$. Let $d_1 \geqslant qk$ and let $\zeta > 0$ be some constant. Let $X$ be a $d$-dimensional simplicial complex as in Theorem 2.34. Let $\mathcal{D}$ be an agreement distribution over $\Delta_{qk+q}(k)$. and let $E \subseteq \operatorname{supp} \mathcal{D}_X$. Then

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

*Proof of Claim 2.35.* 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 2.34 gives us the claim. $\qquad\square$

## 2.8 Suitable complexes

**Definition 2.36** (Suitable complex). Let $d > k > 0$ be integers, and let $\alpha > 0$. Let $X$ be a $d$-dimensional simplicial complex. We say that $X$ is $(d, k, \alpha)$-*suitable* if it has the following properties:

1. There exists some integer $d_1 < d$ with the following properties:

   (a) $k^3 \leqslant d_1 \leqslant d \exp(-\alpha \frac{d_1}{k})$.

   (b) $X$ has $(\exp(-\alpha \frac{d_1}{k}), d_1)$-swap-cosystolic expansion, as in Definition 2.33.

   (c) $X$ is $d_1$-well connected as in Definition 2.32.

2. $X$ is a $\frac{1}{d^2}$-two-sided local spectral expander.

3. $X$ is a clique complex.

We also say that $X$ is suitable with respect to $d_1$.

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 2.37.**

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

1. *The $SL_d(\mathbb{F}_q)$-spherical building is a $(d, k, \alpha)$-suitable.*

2. *$d$-dimensional [LSV05a] complexes, such that the link of every vertex is isomorphic to $SL_{d-1}(\mathbb{F}_q)$-spherical building are $(d, k, \alpha)$-suitable.*

This paper is focused on the 1%-agreement theorem; the proof of swap cosystolic expansion gives rise to other significant challenges so we defer the proof of the swap cosystolic expansion to a companion paper [DD23b].

*Proof.* In both cases we fix $k$ and take $d_1 = k^3$.

1. (a) **Swap-coboundary and cosystolic expansion.** This is the main theorem in [DD23b], which shows that $Fd_1X$ is a $\exp(-\Omega(\sqrt{d_1}))$-cosystolic expander for these complexes (a coboundary expander in the spherical building case).

   (b) **Well connectedness.** Let $s$ be a face in a complex $X$ (which is one of the complexes in the theorem). Then $(\mathrm{F}^{d_1}X)_s = \mathrm{F}^{d_1}(X_s)$. In both cases, this is (isomorphic to) a faces complex of a link in the spherical building. In particular, [DD23b] shows that these are coboundary expanders - a strictly stronger property than simple connectivity.

2. **high dimensional expansion.** High dimensional expansion of the spherical buildings decays with $q$. This bound is independent of the ambient dimension, as proven in [AG22]. Therefore there is a sufficiently large enough $q$ such that the spherical building is a $\frac{1}{d^2}$-high dimensional expander. High dimensional expansion of the [LSV05a] complexes follows from the fact that their links are spectral expanders and Theorem 2.25 by [Opp18].

3. **Clique complexes.** It is easy to show that the spherical building is a clique complex by definition. If $\{u_0, u_1, \ldots, u_k\}$ are such that all edges appear in the spherical building, it means that every two subspaces are contained in one another. Rearranging them by dimension shows that this in indeed a flag, i.e. a face in the spherical building. The fact that quotients of the spherical buildings are clique complexes appear e.g. in [LSV05b].

$\square$

# 3 Main Agreement Theorem

In this section we describe and prove our main agreement theorem. Recall that we defined *suitable complexes* in Definition 2.36.

**Theorem 3.1** (Main)**.** *For every $\varepsilon_0 > 0$, $p > 1$, and $k \geqslant \exp(\text{poly}(1/\varepsilon_0))$ there exists $\alpha = \text{poly}(\varepsilon)$ and $d_0 \in \mathbb{N}$ such that the following holds for any $d \geqslant d_0$. Let $\eta \leqslant \exp(-\text{poly}(1/\varepsilon_0))$ be sufficiently small. Let $\mathcal{D}$ be an agreement distribution on $\Delta_k(pk)$ such that its extension to $\Delta_k(m)$ is $(\eta, \text{poly}(\frac{1}{k}), O(1))$-sound for every $m \geqslant k^3$. Let $X$ be a $(d, k, \alpha)$-suitable complex and let $\mathcal{D}_X$ be the extension of $\mathcal{D}$ to $X$. Then $\mathcal{D}_X$ is $(\varepsilon_0, \gamma, O(1))$-cover sound where $\gamma = \exp(\text{poly}(1/\varepsilon_0))\eta$.*

We derive Corollary 1.6 and Corollary 1.5 as corollaries of this theorem in the end of this section.

**Corollary 3.2** (Formal version of Corollary 1.6)**.** *For every $\varepsilon_0 > 0$ and $k \geqslant \exp(\text{poly}(1/\varepsilon_0))$ there exists $d_0 \geqslant k$ and integer $q_0$ such that the following holds. Let $\eta \leqslant \exp(-\text{poly}(1/\varepsilon_0))$. 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:*

1. *$\mathcal{D}_X$ is $(\eta, \varepsilon_0, O(1))$-sound where $\gamma = \exp(\text{poly}(1/\varepsilon))\text{poly}(1/k)$.*

2. *There is an $\ell = \text{poly}(1/\varepsilon)$ and a list of functions $G_1, G_2, \ldots, G_\ell : Y(0) \to \Sigma$ such that*

$$\mathop{\mathbb{P}}_{\{r_i\} \sim \mathcal{D}}\left[\{f_{r_i}\} \text{ passes the agreement test but for every } G_m \text{ there is at least one } f_{r_i} \text{ s.t. } f_{r_i} \overset{1-\gamma}{\not\approx} G_m|_{r_i}\right] \leqslant \varepsilon^2.$$
(3.1)

**Corollary 3.3.** *For every $\varepsilon_0 > 0$, $\eta > 0$ and $k \geqslant \exp(\text{poly}(1/\varepsilon_0)) \in \mathbb{N}$ there exists $C > 0$ and an infinite family of (non-isomorphic) complexes $X_n$ such that the following holds.*

1. *$|X_n(k)| \leqslant C|X_n(0)|^2$.*

2. *The extended $V$ or $Z$-test on $X_n$ is an $(\eta, \varepsilon_0, O(1))$-sound.*

**Corollary 3.4** (Formal version of Corollary 1.5)**.** *For every $\varepsilon_0 > 0$ and $k \geqslant \exp(\text{poly}(1/\varepsilon_0))$, there exists $d \geqslant k$ 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 $q \geqslant q_0$. Let $\mathcal{D}$ be either the $V$-test or the $Z$-test for sets of dimension $k$. Then $\mathcal{D}_X$ is $(\gamma, \varepsilon_0, O(1))$-cover sound where $\gamma = \exp(\text{poly}(1/\varepsilon))\text{poly}(1/k)$.*

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. We will give a high-level proof, relying on a few claims and lemmas that will be stated as needed. Then we will prove all outstanding claims in the following sections, formally completing the proof. For the rest of the section, we fix $X, \mathcal{D}, \mathcal{F}$ to be as in the theorem statement.

### 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 3.5.** *Let $X, \mathcal{D}, \mathcal{F}$ and $\varepsilon$ be as in Theorem 3.1. Let $d_1$ be an integer for which $X$ is suitable. Let $d_2 = 2d_1 + 1$ and let $d_3 = 3d_1 + 2$. There exists $\ell = \text{poly}(1/\varepsilon)$ 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. Let $\gamma = \eta\exp(\text{poly}(1/\varepsilon))$.*

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

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

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 $\pi_{t,s} : [\ell] \to [\ell]$ such that*

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

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

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

19

## 3.2 A cover $\widetilde{\mathrm{F}^{d_1}X} \twoheadrightarrow \mathrm{F}^{d_1}X$

Using Lemma 3.5 we construct an almost-cycle $\psi \in C^1(\mathrm{F}^{d_1}X, Sym(\ell))$ as follows. For each edge $(s_1, s_2)$ of $\mathrm{F}^{d_1}X$, denoting $t = s_1 \cup s_2$, let

$$\psi((s_1, s_2)) = \pi_{t,s_2}^{-1} \circ \pi_{t,s_1}.$$

**Lemma 3.6.**

$$\mathrm{wt}(\zeta(\psi)) = \mathop{\mathbb{P}}_{(s_1,s_2,s_3) \in \mathrm{F}^{d_1}X(2)} [\psi(s_1, s_2) \circ \psi(s_2, s_3) \neq \psi(s_1, s_3)] = exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k}). \tag{3.2}$$

*Proof of Lemma 3.6.* Recall Definition 2.31 of the faces complex $\mathrm{F}^{d_1}X$ and recall Definition 2.28 of the flags complex $GX$. We wish to use Claim 2.29 for the flag complex of the faces complex, namely for $G(\mathrm{F}^{d_1}X) = G\mathrm{F}^{d_1}X$. We define $\phi : G\mathrm{F}^{d_1}X(1) \to Sym(\ell)$ as follows. Let $\{\{s_1\}, \{s_1, s_2\}, \{s_1, s_2, s_3\}\}$ be a triangle in the flag complex. Let $t = s_1 \cup s_2$ and let $u = s_1 \cup s_2 \cup s_3$. Then

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

Where the permutations are the ones given in Lemma 3.5. By the third item of Lemma 3.5, it holds that with probability $\mathrm{wt}(\delta\phi) = \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ and by Claim 2.29 the same holds for $\delta(\psi)$. □

Using Lemma 3.6 and $\beta = \exp(-\alpha\frac{d_1}{k})$-cosystolic expansion of $\mathrm{F}^{d_1}X$ we have that

**Corollary 3.7.** *There exists a cocycle* $\psi' \in Z^1(\mathrm{F}^{d_1}X, Sym(\ell))$ *such that* $\mathrm{dist}(\psi', \psi) = \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k} + \alpha\frac{d_1}{k})$, *namely,*

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

Here is where we require that $\beta$ is large enough relative $\exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$. By Definition 2.19, $\psi'$ gives rise to an $\ell$-cover $\nu_{\psi'} : \widetilde{\mathrm{F}^{d_1}X} \to \mathrm{F}^{d_1}X$.

## 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 (2.6) holds). So far we managed to construct a cover $\nu = \nu_{\psi'} : \widetilde{\mathrm{F}^{d_1}X} \to \mathrm{F}^{d_1}X$. It seems natural to expect that every such cover $\widetilde{\mathrm{F}^{d_1}X}$ 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 3.8.** *Let* $\nu : \widetilde{\mathrm{F}^{d_1}X} \to \mathrm{F}^{d_1}X$ *be an $\ell$-covering map. Then there exists an $\ell$-covering map* $\rho : Y \to X$ *such that* $FY$ *and* $\widetilde{\mathrm{F}^{d_1}X}$ *are isomorphic.*

$$
\begin{array}{ccc}
Y & \xrightarrow{faces} & FY \overset{\iota}{\cong} \widetilde{\mathrm{F}^{d_1}X} \\
\downarrow{\scriptstyle\rho} & & \downarrow{\scriptstyle\nu} \\
X & \xrightarrow{faces} & \mathrm{F}^{d_1}X
\end{array}
$$

We denote by $\iota : FY \overset{\sim}{\to} \widetilde{\mathrm{F}^{d_1}X}$ the isomorphism. We prove this lemma in Section 5.

## 3.4 The global function $G$

Given Lemma 3.8, 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 3.8. 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* 3.9. There exists a universal constant $c > 0$ such that $\mathbb{P}_{\tilde{s} \in Y(d_1)} \left[ h_{\tilde{s}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{s}} \right] \geqslant 1 - \exp(-\operatorname{poly}(\varepsilon) \left( \frac{d_1}{k} \right)^c) - \frac{1}{d_1^c}$.

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

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

By Lemma 3.5 $\mathbb{P}\left[E_1\right] \leqslant 1 - \operatorname{poly}(\varepsilon)$. Thus to show the theorem it is enough to bound $\mathbb{P}\left[E_2\right] \leqslant \exp(-\operatorname{poly}(\gamma)k) + \exp(-\operatorname{poly}(\varepsilon) \left( \frac{d_1}{k} \right)^c) + \frac{1}{d_1^c}$ for the $c > 0$ in Claim 3.9 ($k$ 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,

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

The first item in this sum is bounded by $\exp(-\operatorname{poly}(\varepsilon) \left( \frac{d_1}{k} \right)^c)$ by Claim 3.9. The second is bounded by $\exp(-\operatorname{poly}(\gamma)k)$ by a standard Chernoff argument.

Thus

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

and Theorem 3.1 follows.

## 3.5 Deriving the corollaries

*Proof of Corollary 3.2.* We begin with the first item, i.e. showing that there exists *one* global function $G : X(0) \to \Sigma$ for $\varepsilon^{O(1)}$ of the $f_r$'s, $G|_r \approx f_r$.

Fix $\varepsilon_0$ and let $d, k, \alpha$ be as in Theorem 3.1 (for agreement distributions for either 2 or 3 sets). By Theorem 2.38 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 $(\varepsilon_1, \eta)$-sound for $\varepsilon_1, \eta = \operatorname{poly}(1/k)$. We take $k$ to be large enough so that $\varepsilon_0 \geqslant \varepsilon_1$ and so that $\gamma = \exp(1/\operatorname{poly}(\varepsilon_0))\eta < 1$. 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 Definition 2.26.

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 connected component in the cover is isomorphic to $X$, 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 (2.6) holds conditioned on sampling $\tilde{r}$ in $G_i$'s component.

Turning to (3.1), by [DG08, Theorem 5.1], if

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

holds, then there is an (inefficient) algorithm that outputs a list of functions $G_1, G_2, \ldots, G_\ell$ for $\ell = \text{poly}(1/\varepsilon)$ such that (3.1) holds. $\qquad \square$

*Proof sketch of Corollary 3.3.* By Corollary 3.2 for every $\varepsilon_0, \eta > 0$ and $k$ sufficiently large, there is an integer $d_0$ and a prime power $q_0$ such that the following holds. Let $X_q$ be the $d_0$-dimensional spherical building over any $\mathbb{F}_q$ for $q \geqslant q_0$. Then the $V$ or $Z$ tests on $X$ are $(\eta, \varepsilon_0, O(1))$-sound.

Let us show that $|X_q(k)| = O_{d_0,k}|X_q(0)|$. The number of flags in a spherical building is $|X(d_0)| \approx q^{d^2/2}$ and for $k < d$ clearly $|X(k)| \leqslant 2^d \cdot |X(d)|$. However, the number of elements in $X(0)$, which are subspaces in $|X(0)| \geqslant \binom{d_0}{d_0/2}_q \geqslant q^{d^2/4}$. The corollary follows. $\qquad \square$

*Proof of Corollary 3.4.* Fix $\varepsilon_0$ and let $d, k, \alpha$ be as in Theorem 3.1 (for agreement distributions for either 2 or 3 sets). By Theorem 2.38 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 $(\varepsilon_1, \eta, O(1))$-sound for $\varepsilon_1, \eta = \text{poly}(1/k)$. We take $k$ to be large enough so that $\varepsilon_0 \geqslant \varepsilon_1$ and so that $\gamma = \exp(1/\text{poly}(\varepsilon_0))\eta < 1$. The corollary follows directly from Theorem 3.1. $\qquad \square$

## 4 From one percent agreement to list agreement

In this section we prove Lemma 3.5.

**Lemma** (Restatement of Lemma 3.5). *Let $X, \mathcal{D}, \mathcal{F}$ and $\varepsilon$ be as in Theorem 3.1. Let $d_1$ be an integer for which $X$ is suitable. Let $d_2 = 2d_1 + 1$ and let $d_3 = 3d_1 + 2$. There exists $\ell = \text{poly}(1/\varepsilon)$ 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. Let $\gamma = \eta \exp(\text{poly}(1/\varepsilon))$.*

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

$$\Pr_{r \in X(k), r \subset s} \left[ L_s^i|_r \stackrel{1-\gamma}{\approx} f_r \right] \geqslant \text{poly}(\varepsilon).$$

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 $\pi_{t,s} : [\ell] \to [\ell]$ such that*

$$\Pr_{s,t} \left[ \forall i = 1, 2, \ldots, \ell, j = \pi_{t,s}(i) \; L_s^i \stackrel{1-\gamma}{\approx} L_t^j|_s \right] \geqslant 1 - \exp(-\text{poly}(\varepsilon)\frac{d_1}{k}).$$

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

$$\Pr_{s \subset t \subset u} [\pi_{s,u} = \pi_{s,t} \circ \pi_{t,u}] \geqslant 1 - \exp(-\text{poly}(\varepsilon)\frac{d_1}{k}).$$

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 $(\varepsilon_0, \eta)$-sound. We also assume that $\text{Agree}_{\mathcal{D}}(\mathcal{F})^M = \varepsilon^M \geqslant \varepsilon_0$ for some constant integer $M > 0$ (which we do not explicitly calculate).

For a function $g : X(0) \to \Sigma$ We denote by $\text{supp}_\zeta(g) = \left\{ r \in X(k) \;\middle|\; f_r \stackrel{1-\zeta}{\approx} g \right\}$ and call this set the $\zeta$-support of $g$. We also denote by $A_\zeta(g) = \left\{ \{r_1, r_2\} \;\middle|\; r_1, r_2 \in \text{supp}_\zeta(g) \text{ and } f_{r_1} = f_{r_2} \right\}$ the "agreeing edge set of $g$" and by $\alpha_\zeta(g) = \Pr_{\{r_1, r_2\} \sim D}[\{r_1, r_2\} \in A_\zeta(g)]$. Here $D$ is the agreement distribution. When the ensemble of functions is not clear from context we denote $\text{supp}_\zeta^{\mathcal{F}}(g), A_\zeta^{\mathcal{F}}(g)$ 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 $\overline{L}_{\tau,\zeta}(t) = \left\{ g : t \to \Sigma \mid \alpha_\zeta^{\mathcal{F}_t}(g) \geqslant \tau \right\}$. 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 $\mathrm{supp}_\zeta(g), \alpha_\zeta(g)$ or $A_\zeta(g)$ 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 3.5

The idea is to use the the agreement theorem at hand to create lists of functions $\overline{L}(s) = \{L_s^1, L_s^2, \ldots, L_s^\ell\} \subseteq \overline{L}_{\tau,\zeta}(s)$ 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 $\tau, \zeta$ and $\gamma$, 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 $\overline{L}_{\tau,\zeta}^0(s)$. 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 $\gamma$-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 - meaning that every function $g$ that has non-trivial agreement with the $f_r$'s has a close by function in the list. That is, if $g : s \to \Sigma$ has that $f_r \approx g|_r$ for $\geqslant \tau = \mathrm{poly}(\varepsilon)$ fraction of the $r \in X(k)$ inside $s$, then either $g \in \overline{L}(s)$ or $g' \in \overline{L}(s)$ 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 $\gamma$-separated if for every two $g_1, g_2 \in \overline{L}_1$ it holds that $\mathrm{dist}(g_1, g_2) > \gamma$.

2. We say that $\overline{L}_1$ is $\gamma$-dense in $\overline{L}_2$ if for every $h \in \overline{L}_2$ there exists $g \in \overline{L}_1$ such that $\mathrm{dist}(g, h) \leqslant \gamma$ (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 $\overline{L}_{\tau,\zeta}(s)$ (to overcome the second issue). In fact, a method by [Din+18a] allows us to find lists that are $\gamma$-dense but $10\gamma$-separated (for a smartly chosen $\gamma$). This is the content of Claim 4.8. We will need this property in order to match between lists. If $\overline{L}(s)$ is $\gamma$-dense in $\overline{L}_{\tau,\zeta}(s)$ and $\overline{L}(t)|_s = \{L_t^i|_s \mid L_t^i \in \overline{L}(t)\}$ is contained in $\overline{L}_{\tau,\zeta}(s)$ then for every $L_t^i \in \overline{L}(t)$ there exists some $L_s^j \in \overline{L}(s)$ such that $L_s^j \overset{1-\gamma}{\approx} L_t^i|_s$.[7]

The $10\gamma$-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 $\gamma$-close to $L_t^i$ then they are $2\gamma$-close to one another, which is a contradiction to $10\gamma$-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 $10\gamma$-separated, we get that this function $\pi_{s,t}$ is injective for most $s \subseteq t$. If $L_t^i, L_t^{i'}$ are $10\gamma$-separated, then for most $s \subseteq t$, it will hold that $L_t^i|_s, L_t^{i'}|_s$ will still be more than $2\gamma$-far. Hence, both $L_t^i|_s, L_t^{i'}|_s$ cannot be $\gamma$-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 $\mathrm{poly}(\frac{1}{\varepsilon})$. 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

---

[7]There is a subtlety here that even if $\alpha_\zeta(L_t^i|_s) \geqslant \tau$, it could be that $\alpha_\zeta(L_t^i) < \tau$. 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 $\alpha_\zeta(L_t^i|_s) = \alpha_\zeta(L_t^i)$.

separated. It could be that there is a "new" function $L_s^j \in \overline{L}_{\tau,\zeta}(s)$ 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 $\overline{L}_{\tau,\zeta}(s)$.

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 $\zeta' = O(\zeta)$, and a function $g : t \to \Sigma$ such that $\alpha_{\zeta'}(g) \geq \tau - \text{poly}(\varepsilon)$, that is $\zeta'$-far from all functions in $\overline{L}_{\tau,\zeta}(t)$. Showing this is where the previous agreement theorems come into play; if it is true that $\overline{L}_{\tau,\zeta}(t)$ is not dense for many of the $s \subseteq t$, this means that even after rerandomizing $f_r$ for $r \in \bigcup_{g \in \overline{L}_{\tau,\zeta}(t)} \text{supp}_\zeta(g)$, the ensemble will pass the agreement test with non-negligible probability. By the agreement soundness guarantee, there is a new function $g$ such that $A_\zeta(g) \geq \tau^3$. We then show that by scaling $\zeta$ by a constant factor, it also holds that $A_{\zeta'}(g) \geq \tau - \text{poly}(\varepsilon)$ (this happens because $g$ needs to agree with the functions $f_r$ for $r \notin \bigcup_{g \in \overline{L}_{\tau,\zeta}(t)} \text{supp}_\zeta(g)$, 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 $\zeta, \tau$ such that there are no new functions like this: We take any sequence $\{(\tau_i, \zeta_i)\}_{i=1}^m$ where $\tau_1 = \varepsilon^2, \tau_{i+1} = \tau_i - \varepsilon^{50}$ and $\zeta_{i+1} = 20\zeta_i$. Suppose that for every such pair there is a "new" $g$ such that $\alpha_{\zeta_{i+1}}(g) \geq \tau_{i+1} \geq \varepsilon^{10}$ that is far from all functions in $\overline{L}_{\tau_i,\zeta_i}(t)$. This will imply that $\mathbb{P}\left[A_{\zeta_i}(g_i) \setminus \bigcup_{j=1}^{i-1} A_{\zeta_j}(g)\right] \geq \varepsilon^{10}$. 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

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

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

We use the pigeonhole principle to find a pair $(\tau_i, \zeta_i)$ and $M_{(\tau_i,\zeta_i)} \subseteq X(d)$ of relative size $\text{poly}(\varepsilon)$, such that for every $t \in M_{(\tau_i,\zeta_i)}$ and for all but a negligible fraction of $s \subseteq t$ it holds that $\overline{L}_{\tau_i,\zeta_i}(t)|_s$ is dense inside $\overline{L}_{\tau_{i+1},\zeta_{i+1}}(s)$. Finally, using this set, we show that if $t \in M_{(\tau_i,\zeta_i)}$ then most $s \subseteq t$ also have the same property (i.e. that $\overline{L}_{\tau_i,\zeta_i}(s)|_u$ is dense inside most $\overline{L}_{\tau_{i+1},\zeta_{i+1}}(u)$ 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 $(\tau_i, \zeta_i)$. Then we find lists $\overline{L}(s)$ that are $\gamma$-dense and $10\gamma$-separated inside $\overline{L}_{\tau_{i+1},\zeta_{i+1}}(s)$ 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 3.5.

## 4.2 Proof of Lemma 3.5

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

**Definition 4.2.** Let $s \subseteq t$, $\zeta > 0$ 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 $|\text{dist}(g_1|_s, g_2|_s) - \text{dist}(g_1, g_2)| \leq \varepsilon^{100}$. 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 $\left|\mathbb{P}\left[A_\zeta(g)\right] - \mathbb{P}\left[A_\zeta(g|_s)\right]\right| \leq \varepsilon^{100}$.

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

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

1. $\overline{L}(t)$ is $9\gamma$-separated.

2. $\overline{L}(t)$ is $\gamma$-dense in $\overline{L}_{\tau-\varepsilon^{100},\zeta}(t)$.

3. $|\overline{L}(t)| = \ell$.

4. For $i = 2, 3$, and $i' < i$. Then the following set $P_t(i')$ has probability $\mathbb{P}_{s \subseteq t, s \in X(i')}[P_t(i')] \geq 1 - \exp(-\Omega(\text{poly}(\varepsilon)\frac{d_1}{k}))$. $P_t(i')$ is all the $s \subseteq t, s \in X(i')$ such that $s$ samples $\overline{L}(t)$'s agreement well.

24

The proof of Lemma 3.5 relies on the following proposition.

**Proposition 4.4.** *There exists parameters $\tau = \mathrm{poly}(\varepsilon), \zeta, \gamma \leqslant \eta \exp(\mathrm{poly}(1/\varepsilon))$ 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 $\mathbb{P}\left[M_i\right] \geqslant 1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$.*

We will also use this claim,

*Claim* 4.5. Let $L_1 = \{g_1, g_2, \ldots, g_{\ell_1}\}, L_2 = \{h_1, h_2, \ldots, h_{\ell_2}\}$ be such that $L_1$ is $\gamma$-dense in $L_2$ and such both $L_1$ and $L_2$ are $2\gamma$-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 3.5.* Recall the definition of $M_i$ is Definition 4.3. Given Proposition 4.4, for every $s \in M_1, t \in M_2$ and $u \in M_3$ we take such a list $\overline{L}(x) = \{L_x^1, L_x^2, \ldots, L_x^\ell\}$, where $x$ is $s, t, u$ respectively (for $x \notin M_i$ we take arbitrary lists). By definition, for every $x \in M_i$ and every function in $\overline{L}_x$ satisfies item (1) in Lemma 3.5 by definition. Let us verify the second and third items. We show the following for pairs $s \subseteq t$, $s \subseteq u$ and $t \subseteq u$, we denote such a pair as $x \subseteq y$ for short. For such a pair where $y \in M_i, x \in P_y(i') \cap M_{i'}$ (where $i, i'$ are the corresponding sets) we define $\pi_{x,y} : [\ell] \to [\ell]$ by

$$\pi_{x,y}(j) = k \Leftrightarrow L_y^j|_x \overset{1-\gamma}{\approx} L_x^k. \tag{4.1}$$

If either $y \notin M_i$ or $x \notin P_y(i') \cap M_{i'}$ 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. If this permutation is well defined for $1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ of the pairs $x \subseteq y$ this implies item (2 in Lemma 3.5.

First we note that if $x$ samples $\overline{L}(y)$'s agreement well then $\overline{L}(y)|_x \subseteq \overline{L}_{\tau-\varepsilon^{100},\zeta}(x)$ and $\overline{L}(y)|_u$ is $8\gamma$-separated. Moreover if $\overline{L}(x)$ is $\gamma$-dense in $\overline{L}_{\tau-\varepsilon^{100},\zeta}(x)$, and $\overline{L}(y)|_x \subseteq \overline{L}_{\tau-\varepsilon^{100},\zeta}(x)$ then in particular $\overline{L}(x)$ is $\gamma$-dense in $\overline{L}(y)|_x$. Finally, by definition $\overline{L}(y)$ is also $9\gamma$-separated. Hence Claim 4.5 implies that this is a well defined and injective function. As both $\overline{L}(y)|_x$ and $\overline{L}(y)$ have the same size $\ell$, this is indeed a permutation.

Finally, we will prove the third item in Lemma 3.5,

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

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 $k$,
$$\left|\mathrm{dist}(L_t^j|_s, L_u^k|_s) - \mathrm{dist}(L_t^j, L_u^k|_t)\right| \leqslant \gamma. \tag{4.2}$$

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

*Claim* 4.6. Fix $t, L$ and $\zeta$. 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 $L_t^{\pi_{s,t}(\pi_{t,u}(j))}|_s \overset{1-\gamma}{\approx} L_s^k$. By (4.2), distances are approximately preserved, so $L_u^j|_s \overset{1-2\gamma}{\approx} L_t^{\pi_{t,u}(j)}|_s$. Using the triangle inequality (where $L_t^{\pi_{t,u}(j)}|_s$ is the point in the middle) 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 $10\gamma$-separation of $\overline{L}(u)$, $k = k'$. $\square$

25

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 $2\gamma$-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 $2\gamma$-separation of $\overline{L}(s)|_u$, this implies that $j = j'$, i.e. the function is an injection. $\square$

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 $\exp(-\operatorname{poly}(\varepsilon)d_1)$ of the $s \subseteq t$ it holds that $|\operatorname{dist}(g, g') - \operatorname{dist}(g|_s, g'|_s)| > \varepsilon^{100}$. Let $A \subseteq t$ be the set of $v \in t$ such that $g(v) \neq g'(v)$. By Theorem 2.34 the fraction of $s \subseteq t$ such that $|\mathbb{P}[A] - \mathbb{P}_{v \subseteq s}[A]| > \varepsilon^{100}$ is $\exp(-\operatorname{poly}(\varepsilon)d_1)$. On the other hand, $\mathbb{P}[A] = \operatorname{dist}(g, g')$ and $\mathbb{P}_{v \subseteq s}[A] = \operatorname{dist}(g|_s, g'|_s)$ giving us the that at most $\binom{|L|}{2} \exp(-\operatorname{poly}(\varepsilon)d_1)$ 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 2.35 there is at most $\exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ of the $s \subseteq t$ such that

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

Another union bound gives us us the claim. $\square$

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

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

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

*Proof of Proposition 4.4.* By Claim 2.35 there are $1 - \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ faces $t \in X(d)$ such that $\operatorname{Agree}_\mathcal{D}(\mathcal{F}_t) \geqslant \varepsilon - \varepsilon^{100}$ (we use Claim 2.35 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

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

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

Moreover, by Claim 4.8 there exists some

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

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

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

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

1. $s \subseteq t$ samples $L'(t) \cup \overline{L}(t)$ well.

2. $L'(t)$ is $20\zeta$-dense in $\overline{L}_{\tau - \varepsilon^{100}, \zeta}(s)$.

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 $\tau, \zeta, \ell$ and $\gamma = 2\gamma'$. We take $\overline{L}(s) = \overline{L}(t)|_s$ for some arbitrary $t \in N_d$ 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 $\overline{L}(s) \subseteq \overline{L}_{\tau - \varepsilon^{100}, \zeta}(t)$. 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. $|\overline{L}(s)| = |\overline{L}(t)|_s| = \ell$).

Next we show density in $\overline{L}_{\tau - \varepsilon^{100}, \zeta}(s)$. Note that $L'(t)|_s$ is $20\zeta$-dense in $\overline{L}_{\tau - \varepsilon, \zeta}(s)$, and $\overline{L}(t)$ is $\gamma'$-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 $\gamma' + \varepsilon^{100} + 20\zeta \leqslant \gamma$-dense in $\overline{L}_{\tau - \varepsilon^{100}, \zeta}(s)$.

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 $2\gamma$-dense. $u \in M_{i'}$, therefore there is some $\overline{L}(u)$ of size $\ell$ that is $\gamma$-dense in $\overline{L}_{\tau - \varepsilon^{100}, \zeta}(u)$, and $11\gamma$-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 $2\gamma$-dense. By Claim 4.5, the fact that $\overline{L}(u)$ is dense in $\overline{L}(t)|_s$ and the fact that both are $2\gamma$-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 $\mathbb{P}[M_i] \geqslant 1 - \exp(-\Omega(\operatorname{poly}(\varepsilon)\frac{d_1}{k})$. Let us begin with $M_1$. Let $B_1$ be the fraction of $s$'s such that $\mathbb{P}_{t \supseteq s}[N_d] \leqslant \operatorname{poly}(\varepsilon)$. Its fraction is $\mathbb{P}[B_1] = O(\frac{d_1}{d \operatorname{poly}(\varepsilon)}) = \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ by Corollary 2.3. Let $B_2 \subseteq X(d_i)$ be the event that more than $\frac{1}{3}$ of the $t \supseteq s, t \in N_d$ have:

1. Either $s$ doesn't sample $\overline{L}(t) \cup L'(t)$ well.

2. $L'(t)$ is $20\zeta$-dense in $\overline{L}_{\tau - \varepsilon^{100}, \zeta}(s)$.

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

Continuing with $M_2$. The same arguments for $M_1$ show that $1 - \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$ of the $s \in X(d_2)$ sample $\overline{L}(t) \cup L'(t)$ well and that $L'(t)$ is $20\zeta$-dense in $\overline{L}_{\tau - \varepsilon^{100}, \zeta}(s)$. 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 $\exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$, albeit with a slightly worse $\operatorname{poly}(\varepsilon)$ than before, since since $\mathbb{P}[M_1] = \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$. Hence $\mathbb{P}[M_2] = 1 - \exp(-\operatorname{poly}(\varepsilon)\frac{d_1}{k})$. The proof for $M_3$ is similar. $\qquad\square$

*Proof of Claim 4.8.* Consider the following method, resembling [Din+18a]. Let $L_1 \subseteq \overline{L}_{\tau, \zeta}(t)$ be a maximal $24\zeta$-separated list. Consider the following chain $L_1 \supseteq L_2 \supseteq L_3 \ldots$ where every $L_i$ is a maximal $24^i\zeta$-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 $23\gamma_i = 24^{i+1}\zeta$-separated. Let us see that is it also $26 \cdot 24^i\zeta$-dense in $\overline{L}_{\tau, \zeta}(t)$. First note that every $L_j$ is $24^i\zeta$-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

$g \in \overline{L}_{\tau,\zeta}(t)$ there is a sequence $g = g_0, g_1, g_2, \ldots, g_i$ where $g_j \in L_j$ for $j > 0$, such that $\mathrm{dist}(g_{j-1}, g_j) \leqslant 24^j \zeta$. Thus concluding that

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

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

$\square$

*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 $3\zeta$-separated, for every distinct $g, g' \in L'$, the fraction of $r \in X(k)$ such that $\mathrm{dist}(g|_r, g'|_r) \leqslant 2\zeta$ (i.e. $g|_r \overset{1-2\zeta}{\approx} g'|_r$) is at most $\exp(-\Omega(\zeta^2 k)) \leqslant \tau^{10}$. This shows that there are almost no intersection between the supports of $g$ and $g'$, since $\mathrm{supp}_\zeta(g) \cap \mathrm{supp}_\zeta(g') \subseteq \{r \in X(k) \mid \mathrm{dist}(g|_r, g'|_r) \leqslant 2\zeta\}$. That is, $\mathbb{P}\left[\mathrm{supp}_\zeta(g) \cap \mathrm{supp}_\zeta(g')\right] \leqslant \exp(-\Omega(\zeta^2 k)) \leqslant \tau^{10}$.

Denote by

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

Then in particular

$$\mathbb{P}\left[D(g)\right] \geqslant \mathbb{P}\left[A_\zeta(g)\right] - \sum_{g' \in L', g' \neq g} \mathbb{P}\left[A_\zeta(g) \cap A_\zeta(g')\right] \tag{4.3}$$

$$\geqslant \mathbb{P}\left[A_\eta(g)\right] - \sum_{g' \in L', g' \neq g} 2\,\mathbb{P}\left[\mathrm{supp}_\zeta(g) \cap \mathrm{supp}_\zeta(g')\right] \tag{4.4}$$

$$\geqslant \tau - 2\tau^6. \tag{4.5}$$

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 $\overline{L}_{\tau^3,\zeta}(t)$ restricted to $s \subseteq t$ is dense in $\overline{L}_{\tau,\zeta}(s)$ for most $s$. This claim is where we use the agreement soundness of the original distribution.

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

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

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

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

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

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

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

28

Let us explain why. If there exists some $h \in \overline{L}_{\tau,\zeta}(s)$ such that $\mathrm{dist}(h,g) \geqslant \gamma + 4\zeta$ for all $g \in L|_s$, then $A_\zeta(h) \geqslant \tau$ and moreover it holds that $\mathbb{P}\left[\mathrm{supp}_\zeta(h) \cap \mathrm{supp}_{\gamma+2\zeta}(g)\right] = \exp(-\mathrm{poly}(\varepsilon)k) \ll \mathrm{poly}(\varepsilon)$ (since whenever $f_r \overset{1-\zeta}{\approx} h|_r$ and $f_r \overset{1-\gamma-2\zeta}{\approx} g|_r$ then it holds that $\mathrm{dist}(g|_r, h|_r) \leqslant \gamma + 3\zeta$ 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 $(r_1, r_2) \in A_\zeta(h)$ have that $r_1, r_2 \notin \bigcup_{g \in L} \mathrm{supp}_{\gamma+3\zeta+\varepsilon^{100}}(g)$. Thus $\mathbb{P}_{r_1,r_2 \subseteq s, r_1,r_2 \sim D}[P] \geqslant \frac{1}{2} A_\zeta(h) \geqslant \frac{1}{2}\tau$. On the other hand, by the first item $\mathbb{P}[P] \leqslant 2\tau^{1.5}$. By the second item $\mathbb{P}_{r_1,r_2 \subseteq s}[P] \leqslant \mathbb{P}[P] + \varepsilon^{100} \leqslant 3\tau^{1.5}$ 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 2.35. 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 $r \in \bigcup_{g \in L} \mathrm{supp}_{\gamma+3\zeta+\varepsilon^{100}}(g)$ (and taking $f'_r = f_r$ for the rest of the faces)[8]. 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 $A_\eta^{\mathcal{F}'}(h) \geqslant \tau^3$. On the other hand, there must by some $g \in L$ such that $\mathrm{dist}(h,g) \leqslant \gamma$. This implies that $\mathbb{P}\left[\mathrm{supp}_\eta^{\mathcal{F}'}(h) \setminus \mathrm{supp}_{\gamma+2\zeta}(g)\right] \leqslant \exp(-\mathrm{poly}(\zeta^2 k)) \ll \tau^3$. This is a contradiction since $A_\eta^{\mathcal{F}'}(h) \subseteq \mathrm{supp}_\eta^{\mathcal{F}'}(h) \setminus \mathrm{supp}_{3\zeta+\gamma}(g)$ (as we erasing $g$'s support). □

*Claim* (Restatement of 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 $\tau_m = \varepsilon^2(1-\varepsilon^{50})^m$ and $\zeta_m = 3\eta 20^m$. There exists $L \subseteq \overline{L}_{\tau_m,\zeta_m}(t)$ such that for $1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ of the $u \subseteq t, u \in X(d_1)$, $L|_u$ is $20\zeta_m$-dense in $\overline{L}_{\tau_{m+1},\zeta_m}(u)$.

*Proof of Claim 4.7.* The proof logic is the following. We construct a sequence of lists $L_1, L_2, \ldots$ where $L_1 \subseteq \overline{L}_{\tau_1,\zeta_1}(t)$, $L_2 \subseteq \overline{L}_{\tau_2,\zeta_2}(t)$ and so on. If $L_m \subseteq \overline{L}_{\tau_m,\zeta_m}(t)$ has the property required in the claim, i.e. that for $1 - \exp(-\mathrm{poly}(\varepsilon)\frac{d_1}{k})$ of the $u \subseteq t$, $L_m|_u$ is $20\zeta_m$ dense inside $\overline{L}_{\tau_{m+1},\zeta_m}(u)$ we conclude; else we continue to the next list $L_{m+1}$. What we need to make sure is that there is an $m \leqslant \frac{1}{\varepsilon^{10}}$ such that this process concludes with $L_m$. To do so, we define a potential function $p : \{1, 2, \ldots, \frac{1}{\varepsilon^{10}}\} \to [0,1]$ that is monotone non-decreasing (and $p \leqslant 1$), and show that if we don't conclude at step $m$, then $p(m+1) \geqslant p(m) + \varepsilon^{10}$, i.e. one of the first $\frac{1}{\varepsilon^{10}}$ first lists must satisfy the property else $p(\frac{1}{\varepsilon^{10}}) > 1$. The potential function is

$$p(m) = \underset{r_1,r_2 \subseteq t, r_1,r_2 \sim D}{\mathbb{P}}\left[\bigcup_{g_j \in L_m} A_{9\zeta_{m-1}}(g_j)\right]. \tag{4.6}$$

That is, $p(m)$ measures how much agreement is explainable by $L_m$, and we show that while the list $L_m$ is not dense for most $u \subseteq t$, then in fact we can find a list $L_{m+1}$ whose functions 'explains' more agreement. Details follow.

Recall that $\tau_m = \varepsilon^2(1-\varepsilon^{10})^m$ and that $\zeta_m = 3\eta \cdot 20^m$. Fix some $t$. For every $\tau_m, \zeta_m$ we choose some (arbitrary) $L_m^{aux}(t) \subseteq \overline{L}_{\tau_m^3,\zeta_m}(t)$ that is $3\zeta$-dense and $3\zeta$-separated. We will sequentially construct lists $L_m = \{g_1, g_2, \ldots, g_m\} \subseteq \overline{L}_{\tau_m,\zeta_m}(t)$, for $m = 1, 2, \ldots$ until $L_m$ satisfies the claim. The potential function $p$ shall be defined as in (4.6) with respect to these lists, once we construct them.

For $m = 0$ we take $L_0 = \emptyset \subseteq \overline{L}_{\tau_0,\zeta_0}(t)$. Let us describe how to construct $L_m$ using $L_{m-1}$, provided that $L_{m-1}$ is not $\zeta_m = 20\zeta_{m-1}$-dense in $\overline{L}_{\tau_m,\zeta_{m-1}}(u)$ 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 $20\zeta_{m-1}$-dense in $\overline{L}_{\tau_m,\zeta_{m-1}}(u)$.
   – $u$ samples $L_m^{aux} \cup L_{m-1}$ well.
   – $L_m^{aux}|_u$ is $7\zeta_m$-dense in $\overline{L}_{\tau_m,\zeta_{m-1}}(u)$.

---

[8]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 $\Sigma$ to a large enough set.
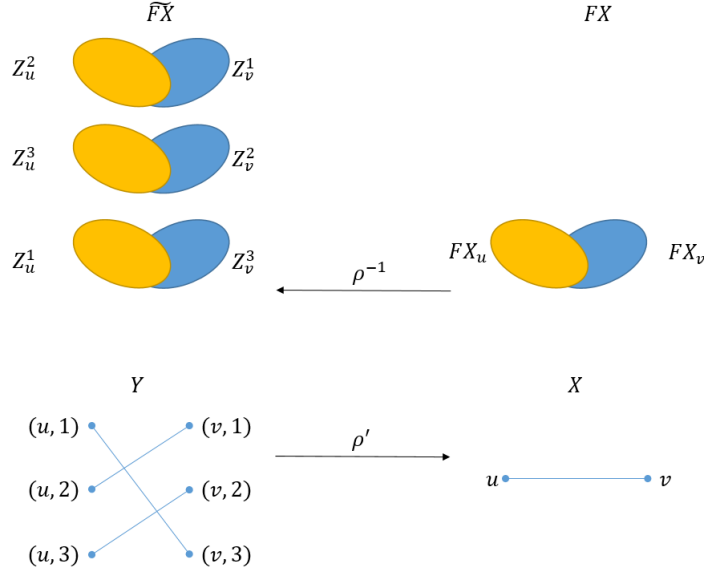
Figure 4: Constructing a cover of $X$ from a cover of $\mathrm{F}^{d_1}X$

   – For every $g \in L_m^{aux}(t)$ it holds that $\left| A_{9\zeta_{m-1}}^{\mathcal{F}_t}(g) - A_{9\zeta_{m-1}}^{\mathcal{F}_u}(g|_u) \right| \leqslant \varepsilon^{100}.$[9]

2. Find some $h \in \overline{L}_{\tau_m, \zeta_{m-1}}(u)$ that is $20\zeta_{m-1}$-far from every element in $L_{m-1}$. There exists such an $h$ since $L_{m-1}$ is not $20\zeta_{m-1}$-dense in $\overline{L}_{\tau_m, \zeta_{m-1}}(u)$.

3. Take some $g = g_m \in L_m^{aux}$ such that $\mathrm{dist}(g|_u, h_u) \leqslant 7\zeta_{m-1}$. Set $L_m = L_{m-1} \cup \{g_m\}$. This step is possible since $L_m^{aux}(t)|_u$ is $7\zeta_m$-dense inside $\overline{L}_{\tau_m, \zeta_m}(u)$.

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

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

That is, $g_m \in \overline{L}_{\tau_m, \zeta_m}(t)$ (and thus every $L_m \subseteq \overline{L}_{\tau_m, \zeta_m}(t)$).

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

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

# 5  Constructing a cover

In this subsection we prove Lemma 3.8. 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 $\iota : \mathrm{F}^{d_1}Y \to \widetilde{\mathrm{F}^{d_1}X}$. In

---

[9]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 2.35 and Claim 4.10. Hence if $L_m$ is not $20\zeta_{m-1}$-dense for for $1 - \exp(-\Omega(\mathrm{poly}(\varepsilon)\frac{d_1}{k}))$ we can find a $u \subseteq t$ such that this holds.

addition to proving Lemma 3.8, we will use the explicit definition of $\iota$ for proving Claim 3.9 later on.

Let $X$ be a $d$-dimensional clique complex. Let $d_1$ be such that $3d_1 \leqslant d - 2$. Let $\mathrm{F}^{d_1}X = F(X, d_1)$ be as in Definition 2.31 and assume that $X$ is well connected as in Definition 2.32. Fix $\nu : \widetilde{\mathrm{F}^{d_1}X} \to \mathrm{F}^{d_1}X$ to be an $\ell$-cover. In this section, we will show how to construct an $\ell$-cover of $\rho : Y \to X$ such that $\widetilde{\mathrm{F}^{d_1}X} \cong \mathrm{F}^{d_1}Y$.

The main idea will be to encode the vertices $v \in X(0)$ by the links $\mathrm{F}^{d_1}X_v \subseteq \mathrm{F}^{d_1}X$ (these are the $\mathrm{F}^{d_1}X_v = F(X_v, d_1)$ as in Definition 2.31 which are contained as sub complexes inside $\mathrm{F}^{d_1}X$). We observe that this encoding has the property that if $v \sim u$ then $\mathrm{F}^{d_1}X_v \cap \mathrm{F}^{d_1}X_u = \mathrm{F}^{d_1}X_{uv}$. We will use these non-empty intersections to define the edges in the cover $Y$. For every $r \in X$, let $Z_r = \nu^{-1}(\mathrm{F}^{d_1}X_r) \subset \widetilde{\mathrm{F}^{d_1}X}$. 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 $\ell$ connected-components[10] $Z_r^1, Z_r^2, ..., Z_r^\ell \subseteq \widetilde{\mathrm{F}^{d_1}X}$, such that $\nu$ is an isomorphism between each component and $\mathrm{F}^{d_1}X_r$, $\nu|_{Z_r^i} : Z_r^i \xrightarrow{\sim} \mathrm{F}^{d_1}X_r$.

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 $\pi_{r,s}(i) = j$.

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

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

where $\pi_{uv,u}, \pi_{uv,v}$ 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 $\{(v, i), (u, j)\} \in Y(1)$ if $vu \in X(1)$ and $j = \psi(vu).i$. 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 $\mathrm{F}^{d_1}X$ (see below). For this we recall the following corollary:

**Corollary** (Restatement of Corollary 2.30). *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 $\phi(u, uv) = \pi_{uv,u}$ and thus $\psi(uv) = \pi_{uv,u}^{-1} \circ \pi_{uv,v}$.

*Proof of Lemma 5.2.* Let us show that $Y$ is a cover. This is equivalent to showing that $\psi$ is a cocycle. Let $X^{\leqslant 2}$ be the two skeleton of $X$ and let $G$ be its flag-complex as in Definition 2.28. Let $\phi : GX(1) \to Sym(\ell)$ be given by $\phi(r, s) = \pi_{s,r}$ for every $r \subseteq s$ (where $\pi_{s,r}$ are the permutations in Claim 5.1). If we show that $\phi \in Z^1(GX, Sym(\ell))$ then by Corollary 2.30 $\psi$ is also a cocycle. This amounts to show that for every $uvw \in X$ it holds that $\pi_{uvw,uv} \circ \pi_{uv,u} = \pi_{uvw,u}$.

Indeed, fix $i$ and let $k = \pi_{uvw,u}(i)$, and $k' = \pi_{uvw,uv} \circ \pi_{uv,u}(i)$. 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 $\ell$ 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 $Z_r^i(0) := Z_v^i(0) \cap \nu^{-1}(\mathrm{F}^{d_1}X_r(0))$. As every $Z_v^i$ is isomorphic to $\mathrm{F}^{d_1}X_v$ then in particular, the induced sub-complex $Z_r^i$ is isomorphic to $\mathrm{F}^{d_1}X_r$. By well connectedness of $X$, $\mathrm{F}^{d_1}X_r$ is connected and thus we get that these are $\ell$ connected-components (they cannot be connected to one another because each lies in a different $Z_v^i$).

---

[10]When we say "connected components" we mean that the induced complex $Z_r$ has these $\ell$ connected components. Of course it may hold that there are paths from $Z_i$ to $Z_j$ in $\widetilde{\mathrm{F}^{d_1}X}$, but they must go through some vertex $v \in \widetilde{\mathrm{F}^{d_1}X} \setminus Z_r$.
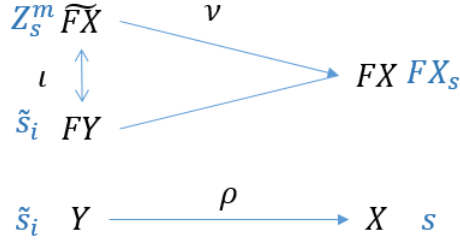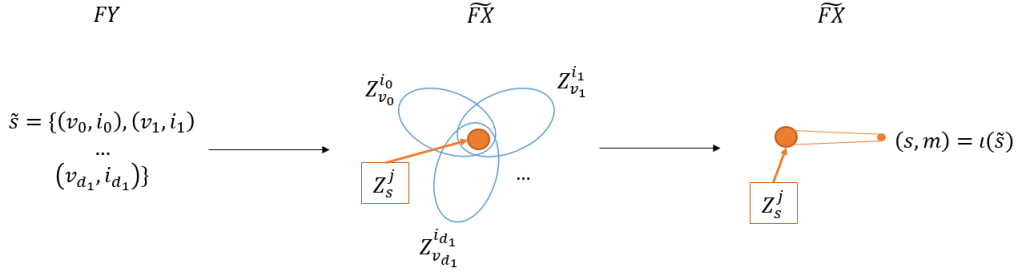
Figure 5: Commutative diagram



Figure 6: From $\tilde{s}$ to $\iota(\tilde{s}) = (s, m)$

Now we show the first item for every vertex $v \in X(0)$. Towards this end, we note that the restriction $\nu|_{Z_v} : Z_v \to \mathrm{F}^{d_1} X_v$ is also a covering map (for a proof of this statement, see Claim 2.22). From the well connectivity of $X$, $\mathrm{F}^{d_1} X_v$ is simply connected, so $Z_v$ must decompose to $\ell$ 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 the first item, the restriction $\nu|_{Z_r^i}$ is an isomorphism between $Z_r^i$ and $\mathrm{F}^{d_1} X_r$. In particular, it follows from the definition of an isomorphism that $\nu^{-1}(\mathrm{F}^{d_1} X_s) \cap Z_r^i$ is isomorphic (by $\nu|_{Z_r^i}$) to $\mathrm{F}^{d_1} X_s \subset \mathrm{F}^{d_1} X_r$, and thus $Z_s^j$ is one of the connected components $Z_s^j \subseteq \nu^{-1}(\mathrm{F}^{d_1} X_s)$ promised by the first item we already proved. This is the index $j$ we needed. Note that it is unique since $Z_s^j = \nu^{-1}(\mathrm{F}^{d_1} X_s) \cap Z_r^i$ (so there are no more faces in $\nu^{-1}(\mathrm{F}^{d_1} X_s)$ and $Z_r^i$ that are not already in $Z_s^j$). $\quad\square$

## 5.1 The isomorphism between $\mathrm{F}^{d_1}Y$ to $\widetilde{\mathrm{F}^{d_1}X}$

We continue with the notation in the previous subsection. In this subsection we describe the isomorphism $\iota : \mathrm{F}^{d_1}Y \to \widetilde{\mathrm{F}^{d_1}X}$ explicitly. See Figure 5 for the relations between $X, Y, \mathrm{F}^{d_1}X, \mathrm{F}^{d_1}Y$ and $\widetilde{\mathrm{F}^{d_1}X}$.

The idea is simple. Given some $\tilde{s} = \{(v_0, i_0), (v_1, i_1), ..., (v_{d_1}, i_{d_1})\} \in \mathrm{F}^{d_1}Y(0)$, such that $\rho(\tilde{s}) = s = \{v_0, v_1, \ldots, v_{d_1}\}$, 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 $\mathrm{F}^{d_1} X_s$. We will define $\iota(\tilde{s}) = (s, m)$ which is connected to this copy. See Figure 6 for a pictorial description of $\iota$. We first claim that $\iota$ is well defined.

*Claim* 5.3. Let $\tilde{s} \in \mathrm{F}^{d_1}Y(0)$, $\tilde{s} = \{(v_0, i_0), (v_1, i_1), ..., (v_{d_1}, i_{d_1})\}$ 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.

**Lemma 5.4.** $\iota : \mathrm{F}^{d_1}Y \to \widetilde{\mathrm{F}^{d_1}X}$ *is an isomorphism.*

*Proof of Lemma 5.4.* Let us start by explaining why this is a bijection. First note that for every $s \in \mathrm{F}^{d_1}X(0)$, $\iota$ maps the $\ell$-elements of $\rho^{-1}(s) \subseteq \mathrm{F}^{d_1}Y(0)$ to the $\ell$-elements of $\nu^{-1}(s) \subseteq \widetilde{\mathrm{F}^{d_1}X}$. Thus if we show that $\iota$ 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

32

$\iota(\tilde{s}_i) = (s, m_i)$. We'll show that $m_1 \neq m_2$. Take some vertex $v \in X(0)$ in $s \in \mathrm{F}^{d_1}X(0)$, 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 $\mathrm{F}^{d_1}Y$ and $\widetilde{\mathrm{F}^{d_1}X}$ 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 $\mathrm{F}^{d_1}Y$ of a clique complex is also a clique complex.

2. $\mathrm{F}^{d_1}X$ is a clique complex since it is a face complex of a clique complex. By Claim 2.23 the cover $\widetilde{\mathrm{F}^{d_1}X}$ is also a clique complex.

Thus it is enough to show that the bijection $\iota$ is a graph isomorphism between $\mathrm{F}^{d_1}Y^{\leqslant 1}$ to $(\widetilde{\mathrm{F}^{d_1}X})^{\leqslant 1}$.

Let us begin by showing that $\iota$ is a graph homomorphism. Let $\{\tilde{s}_1, \tilde{s}_2\} \in \mathrm{F}^{d_1}Y(1)$ 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 $\{(s_1, j_1), (s_2, j_2)\} \in \widetilde{\mathrm{F}^{d_1}X}(1)$. Since $\{\tilde{s}_1, \tilde{s}_2\} \in \mathrm{F}^{d_1}Y(1)$, then in particular they are contained in some triangle $\{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3\} \in \mathrm{F}^{d_1}Y(2)$[11] and denote by $\nu(\tilde{s}_3) = s_3$. Let $Z_{s_1}^{j_1}, Z_{s_2}^{j_2}$ be the copies of $\mathrm{F}^{d_1}X_{s_1}, \mathrm{F}^{d_1}X_{s_2}$ 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 $\iota$). If we show that the same copy of $s_3$ is in the intersection of $Z_{s_1}^{j_1}, Z_{s_1}^{j_2} \subseteq \widetilde{\mathrm{F}^{d_1}X}$, namely that there is some $(s_3, k) \in Z_{s_1}^{j_1} \cap Z_{s_2}^{j_2}$ this will imply that $\{(s_1, j_1), (s_2, j_2)\} \in \widetilde{\mathrm{F}^{d_1}X}(1)$ (because this implies that both $(s_1, j_1), (s_2, j_2)$ are in the link of $(s_3, k) \in \widetilde{\mathrm{F}^{d_1}X}$. This link is isomorphic by $\nu$ to the link of $s_3 \in \mathrm{F}^{d_1}X$. There we know that $\{s_1, s_2\}$ is an edge, so in particular $\{(s_1, j_1), (s_2, j_2)\} \in \widetilde{\mathrm{F}^{d_1}X}(1)$).

Take two vertices $(v_1, k_1), (v_2, k_2) \in Y$ such that $(v_i, k_i) \in \tilde{s}_i$. Because $\{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3\} \in \mathrm{F}^{d_1}Y(2)$ implies that $\tilde{s}_1 \cup \tilde{s}_2 \cup \tilde{s}_3 \in Y$. This means that

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

2. $s_3 \in \mathrm{F}^{d_1}X_{v_1 v_2}$.

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 $s_3 \in \mathrm{F}^{d_1}X_{v_1 v_2}$ 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 $\iota$, it holds that $Z_{s_1}^{j_1} \subseteq Z_{v_1}^{k_1}$. As $Z_{v_1}^{k_1}$ is isomorphic by $\nu$ to $\mathrm{F}^{d_1}X_{v_1}$ and both $(s_3, k), (s_3, k') \in Z_{v_1}^{k_1}$ are sent to $s_3$ by $\nu$, 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 $\widetilde{\mathrm{F}^{d_1}X}$ and $\mathrm{F}^{d_1}Y$ are $\ell$-covers of $\mathrm{F}^{d_1}X$. In particular, the degree of every $\tilde{s} \in \mathrm{F}^{d_1}Y$ is equal to the degree of $\iota(\tilde{s}) \in \widetilde{\mathrm{F}^{d_1}X}$ (both are equal to the degree of $\rho(s)$ in $\mathrm{F}^{d_1}X$). Any bijective graph isomorphism that preserves degrees must be an isomorphism. $\qquad\square$

*Proof of Claim 5.3.* Let $j$ be such that $Z_s^j \subseteq Z_{v_0}^{i_0}$ (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 $Z_{v_0 v_m}^k \subseteq Z_{v_0}^{i_0} \cap Z_{v_m}^{i_m}$ (there is such a $k$ by Lemma 5.2). Then by the fact that $\mathrm{F}^{d_1}X_s \subseteq \mathrm{F}^{d_1}X_{v_0 v_m} \subseteq \mathrm{F}^{d_1}X_{v_0}$ and $\mathrm{F}^{d_1}X_s \subseteq \mathrm{F}^{d_1}X_{v_0 v_m} \subseteq \mathrm{F}^{d_1}X_{v_m}$, it holds by the isomorphism that $\nu$ induces that $Z_s^j \subseteq Z_{v_0 v_m}^k \subseteq Z_{v_0}^{i_0}$, and 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 $Z_s^j \subseteq Z_{v_0}^{i_0}$, and $\bigcap_{m=0}^{d_1} Z_{v_m}^{i_m} \subseteq Z_{v_0}^{i_0}$ so there must be only a single such $j$. $\qquad\square$

---

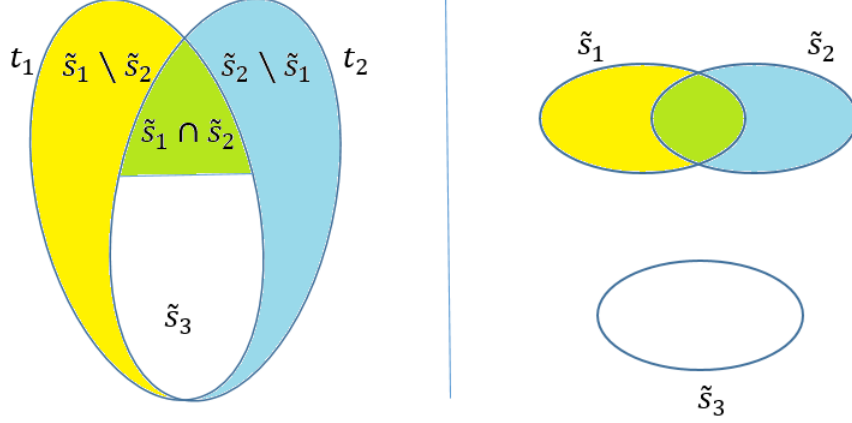[11]$Y$ is pure so $\mathrm{F}^{d_1}Y$ is also pure.

Figure 7: 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 $\tilde{s_2}$).

# 6 A global function on $Y$

We remind the reader of our convention that $\pi_{x,y} = \pi_{y,x}^{-1}$. Let us rephrase Corollary 3.7.

**Lemma 6.1.** *There exists a cocycle $\psi \in Z^1(X, Sym(\ell))$ such that $\mathbb{P}_{s_1,s_2,t=s_1 \cup s_2}\left[\psi(s_1, s_2) = \pi_{t,s_2}^{-1} \circ \pi_{t_1,s_1}\right] = 1 - \exp(-\Omega(d_1/k))$, where $\pi_{t,s_i}$ are the permutations over the list of functions of $s_i$ and $t$ respectively as in Lemma 3.5.*

Let us denote the $\ell$-cover that this cocycle induces by $\rho : Y \to X$ such that we have an isomorphism $\iota : \mathrm{F}^{d_1}Y \to \widetilde{\mathrm{F}^{d_1}X}$. 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 3.9). There exists a universal constant $c > 0$ such that $\mathbb{P}_{\tilde{s} \in Y(d_1)}\left[h_{\tilde{s}} \overset{1-O(\gamma)}{\approx} G|_{\tilde{s}}\right] \geq 1 - \exp(-\operatorname{poly}(\varepsilon)\left(\frac{d_1}{k}\right)^c) - \frac{1}{d_1^c}$.

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

**Theorem 6.2** ([DH23, Claim 8.5]). *There exists a universal constant $c > 0$ such that the following holds. Let $Y$ be a $\frac{1}{d_1^3}$-one sided high dimensional expander. 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*

$$\underset{\tilde{s_1},\tilde{s_2} \sim DU_n}{\mathbb{P}}\left[h_{\tilde{s_1}}|_{\tilde{s_1} \cap \tilde{s_2}} \overset{1-\eta}{\approx} h_{\tilde{s_2}}|_{\tilde{s_1} \cap \tilde{s_2}}\right] \geq 1 - \gamma,$$

*then*

$$\underset{\tilde{s} \in Y(d_1)}{\mathbb{P}}\left[G|_{\tilde{s}} \overset{1-8(\eta+\gamma)}{\approx} h_{\tilde{s}}\right] \geq 1 - \frac{1}{d_1^c} - \gamma^c.$$

*Proof of Claim 3.9.* We want to use Theorem 6.2, and thus we need to show that

$$\underset{\tilde{s_1},\tilde{s_2} \sim DU_n}{\mathbb{P}}\left[h_{\tilde{s_1}}|_{\tilde{s_1} \cap \tilde{s_2}} \overset{1-O(\eta)}{\approx} h_{\tilde{s_2}}|_{\tilde{s_1} \cap \tilde{s_2}}\right] \geq 1 - \exp(-\Omega(d_1/k)),$$

Let us define the distribution $(\tilde{s_1}, \tilde{s_2}, \tilde{s_3}) \sim P$ where:

34

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. $\tilde{s_3} \uplus (\tilde{s_1} \cup \tilde{s_2}) \in X$, i.e. $\tilde{s_3}$ is sampled via a swap walk step from $(\tilde{s_1} \cup \tilde{s_2})$.

Let us denote by $\tilde{t_i} = \tilde{s_i} \uplus \tilde{s_3}$ for $i = 1, 2$, and let us denote $\tilde{r} = \tilde{s_1} \cap \tilde{s_2}$. See Figure 7 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 $j_i' = \pi_{t_i, s_i}(j_i)$. We define the following two "bad" events

1. $E = \{L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}|_{s_i} \text{ or } L_{s3}^{j3} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}|_{s3} \text{ for } i = 1 \text{ or } i = 2\}$.

2. $F = \{L_{t_1}^{j_1'}|_{s3} \overset{1-2\eta}{\approx} L_{t_2}^{j_2'}|_{s3} \text{ and } L_{t_1}^{j_1'}|_r \overset{1-9\eta}{\not\approx} L_{t_2}^{j_2'}|_r\}$.

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 $L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}|_{s_i}$, 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 $L_{s3}^{j3} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}|_{s3}$ 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 $j_i' = \pi_{t, s_i}(j_i)$. 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, $L_{t_1}^{j_1'}|_{s3} \overset{1-\eta}{\approx} L_{s3}^{j3} \overset{1-\eta}{\approx} L_{t_2}^{j_2'}|_{s3}$ and hence $L_{t_1}^{j_1'}|_{s3} \overset{1-2\eta}{\approx} L_{t_2}^{j_2'}|_{s3}$. 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. $\qquad\square$

*Proof of Claim 6.3.* Let us begin with bounding the probability that $L_{s_i}^{j_i} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}|_{s_i}$ 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

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

by Lemma 3.5. When this occurs then $L_{s_i}^x \overset{1-\eta}{\approx} L_{t_i}^{\pi_{t_i, s_i}(x)}|_{s_i}$ for all indexes $x$, and in particular for $j_i$.

We move towards bounding the probability that $L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{j_i'}|_{s_3}$ 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 $\mathbb{P}\left[ L_{s_3}^{j_3} \overset{1-\eta}{\not\approx} L_{t_i}^{\pi_{t_i,s_3}(j_3)} \right] = \exp(-\Omega(d_1/k))$.

Moreover, by Lemma 6.1, with probability $1 - \exp(-\Omega(d_1/k))$ it holds that $\psi(s_i, s_3) = \pi_{t_i,\tilde{s}_3}^{-1} \circ \pi_{t_i,\tilde{s}_i}$ (where $\psi$ is the cocycle that defined the cover $\rho$). Recall that $\iota(\tilde{s}_i) = (s_i, j_i)$ where $\iota : \mathrm{F}^{d_1}Y \to \widetilde{\mathrm{F}^{d_1}X}$ and $\tilde{s}_i, \tilde{s}_3 \in \mathrm{F}^{d_1}Y(0)$ are neighbors in $\mathrm{F}^{d_1}Y$ which implies that for $i = 1, 2$, $(s_i, j_i) \sim (s_3, \psi(s_i, s_3).j_i)$ are neighbors in $\widetilde{\mathrm{F}^{d_1}X}$. Thus with probability $1 - \exp(-\Omega(d_1/k))$,

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

for $i = 1, 2$. When this occurs, then
$$\pi_{t_i,s_3}(j_3) = \pi_{t_i,s_i}(j_i) = j_i',$$

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

$\square$

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

$$\underset{\tilde{s}_3, r}{\mathbb{P}} \left[ f|_{s_3} \overset{1-2\eta}{\approx} g|_{s_3} \text{ and } f|_r \overset{1-9\eta}{\not\approx} g|_r \right] \leqslant \exp(-\Omega(\eta^2|s|))$$

which proves the claim since $|s| = \Omega(d_1)$ and $\eta$ 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.

$\square$

# 7 Covers and Counterexamples

## 7.1 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 $\zeta < \frac{1}{2}$, $k \in \mathbb{N}$, and let $\lambda \leqslant \frac{0.99}{k}$. Let $X$ be a $k$-dimensional $\lambda$-two-sided high dimensional expander, and assume $X$ has a connected $\ell$-cover for some $\ell > 1$ (e.g. $\ell = 2$). Then there exists an ensemble of functions $\mathcal{F} = \{ f_r : r \to \{0, 1\} \mid r \in X(k) \}$ such that $Agree(\mathcal{F}) \geqslant \frac{1}{\ell}$, and yet for every $G : X(0) \to \{0, 1\}$ it holds that*

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

*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 $\rho$, 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 $\mathbb{E}_{\mathcal{F}}\left[\text{Agree}_{\mathcal{D}}(\mathcal{F})\right] \geqslant \frac{1}{2}$.

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 $\mathbb{E}_{\mathcal{F}}\left[\text{Agree}_{\mathcal{D}}(\mathcal{F})\right] \geqslant \frac{1}{2}$. The expectation is over the choice of preimages $\tilde{r}$ for every $r \in X(k)$.

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

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 $\mathbb{P}_{\mathcal{F}}\left[f_r = f_{r'}\right] \geqslant \frac{1}{2}$.

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 $\mathbb{P}_{\mathcal{F}}\left[f_r = f_{r'}\right] \geqslant \frac{1}{2}$.

Thus we take $\mathcal{F}$ be any such assignment where $\text{Agree}_{\mathcal{D}}(\mathcal{F}) \geqslant \frac{1}{2}$ (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

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

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 $G|_r \overset{1-\zeta}{\approx} f_r$ and $f_r$ answers according to the preimage $\tilde{r}$, then $H|_{\tilde{r}} \overset{1-\zeta}{\approx} \tilde{G}|_{\tilde{r}}$. 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

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

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 $\mathbb{P}_{\tilde{r} \in Y(k)}\left[\tilde{G}|_{\tilde{r}} \overset{1-\zeta}{\approx} H|_{\tilde{r}}\right] = \exp(-\Omega_\zeta(k))$.

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

Thus by Theorem 2.34, it holds that

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

and by (7.1),

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

$\square$

*Remark* 7.1. A similar argument can generalized for complexes $X$ with connected $\ell$-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

$$\Pr_{r \in X(k)} \left[ G|_r \overset{1-\zeta}{\approx} g_r \right] = \exp(-\Omega_\zeta(k))$$

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

## 7.2 Gap amplification fails below $1/2$

In [Bog05], Bogdanov addresses the question of whether the soundness of a PCP can be reduced to any $\varepsilon > 0$ by a step of gap amplification a la [Din07]. The starting point is a constraint graph $\mathcal{G} = (V, E, \{\Sigma_v\}, \{\pi_{uv}\})$ where $V$ is a set of vertices, $E$ a set of edges, $\Sigma_v$ a finite set of labels to be assigned to $v$ and $\pi_{uv} \subset \Sigma_u \times \Sigma_v$ a constraint specifying which pairs of labels are considered satisfying. The value of $\mathcal{G}$ is the maximal fraction of satisfiable constraints, namely

$$val(\mathcal{G}) = \max_{H:V \to \{0,1\}} \Pr_{uv \in E}[(H(u), H(v)) \in \pi_{uv}].$$

Our initial assumption is that it is NP-hard to decide if $val(\mathcal{G}) = 1$ or $val(\mathcal{G}) < 1 - \varepsilon_0$.

The gap amplification, or "powering" step (with parameter $\ell \in \mathbb{N}$), constructs a family $X$ of subsets of $V$ corresponding to balls in $\mathcal{G}$ as follows. For each $v \in V$, we define $s_v = \{u \in V \mid \operatorname{dist}(u, v) \leqslant \ell\}$. We let $X = \{s_v \mid v \in V\}$. Furthermore, for a set $s_v$ we will only allow local functions $f_{s_v} : s_v \to \{0,1\}$ that satisfy all $\mathcal{G}$ constraints inside $s_v$. Formally, let

$$\Sigma_v = \{f_v : s_v \to \{0,1\} \mid f_v \text{ satisfies all } \mathcal{G} \text{ constraints in } s_v\}. \tag{7.2}$$

A valid ensemble of local functions on $X$ is $\{f_{s_v} : s_v \to \{0,1\} \mid f_{s_v} \in \Sigma_v\}$. An agreement test for $X$ specifies a distribution over pairs of sets in $X$, but we do not specify the precise distribution because the counterexample works for any distribution.

– The completeness of this transformation is that whenever there is an assignment satisfying all of the constraints of $\mathcal{G}$ then there is a valid ensemble of local functions $\{f_s\}$ with $\operatorname{Agree}(\{f_s\}) = 1$.

– The soundness of this transformation is that whenever there is a valid ensemble $\{f_s\}$ with $\operatorname{Agree}(\{f_s\}) > \varepsilon$ there must be some global $H : V \to \{0,1\}$ that satisfies more than $1 - \varepsilon_0$ of the constraints of $\mathcal{G}$, namely, $val(\mathcal{G}) > 1 - \varepsilon_0$ which means, by our assumption, that $val(\mathcal{G}) = 1$.

The "soundness" of the gap amplification transformation is the value of $\varepsilon$ for which we can guarantee that

$$\operatorname{Agree}(\{f_s\}) > \varepsilon \implies val(\mathcal{G}) > 1 - \varepsilon_0.$$

Contrapositively, a counterexample to gap amplification is an ensemble that has high agreement yet $val(\mathcal{G}) < 1 - \varepsilon_0$.

In Bogdanov's example, the graph $\mathcal{G}$ is chosen to be a high-girth expander graph, the alphabet is $\Sigma_v = \{0, 1\}$ for all $v \in V$, and all edges carry inequality constraints, namely $\pi_{uv} = \{(0,1), (1,0)\}$ for all $uv \in E$. The expansion asserts that $val(\mathcal{G}) \approx 1/2$ since every cut in the graph violates about half of the edges or more.

Looking at valid ensembles, we observe that the set $\Sigma_v$ of possible local functions for a given $s_v$ consists of only two assignments, per (7.2). One where $v \in s_v$ is given 0 and one where $v \in s_v$ is given 1. The unique constraints along the edges propagate the values uniquely from $v$ to the rest of $s_v$. Bogdanov chooses an ensemble at random by choosing, for each $s_v$ one of the two available local functions independently at random. Clearly, for each edge separately, there is probability $1/2$ of choosing a matching pair, so the overall local agreement is $\approx 1/2$, whereas, with high probability there is no global assignment agreeing with even an $\varepsilon$ faction of labels, assuming $\mathcal{G}$ is a good expander.

Now let us recast this in terms of covers (and then generalize). We first construct the double cover $\mathcal{G}'$ of $\mathcal{G}$. Let $\mathcal{G}'$ be a bipartite graph with vertices $V' = V \times \{0,1\}$ and edges $E' = \{\{(u,0),(v,1)\} \mid uv \in E\}$ (and still putting inequality constraints on all edges).

Observe that this is a 2-cover a la Definition 2.12. Moreover, since $\mathcal{G}'$ is bipartite, we actually have $val(\mathcal{G}) = 1$. This is because all constraints are satisfied by the global assignment $H : V' \to \{0,1\}$ defined by $H(v,b) = b$.

Now we follow the recipe of the proof of Lemma 1.2 (see Figure).

$$
\begin{array}{ccc}
Y & \xleftarrow{\ powering\ } & \mathcal{G}' \\
\downarrow{\scriptstyle \rho} & & \downarrow{\scriptstyle \rho} \\
X & \xleftarrow{\ powering\ } & \mathcal{G}
\end{array}
$$

Assuming that the girth of $\mathcal{G}$ is larger than $\ell$, the family $Y = \left\{ \tilde{s}_{(v,b)} \,\middle|\, (v,b) \in V' \right\}$ is a 2-cover of the family $X$ in that there is a 2-to-1 map $\rho : V' \to V$ such that the image of every $\tilde{s} \in Y$ is some $s \in X$, and the preimage of each $s \in X$ is two disjoint sets $\tilde{s}_1, \tilde{s}_2 \in Y$[12]. Specifically $\rho(\tilde{s}_{(v,b)}) = s_v$, and $\rho^{-1}(s_v) = \{\tilde{s}_{(v,0)}, \tilde{s}_{(v,1)}\}$. From the global assignment $H : V' \to \{0,1\}$ we can construct a perfectly agreeing ensemble (on $Y$) by setting $f_{\tilde{s}_{(v,b)}} = H|_{\tilde{s}_{(v,b)}}$.

We choose, as in the proof of Lemma 1.2, for each $s_v \in X$ an assignment let

$$
f_{s_v} = \begin{cases} f_{\tilde{s}_{(v,0)}}, & \text{with probability } 1/2 \\ f_{\tilde{s}_{(v,1)}}, & \text{with probability } 1/2 \end{cases}.
$$

Finally, observe that this is exactly the same assignment as in the direct description of Bogdanov's counterexample.

**Discussion**

One possible conclusion of this section is that the failure of gap amplification to go below $1/2$ is due to the graph-based nature of the family $X$ which allows existence of connected covers with perfect assignments. For a better gap amplification we need to start with PCP instances $\mathcal{G}$ that are free of such covers.

An additional observation is that Bogdanov's example generalizes to any unique games instance $\mathcal{G}$, where the size of the alphabet $t = |\Sigma|$ dictates the size of the $t$-to-1 cover. Even for $\mathcal{G}$ whose constraints are not unique, one can easily sparsify them into unique constraints and continue the construction as before. So without revising the powering construction itself, no choice of $\mathcal{G}$ can help.

# References

[AG22]    Dorna Abdolazimi and Shayan Oveis Gharan. "An Improved Trickle-Down Theorem for Partite Complexes." In: *arXiv preprint arXiv:2208.04486* (2022).

[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.

[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.

[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.

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

---

[12]This is also formally a cover, only that $X, Y$ are cell complexes, not simplicial complexes.

[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.

[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.

[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. "Swap cosystolic expansion." 2023. eprint: `arXiv:2312.15325`.

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

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

[Dik+18]    Yotam Dikstein, Irit Dinur, Yuval Filmus, and Prahladh Harsha. "Boolean function analysis on high-dimensional expanders." In: *Proc. 20th 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`.

[Din07]     Irit Dinur. "The PCP theorem by gap amplification." In: *J. ACM* 54.3 (2007), p. 12. DOI: `10.1145/1236457.1236459`. URL: `https://doi.org/10.1145/1236457.1236459`.

[DK17]      Irit Dinur and Tali Kaufman. "High dimensional expanders imply agreement expanders." In: *Proc. 58th 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.

[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`.

[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`.

[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`.

[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. 48th ACM Symp. on Theory of Computing (STOC)*. 2016, pp. 36–48. DOI: 10.1145/2897518.2897543. eprint: 1510.00839.

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

[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.

[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.

[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.

[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.

[KO18]     Tali Kaufman and Izhar Oppenheim. "Construction of new local spectral high dimensional expanders." In: *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*. 2018, pp. 773–786.

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

[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.

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

[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.

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

[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.

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