COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **38** (3) (2018) 547–599
DOI: 10.1007/s00493-016-3383-0

STRONG INAPPROXIMABILITY RESULTS ON BALANCED RAINBOW-COLORABLE HYPERGRAPHS

VENKATESAN GURUSWAMI*, EUIWOONG LEE†

Received January 12, 2015
Revised September 21, 2015
Online First December 14, 2017

Consider a K-uniform hypergraph $H=(V,E)$. A coloring $c\colon V\to\{1,2,\ldots,k\}$ with k colors is *rainbow* if every hyperedge e contains at least one vertex from each color, and is called *perfectly balanced* when each color appears the same number of times. A simple polynomial-time algorithm finds a 2-coloring if H admits a perfectly balanced rainbow k-coloring. For a hypergraph that admits an *almost balanced rainbow* coloring, we prove that it is NP-hard to find an independent set of size ϵ, for any $\epsilon>0$. Consequently, we cannot *weakly color* (avoiding monochromatic hyperedges) it with $O(1)$ colors. With $k=2$, it implies strong hardness for discrepancy minimization of systems of bounded set-size.

One of our main technical tools is based on *reverse hypercontractivity*. Roughly, it says the *noise operator* increases q-norm of a function when $q<1$, which is enough for some special cases of our results. To prove the full results, we generalize the reverse hypercontractivity to more general operators, which might be of independent interest. Together with the generalized reverse hypercontractivity and recent developments in inapproximability based on invariance principles for correlated spaces, we give a *recipe* for converting a promising test distribution and a suitable choice of an outer PCP to hardness of finding an independent set in the presence of highly-structured colorings. We use this recipe to prove additional results almost in a black-box manner, including: (1) the first analytic proof of $(K-1-\epsilon)$-hardness of K-Hypergraph Vertex Cover with more structure in completeness, and (2) hardness of $(2Q+1)$-SAT when the input clause is promised to have an assignment where every clause has at least Q true literals.

Mathematics Subject Classification (2000): 05C15, 68Q15, 05C65, 60E15

* Supported in part by NSF grant CCF-1115525. Most of this work was done while visiting Microsoft Research New England.

† Supported by a Samsung Fellowship, US-Israel BSF grant 2008293, and NSF CCF-1115525. Most of this work was done while visiting Microsoft Research New England.

1. Introduction

The problem of coloring a hypergraph with few colors is a fundamental optimization problem. A K-uniform hypergraph $H = (V, E)$ is said to be k-colorable if there exists a coloring $c\colon V \to \{1,\dots,k\}$ of its vertices with k colors so that no hyperedge is monochromatic.

The problem of determining if a K-uniform hypergraph is 2-colorable is a classic NP-hard problem when $K \geq 3$. By now, strong inapproximability results are known which show that coloring 2-colorable hypergraphs with any fixed constant number of colors is NP-hard – this was first shown for 4-uniform hypergraphs [15,18] and subsequently also for the 3-uniform case [12]. The best known algorithmic results require $n^{\Omega(1)}$ colors, with the exponent tending to 1 as the uniformity k of the hypergraph increases [8,1]. Recently, even coloring 2-colorable hypergraphs with super-polylogarithmically many colors was shown to be hard (for the 8-uniform case) [9,14]. This situation contrasts with graphs ($K = 2$) where it is not known to be hard to color 3-colorable graphs with just 5 colors unless we assume much stronger conjectures [11].

In this work, we are interested in the question of whether coloring a hypergraph remains hard even if we are promised that the hypergraph admits a coloring with natural stronger properties. One such notion, called strong k-colorability, insists that for each hyperedge, all its vertices get different colors. Note that in the case of graphs ($K = 2$), the notions of colorability and strong colorability coincide. Strong coloring of a K-uniform hypergraph $H = (V, E)$ is the same as coloring the graph $G = (V, E')$ with the same vertex set and $E' = \{(u,v)\colon \exists e \in E$ such that $\{u,v\} \subseteq e\}$ (i.e., we make each hyperedge into a K-clique). The minimum possible number of colors needed to strongly color a K-uniform hypergraph is of course K. It is not hard to see that given a strongly K-colorable K-uniform hypergraph H, one can efficiently find a 2-coloring of its vertices such that no hyperedge is monochromatic. See Section 1.2 for more discussion.

There are two natural notions which are weaker than strong colorability but yet impose richer requirements on the coloring than just avoiding monochromatic edges:

- Rainbow k-coloring: Every hyperedge contains a vertex of each of the k colors.
- Balanced/Low-discrepancy 2-coloring: In every hyperedge, there are a roughly equal number of vertices of each of the two colors.

Note that rainbow 2-coloring is the same as normal 2-coloring, and the existence of a rainbow k-coloring for $k \geq 2$ implies that the hypergraph is

2-colorable. We can combine the above two notions and require that every hyperedge has to have roughly the same number of vertices of each color.

Both the above notions of coloring have been studied before. For rainbow k-coloring, it is known as *polychromatic* coloring where the basic question is: given a certain family of hypergraphs (often interpreted as set systems representing geometric objects), what is the smallest K that guarantees rainbow k-coloring? We refer to the recent work of Bollobás et al. [5] and references therein. Finding a good balanced 2-coloring is known as minimizing *discrepancy*, where the ideas of semidefinite programming [3] and random walks [22] have been successfully applied. There are tight hardness results for general hypergraphs ([7], no constraint on the size of edges) and r-uniform hypergraphs [2], where a hypergraph is not 2-colorable in the soundness case. Our goal is to show that a hypergraph is not $O(1)$-colorable in the soundness case.

Our main result in this work is to prove a strong hardness result that rules out coloring a hypergraph with $O(1)$ colors even when it is promised to have a rainbow k-coloring with good balance between colors (for any $k \geq 3$) – see Theorem 1.1 below for a formal statement. It is worth emphasizing that prior to this work, even hardness of 2-coloring a rainbow 3-colorable hypergraph was not known. Indeed such a result seemed out of reach of the sort of Fourier-based PCP techniques used for hardness of hypergraph coloring in [15] and follow-ups. In this work we leverage invariance principle based techniques to analyze test distributions that ensure balanced rainbow colorability (further details about our methods and those in recent technically related works appears in Section 2). One of our contributions is to distill a general recipe for combining test distributions with suitable outer PCPs (various forms of smooth Label Cover) to establish such inapproximability results. This makes our approach quite flexible and can also be readily applied to several other problems as described in Section 1.1.

1.1. Our results and corollaries

Given a hypergraph $G=(V,E)$ and a subset $I \subseteq V$, we say that I *induces* a hyperedge e when $e \subseteq I$. The hypergraph induced by I is (I, E_I) where E_I is the set of hyperedges induced by I.

The following is our main theorem. Note that in any result in this section that guarantees a coloring with some desired properties in the completeness case, each color contains the same fraction of vertices.

Theorem 1.1. *For any $\epsilon > 0$ and $Q, k \geq 2$, given a Qk-uniform hypergraph $H=(V,E)$, it is NP-hard to distinguish the following cases.*

- *Completeness: There is a k-coloring $c\colon V \to [k]$ such that for every hyperedge $e \in E$ and color $i \in [k]$, e has at least $Q-1$ vertices of color i.*
- *Soundness: Every $I \subseteq V$ of measure ϵ induces at least a fraction $\epsilon^{O_{Q,k}(1)}$ of hyperedges. In particular, there is no independent set of measure ϵ, and every $\lfloor \frac{1}{\epsilon} \rfloor$-coloring of H induces a monochromatic hyperedge.*

Fixing $Q=2$ gives a hardness of rainbow coloring with K optimized to be $2k$.

Corollary 1.2. *For all integers $c,k \geq 2$, given a $2k$-uniform hypergraph H, it is NP-hard to distinguish whether H is rainbow k-colorable or is not even c-colorable.*

On the other hand, fixing $k=2$ gives a strong hardness result of discrepancy minimization (with 2 colors). A coloring is said to have discrepancy Δ when in each hyperedge, the difference between the maximum and the minimum number of occurrences of a single color is at most Δ.

Corollary 1.3. *For any $c,Q \geq 2$, given a $2Q$-uniform hypergraph $H=(V,E)$, it is NP-hard to distinguish whether H is 2-colorable with discrepancy 2 or is not even c-colorable.*

The above result strengthens the result of Austrin et al. [2] that shows hardness of 2-coloring in the soundness case. However, their result also holds in $(2Q+1)$-uniform hypergraphs with discrepancy 1, which is not covered by the results in this work.

Our techniques are general – different combinations of test distributions and outer PCPs, plugged into our general *recipe*, yields the following additional results.

Hypergraph vertex cover. Rainbow k-coloring has a tight connection to Hypergraph Vertex Cover, because it partitions the set of vertices into k disjoint vertex covers. In particular, Corollary 1.2 implies that K-Hypergraph Vertex Cover is NP-hard to approximate within a factor of $\left(\frac{K}{2}-\epsilon\right)$, but the better inapproximability factor of $(K-1-\epsilon)$ is already established by the classical result of Dinur et al. [10]. We give the first analytic proof of the same theorem, with two slight improvements: the size of the minimum vertex cover in the completeness case is improved to $\frac{1}{K-1}$ from $\left(\frac{1}{K-1}+\epsilon\right)$, and in the soundness case every set of measure ϵ induces $\epsilon^{O_K(1)}$ fraction of hyperedges.

Theorem 1.4. *For any $\epsilon > 0$ and $K \geq 3$, given a K-uniform hypergraph $H = (V, E)$, it is NP-hard to distinguish the following cases.*

- *Completeness: There is a vertex cover of measure $\frac{1}{K-1}$.*
- *Soundness: Every $I \subseteq V$ of measure ϵ induces at least a fraction $\epsilon^{O_K(1)}$ of hyperedges.*

Bansal and Khot [4] and Sachdeva and Saket [28] focused on *almost* rainbow k-colorable hypergraphs (where one is allowed to remove a small fraction of vertices and all incident hyperedges to ensure rainbow colorability) to show hardness of scheduling problems. This notion allows us to prove the following more structured hardness as well as $(K-1-\epsilon)$-inapproximability for hypergraph vertex cover. It improves [28] in the number of colors used, and almost matches [4] which is based on the Unique Games Conjecture.

Theorem 1.5. *For any $\epsilon > 0$ and $K \geq 3$, given a K-uniform hypergraph $H = (V, E)$, it is NP-hard to distinguish the following cases.*

- *Completeness: There exists $V^* \subseteq V$ of measure ϵ and a coloring $c\colon [V \setminus V^*] \to [K-1]$ such that for every hyperedge of the induced hypergraph on $V \setminus V^*$, $K-2$ colors appear once and the other color twice. Therefore, H has a vertex cover of size at most $\frac{1}{K-1} + \epsilon$.*
- *Soundness: There is no independent set of measure ϵ.*

Q-out-of-$(2Q+1)$-SAT. Q-out-of-$(2Q+1)$-SAT refers to the problem of finding a satisfying assignment in a $(2Q+1)$-CNF formula, given the promise that some assignment makes each clause have at least Q true literals. We give an analytic proof following our recipe of the following result, which was first established based on simpler combinatorial techniques in Austrin et al. [2].

Theorem 1.6. *For $Q \geq 2$, there exists $\epsilon > 0$ depending on Q such that given a $(2Q+1)$-CNF formula, it is NP-hard to distinguish the following cases.*

- *Completeness: There is an assignment such that each clause has at least Q true literals.*
- *Soundness: No assignment can satisfy more than a fraction $(1-\epsilon)$ of clauses.*[1]

[1] An explicit value of ϵ as a function of Q in the soundness is $\exp(-O(Q \log Q))$, which is better than the value $\exp(-O(Q^c))$ for some large absolute constant c implicit in the proof of [2].

1.2. Discussion and open problems: coloring highly structured hypergraphs

The algorithmic and hardness results of highly structured hypergraphs are summarized in Table 1.

Promised Coloring Structure	Algorithm	Hardness
Rainbow K-colorable (K-partite)	2-colorable	Not rainbow K-colorable (Almost, UG) Not weak $O(1)$-colorable [4]
Rainbow $(K-1)$-colorable		**(Almost) Not weak $O(1)$-colorable**
Rainbow $\frac{K}{2}$-colorable with perfect balance	2-colorable	
Rainbow $\frac{K}{2}$-colorable with discrepancy 2		**Not rainbow $\frac{K}{2}$-colorable**
Rainbow $\frac{K}{2}$-colorable with discrepancy $\frac{K}{2}$		**Not weak $O(1)$-colorable**
2-colorable with perfect balance	2-colorable	
2-colorable with discrepancy 1		Not 2-colorable [2]
2-colorable with discrepancy 2		**Not weak $O(1)$-colorable**

Table 1. Summary of algorithmic and hardness results for coloring a highly structured K-uniform hypergraph. Almost means that $\epsilon > 0$ fraction of vertices and incident hyperedges must be deleted to have the structure. UG indicates that the result is based on the Unique Games Conjecture. The results of this work are in boldface.

Fix $K \geq 3$ to be the uniformity of the hypergraph. To the best of our understanding, there is only one general situation under which a K-uniform hypergraph H can be efficiently 2-colored: when $K = Qk$ and H admits a perfectly balanced k-rainbow coloring. By semidefinite programming, we can find a unit vector for each vertex with the guarantee that the K vectors in each hyperedge sum to zero, and the hyperplane rounding will give us a 2-coloring without monochromatic edges (trivially of discrepancy $K-2$). However, the complexity of finding a slightly more structured coloring (e.g., rainbow 3-coloring or 2-coloring with discrepancy less than $K-2$) is wide open. Via a simple reduction from K-colorability on graphs, one can show that finding a rainbow K-coloring (on K-uniform hypergraphs) if one exists is NP-hard. It is, however, consistent with current knowledge (though highly unlikely in our opinion) that a perfectly balanced $\frac{K}{Q}$-coloring ($Q \geq 2$) can be reconstructed in polynomial time.

If we relax the perfect balance promise in the completeness case in certain ways, our results show that the resulting hypergraph becomes hard to even weakly $O(1)$-color. One interesting open question is to show this when there is a 2-coloring of discrepancy 1 (without relying on any unproven conjectures). Another tantalizing challenge is to show hardness of $O(1)$-coloring (or even 2-coloring) when the hypergraph is rainbow $(K-1)$-colorable. We are able to show hardness in the almost rainbow $(K-1)$-colorable case – can we avoid this and achieve perfect completeness?

2. Techniques and related work

We now briefly discuss some closely related works, and then illustrate our main ideas and general recipe in a simple setting.

2.1. Related work

Our work is inspired by recent developments concerning the inapproximability of Hypergraph Vertex Cover and the Constraint Satisfaction Problem (CSP). At a high level, Theorem 1.1 looks similar to the result of Sachdeva and Saket [28] who proved almost the same statement *without perfect completeness* – we need to delete $\epsilon > 0$ fraction of vertices and all incident hyperedges to have a similar guarantee in the completeness case. Achieving perfect completeness is a nontrivial task, as manifested in k-CSP – approximating a $(1-\epsilon)$-satisfiable instance of k-CSP is NP-hard within a factor of $\frac{2k}{2^k}$ [6], while the best inapproximability factor for perfectly satisfiable k-CSP is $\frac{2^{O(k^{1/3})}}{2^k}$ [19].

In CSP, significant research efforts have been made for proving every predicate strictly dominating parity is *approximation resistant* (i.e., no efficient algorithm can beat the ratio achieved by simply picking a random assignment) even on satisfiable instances. O'donnell and Wu [27] proved this assuming the d-to-1 conjecture for $k=3$, and recently this was proven to be true assuming only $\mathbf{P} \neq \mathbf{NP}$ by Håstad ($k=3$, [16]) and Wenner ($k \geq 4$, [31]). Many of these works are based on invariance principle based techniques, and it is natural to ask whether they let us to achieve perfect completeness in Hypergraph Coloring as well. To the best of our knowledge, our work is the first to apply invariance based techniques to prove NP-hardness of Hypergraph Coloring/Vertex Cover problems (Khot and Saket [21] used them to prove hardness of finding an independent set in 2-colorable 3-uniform hypergraphs, assuming the d-to-1 conjecture).

Fourier-analytic proofs of harndess of K-Hypergraph Vertex Cover are known for small K [15,18,20,29]. Even though they cannot be easily generalized to large K, the recent work of Saket [29] for $K=4$ uses general *reverse hypercontractivity* studied by Mossel et al. [23], and we extend his result to present a framework to study general K-uniform hypergraphs. This generalized reverse hypercontractivity might have other applications in hardness of approximation or in other areas of theoretical computer science. In the rest of the section, for simplicity of illustration we fix $Q=k=2$ (so that the test distribution becomes that of [29]) and give a high level glimpse into our proof strategy.

2.2. Techniques

We reduce Label Cover to 4-uniform hypergraph coloring. Given a Label Cover instance based on a bipartite graph $G=(U\cup V,E)$ with projections $\pi_e\colon [R]\to[L]$ (see Section 3 for the formal definition), let U be the *small side* and V be the *big side*. Let $\Omega=\{1,2\}$. Our hypergraph $H=(V',E')$ is defined by $V':=V\times\Omega^R$, and E' is described by the following procedure to sample a hyperedge.

- Sample $u\in U$ and its neighbors $v,w\in V$.
- Sample $x_1,x_2,y_1,y_2\in\Omega^R$ as the following: for $1\le i\le L$,
  - With probability half, $(x_1)_{\pi^{-1}_{(u,v)}(i)},(x_2)_{\pi^{-1}_{(u,v)}(i)},(y_1)_{\pi^{-1}_{(u,w)}(i)}$ are sampled i.i.d., but $(y_2)_j=3-(y_1)_j$ for every $j\in\pi^{-1}_{(u,w)}(i)$.
  - With probability half, $(y_1)_{\pi^{-1}_{(u,w)}(i)},(y_2)_{\pi^{-1}_{(u,w)}(i)},(x_1)_{\pi^{-1}_{(u,v)}(i)}$ are sampled i.i.d., but $(x_2)_j=3-(x_1)_j$ for every $j\in\pi^{-1}_{(u,v)}(i)$.
- Output a hyperedge containing four vertices (v,x_1), (v,x_2), (w,y_1), (w,y_2).

Completeness is obvious from the above distribution. For each *block* that corresponds to $\pi^{-1}_{(u,v)}(i)$ or $\pi^{-1}_{(u,w)}(i)$, one of (x_1,x_2) and (y_1,y_2) is allowed to be sampled independently, but the other pair has to satisfy that two points are different in every coordinate in that block.

For soundness, let I be a large independent set, let $f_v\colon\Omega^R\to\{0,1\}$ be the indicator function of $I\cap(\{v\}\times[k]^R)$. Then I satisfies the following two properties.

$$\mathop{\mathbb{E}}_{v,x_1}[f_v(x_1)] \gg 0, \qquad \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{x_1,x_2,y_1,y_2}[f_v(x_1)f_v(x_2)f_w(y_1)f_w(y_2)] = 0.$$

These two properties seem to be contrary for randomly chosen I, so I with the above two properties should exploit some structure of the reduction. We prove that the existence of such I leads to a good decoding strategy to the Label Cover instance. This implies that there is no large independent set if the Label Cover does not admit a good labeling.

2.2.1. Dealing with noise and influences Before proceeding to the analysis, we discuss two issues that highlight technical difficulties in proving NP-hardness (as opposed to Unique Games-hardness) of coloring with perfect completeness (as opposed to imperfect completeness) in terms of noise.

Strong vs Weak Noise. Given a function $f\colon \Omega^R \to [0,1]$, consider the *noise operator* $T_{1-\gamma}$ defined by $T_{1-\gamma}f(x) = \mathbb{E}_y[f(y)|x]$ where y resamples each coordinate of x with probability γ. It is central to most decoding strategies that we actually analyze noised functions $T_{1-\gamma}f_v$ and $T_{1-\gamma}f_w$ instead of the original functions. We call the step of passing from the original functions to the noised functions *strong noise*. The easiest way to give strong noise is to explicitly include it in the test distribution, independently for all points. However, such explicit and strong noise breaks perfect completeness, since all points might be noised together and we cannot control the behavior.

To deal with this issue, we call *weak noise* to be a property inherent in the test distribution, bounding the correlation between the points we sample. In the test distribution we gave above, it refers to sampling exactly one of (x_1,x_2) or (y_1,y_2) completely independently (for each block). The fact that only one pair is noised is not strong enough to be directly applicable to decoding, but the bounded correlation allows us to apply the result of Mossel [23] to show that the expected value of the product does not change much we replace each f by the noised version only for the sake of analysis. This idea of *smoothing a function in the analysis* allows us maintain perfect completeness.

Block Noise, Block Influence. Consider the projections $\pi_{(u,v)}, \pi_{(u,w)}\colon [R] \to [L]$. Let $d > 1$ be the degree of the projections. d coordinates of x_1, x_2 and d coordinates of y_1, y_2 must be treated in the same *block* which is often regarded as one coordinate.

The aforementioned result of Mossel in fact shows that we can replace f by $\overline{T}_{1-\gamma}f$, where $\overline{T}_{1-\gamma}$ is the *block noise* operator when we view each block as one coordinate. This is not strong enough for our decoding strategy, but the idea of Wenner [31] lets us to replace $\overline{T}_{1-\gamma}f$ by the *individually noised* function $T_{1-\gamma}f$ if f almost depends on only *shattered* parts (roughly, shattered parts of a function under a projection do not distinguish whether the projection is 1-to-1 or not). This shattering behavior can be achieved by Smooth Label Cover defined by Khot [20].

At the end of analysis, our invariance principle will show that $\sum_{1\le i\le L}\overline{\mathsf{Inf}}_i[T_{1-\gamma}f_v]\overline{\mathsf{Inf}}_i[T_{1-\gamma}f_w]$ is large where $\overline{\mathsf{Inf}}$ indicates the influence when we view each block as one coordinate. It turns out to suffice to deal with these *block noises*, since they appear only in the analysis of the decoding; our decoding procedure itself does not depend on the projections, and the goal of the decoding is to have two vertices output the coordinates in the same block. To summarize, we put an effort to pass from block noise to individual noise in the beginning of our analysis, but we keep block influence to the end of analysis where it is naturally integrated with the decoding.

2.2.2. Recipe We briefly discuss the five main steps in the soundness analysis and how they relate to each other. We view distilling and clearly articulating this recipe and highlighting its versatility also as one of the contributions of this work.

1. Fixing a good pair: Given an independent set I of measure ϵ, using smoothness of Label Cover, we show that in the original instance of Label Cover, there is a large fraction $u\in U$ and its neighbors $v,w\in V$ with the following properties. $\mathbb{E}[f_v],\mathbb{E}[f_w]\ge\frac{\epsilon}{2}$, and they almost depend on shattered parts. In the subsequent steps, we fix such u,v,w and analyze the probability that either (u,v) or (u,w) is satisfied by our decoding strategy.

2. Lower bounding in each hypercube: In Theorem 4.8, we show

$$\mathbb{E}[f_v(x_1)f_v(x_2)],\mathbb{E}[f_w(y_1)f_w(y_2)]\ge\zeta(\epsilon)>0.$$

It uses *reverse hypercontractivity* [24,25], which is discussed in Section 4. Roughly, it says the noise operator T_ρ increases q-norm $\|T_\rho f\|_q$ when $q<1$, so that $\|Tf\|_q\ge\|f\|_p$ for some $q<p<1$ depending on ρ (note that $\|f\|_q\le\|f\|_p$). The case $k=2$ follows directly from the previous result, but for larger k we generalize the reverse hypercontractivity to more general operators, even between different spaces. This step does not depend on noise or the degree of projections (e.g., the same ζ works for $T_{1-\gamma}f$ and $\overline{T}_{1-\gamma}f$).

3. Smoothing functions (based on 1.): Based on the bounded correlation of the test distribution, we use the result of Mossel [23] to pass from f to $\overline{T}_{1-\gamma}f$. The fact that f_v,f_w almost depend on shattered parts allows us to use Theorem 5.5 to pass from $\overline{T}_{1-\gamma} f$ to $T_{1-\gamma} f$. Therefore we have

$$\begin{aligned} &\mathop{\mathbb{E}}_{x_1,x_2,y_1,y_2}[f_v(x_1)f_v(x_2)f_w(y_1)f_w(y_2)] \\ \approx &\mathop{\mathbb{E}}_{x_1,x_2,y_1,y_2}[T_{1-\gamma}f_v(x_1)T_{1-\gamma}f_v(x_2)T_{1-\gamma}f_w(y_1)T_{1-\gamma}f_w(y_2)]. \end{aligned}$$

For simplicity, let $f' = T_{1-\gamma} f$.

4. Invariance (based on 2. and 3.): Since I is independent, the above results imply

$$\begin{aligned} 0 \approx &\mathop{\mathbb{E}}_{x_1,x_2,y_1,y_2}[f'_v(x_1)f'_v(x_2)f'_w(y_1)f'_w(y_2)] \ll \zeta^2 \\ \leq &\mathop{\mathbb{E}}_{x_1,x_2}[f'_v(x_1)f'_v(x_2)] \mathop{\mathbb{E}}_{y_1,y_2}[f'_w(y_1)f'_w(y_2)]. \end{aligned}$$

In Theorem 5.6, we use an invariance principle inspired by that of Wenner [31] and Chan [6] to conclude that $\sum_{1\leq i\leq L} \overline{\mathsf{Inf}}_i[f'_v]\,\overline{\mathsf{Inf}}_i[f'_w] \geq \tau$, which implies that there are matching (blocks of) influential coordinates. The crucial property we used is that x_i is independent of (y_1, y_2) – one point is independent of the joint distribution of the points not in the same hypercube.

5. Decoding Strategy (based on 3. and 4.): The standard decoding strategy based on Fourier coefficients of f shows that either (u, v) or (u, w) will be satisfied with good probability. As previously discussed, $\sum_{1\leq i\leq L} \overline{\mathsf{Inf}}_i[f'_v]\,\overline{\mathsf{Inf}}_i[f'_w] \geq \tau$ gives large common block influences of individually noised functions, and they are sufficient for the decoding.

2.2.3. Organization Section 3 introduces basic definitions and their properties used in the paper. Our main technical tool, generalized reverse hypercontractivity, is introduced in Section 4. Section 5 proves the main Theorem 1.1, deferring the technical proofs about Label Cover, invariance/noise to A, B respectively. In Appendix C, and D, we show the versatility of our approach by proving Theorem 1.4, 1.5, and 1.6, using the same procedure.

3. Preliminaries

For a positive integer k, let $[k] := \{1, 2, \ldots, k\}$. Let $\mathbb{S}_k$ be the set of k-permutations – $(x_1, \ldots, x_k) \in [k]^k$ such that $x_i \neq x_j$ for all $i \neq j$. For a vector $x \in \mathbb{R}^m$ and $S \subseteq [m]$, x_S denotes the projection of x onto the coordinates in S. Definitions and simple properties introduced from Section 3.1 to Section 3.4 are from Mossel [23].

3.1. Correlated spaces

Given a probability space (Ω, μ) (we always consider finite probability spaces), let $\mathcal{L}(\Omega)$ be the set of functions $\{f\colon \Omega \to \mathbb{R}\}$ and for an interval $I \subseteq \mathbb{R}$, $\mathcal{L}_I(\Omega)$ be the set of functions $\{f\colon \Omega \to I\}$. A collection of probability spaces are said to be correlated if there is a joint probability distribution on them. We will denote k correlated spaces $\Omega_1, \dots, \Omega_k$ with a joint distribution μ as $(\Omega_1 \times \dots \times \Omega_k; \mu)$. Note that the definition of correlated spaces includes the joint distribution. Two instantiations of correlated spaces, even though they are defined on the same underlying sets, are considered different when their distributions are not the same.

Given two correlated spaces $(\Omega_1 \times \Omega_2, \mu)$, we define the correlation between Ω_1 and Ω_2 by

$$\rho(\Omega_1, \Omega_2; \mu) := \sup\left\{\mathsf{Cov}[f, g]\colon f \in \mathcal{L}(\Omega_1), g \in \mathcal{L}(\Omega_2), \mathsf{Var}[f] = \mathsf{Var}[g] = 1\right\}.$$

The following lemma of Wenner [31] gives a convenient way to bound the correlation.

Lemma 3.1 (Corollary 2.18 of [31]). *Let $(\Omega_1 \times \Omega_2, \mu)$ and $(\Omega_1 \times \Omega_2, \mu')$ be two distinct instantiations of correlated spaces such that the marginal distribution of at least one of Ω_1 and Ω_2 is identical on μ and μ'. For any $0 \le \delta \le 1$, consider another correlated spaces $(\Omega_1 \times \Omega_2, \delta\mu + (1-\delta)\mu')$. Then,*

$$\rho(\Omega_1, \Omega_2; \delta\mu + (1-\delta)\mu') \le \sqrt{\delta \cdot \rho(\Omega_1, \Omega_2; \mu)^2 + (1-\delta) \cdot \rho(\Omega_1, \Omega_2; \mu')^2}.$$

Given k correlated spaces $(\Omega_1 \times \dots \times \Omega_k, \mu)$, we define the correlation of these spaces by

$$\rho(\Omega_1, \dots, \Omega_k; \mu) := \max_{1 \le i \le k} \rho\left(\prod_{1 \le j \le i-1} \Omega_j \times \prod_{i+1 \le j \le k} \Omega_j, \Omega_i; \mu\right).$$

3.2. Operators

Let $(\Omega_1 \times \Omega_2, \mu)$ be two correlated spaces. The *Markov operator* associated with them is the operator mapping $f \in \mathcal{L}(\Omega_1)$ to $Tf \in \mathcal{L}(\Omega_2)$ by

$$(Tf)(y') = \mathop{\mathbb{E}}_{(x,y)\sim\mu}[f(x) | y = y'].$$

The *noise operator* or *Bonami-Beckner operator* T_ρ $(0 \le \rho \le 1)$ associated with a single probability space (Ω, μ) is the Markov operator associated with $(\Omega \times \Omega, \nu)$, where $\nu(x,y) = (1-\rho)\mu(x)\mu(y) + \rho\mathbb{I}[x=y]\mu(x)$ and $\mathbb{I}[\cdot]$ is the indicator function – ν samples (x,y) independently with probability $1-\rho$, and samples $x = y$ with probability ρ. Note that $T_\rho f(y) = \rho f(y) + (1-\rho)\mathbb{E}_\mu[f(x)]$.

3.3. Functions and influences

Let (Ω,μ) be a probability space. Given a function $f \in \mathcal{L}(\Omega)$ and $p \in \mathbb{R}$, let $\|f\|_p := \mathbb{E}_{x\sim\mu}[|f(x)|^p]^{1/p}$. We also use $\|f\|_{p,\mu}$ for the same quantity if it is instructive to emphasize μ. We note that $\|f\|_p$ for $p<0$ is also used throughout the paper, but in this case we ensure that $f>0$. For $f,g\in\mathcal{L}(\Omega)$, $\langle f,g\rangle := \mathbb{E}_{x\sim\mu}[f(x)g(x)]$.

Consider a product space $(\Omega^R,\mu^{\otimes R})$ and $f \in \mathcal{L}(\Omega^R)$. The *Efron-Stein decomposition* of f is given by

$$f(x_1,\dots,x_R) = \sum_{S\subseteq[R]} f_S(x_S),$$

where (1) f_S depends only on x_S and (2) for all $S \not\subseteq S'$ and all $x_{S'}$, $\mathbb{E}_{x'\sim\mu^{\otimes R}}[f_S(x')|x'_{S'}=x_{S'}]=0$.

The *influence* of the jth coordinate on f is defined by

$$\mathsf{Inf}_j[f] := \mathop{\mathbb{E}}_{x_1,\dots,x_{j-1},x_{j+1},\dots,x_R}\Big[\mathop{\mathsf{Var}}_{x_j}[f(x_1,\dots,x_R)]\Big].$$

Given the noise operator T_ρ for (Ω,μ), we let $T_\rho^{\otimes R}$ be the noise operator for $(\Omega^R,\mu^{\otimes R})$ (i.e., noising each coordinate independently) and call it T_ρ. The noise operator and the influence has a convenient expression in terms of the Efron-Stein decomposition:

$$T_\rho[f] = \sum_S \rho^{|S|} f_S; \qquad \mathsf{Inf}_j[f] = \Bigg\|\sum_{S\colon j\in S} f_S\Bigg\|_2^2 = \sum_{S\colon j\in S}\|f_S\|_2^2.$$

The following lemma lets us to reason about the influences of the product of functions. The proof is in Section B.1.

Lemma 3.2 ([30]). *Let $(\Omega_1\times\cdots\times\Omega_k,\mu)$ be k probability spaces and $(\Omega_1^L\times\cdots\times\Omega_k^L,\mu^{\otimes L})$ be the corresponding product spaces. Let $f_i\in\mathcal{L}_{[-1,1]}(\Omega_i^L)$, and $F\in\mathcal{L}_{[-1,1]}(\Omega_1^L\times\cdots\times\Omega_k^L)$ such that $F(x_1,\dots,x_k)=\prod_{1\le i\le k} f_i(x_i)$. Then for $1\le j\le L$, $\mathsf{Inf}_j(F)\le k\sum_{i=1}^k \mathsf{Inf}_j(f_i)$.*

3.4. Blocks

Let R,L,d be positive integers satisfying $R=dL$. Let $(\Omega^R,\mu^{\otimes R})$ be a product space and $\pi\colon[R]\to[L]$ be a projection such that $|\pi^{-1}(j)|=d$ for $1\le j\le L$. Define $\overline{\Omega}:=\Omega^d$. Given $x\in\Omega^R$, we *block* x to have $\overline{x}\in(\overline{\Omega})^L$ defined by

$$\overline{x}_j := (x_{j'})_{\pi(j')=j}.$$

Given $f \in \mathcal{L}(\Omega^R)$, its *blocked version* $\overline{f} \in \mathcal{L}(\overline{\Omega}^L)$ is defined by $\overline{f}(\overline{x}) := f(x)$ for any $x \in \Omega^R$. These blocked versions of functions and arguments depend on the projection π. For each function f, the associated projection will be clear from the context, and the same projection is used to block its argument x. The influence $\mathsf{Inf}_j[\overline{f}]$ and the noise operator $T_\rho \overline{f}$ are naturally defined. Define

$$\overline{\mathsf{Inf}}_j[f] := \mathsf{Inf}_j[\overline{f}], \quad \forall j \in [L]; \qquad (\overline{T}_\rho f)(x) := (T_\rho \overline{f})(\overline{x}), \quad \forall x \in \Omega^R,$$

and call them *block influence* and *block noise operator* respectively. They also have the following nice expressions in terms of f's Efron-Stein decomposition.

$$\overline{T}_\rho f = \sum_S \rho^{|\pi(S)|} f_S; \qquad \overline{\mathsf{Inf}}_j[f] = \sum_{S\colon S \cap \pi^{-1}(j) \neq \emptyset} \|f_S\|_2^2.$$

A subset $S \subseteq [R]$ is said to be *shattered* by π if $|S| = |\pi(S)|$. For a positive integer J, define the *bad* part of f_v under π and J as

$$f^{\mathsf{bad}} = \sum_{S\colon \text{not shattered and } |\pi(S)| < J} f_S.$$

3.5. Q-Hypergraph label cover

An instance of Q-Hypergraph Label Cover is based on a Q-uniform hypergraph $H = (V, E)$. Each hyperedge-vertex pair (e, v) such that $v \in e$ is associated with a projection $\pi_{e,v}\colon [R] \to [L]$ for some positive integers R and L. A labeling $l\colon V \to [R]$ *strongly satisfies* $e = \{v_1, \ldots, v_Q\}$ when $\pi_{e,v_1}\big(l(v_1)\big) = \cdots = \pi_{e,v_Q}(l(v_Q))$. It *weakly satisfies* e when $\pi_{e,v_i}(l(v_i)) = \pi_{e,v_j}(l(v_j))$ for some $i \neq j$. The following are two desired properties of instances of Q-Hypergraph Label Cover.

- ϵ-weakly dense: any subset of V of measure at least $\epsilon' \geq \epsilon$ induces at least $\frac{(\epsilon')^Q}{2^{Q+1}}$ fraction of hyperedges.
- T-smooth: for all $v \in V$ and $i \neq j \in [R]$, $\Pr_{e \in E\colon e \ni v}[\pi_{e,v}(i) = \pi_{e,v}(j)] \leq \frac{1}{T}$.

The following theorem asserts that it is NP-hard to find a good labeling in such instances. The proof is in Appendix A.1, and closely follows the work of Gopalan et al. [13] that proves the hardness of the same problem without T-smoothness.

Theorem 3.3. *For any $Q \geq 2, T \geq 1$ and $\eta, \epsilon > 0$, given an instance of Q-Hypergraph Label Cover that is ϵ-weakly-dense and T-smooth, it is NP-hard to distinguish*

- *Completeness: There exists a labeling l that strongly satisfies every hyperedge.*
- *Soundness: No labeling l can weakly satisfy a fraction η of hyperedges.*

4. Reverse hypercontractivity

The version of reverse hypercontractivity we use is stated below.

Theorem 4.1 ([25]). *Let (Ω,μ) be a probability space. Fix $0\le\rho<1$. There exist $q<0<p<1$ such that for any $f\in\mathcal{L}_{[0,\infty)}(\Omega)$,*

$$\|T_\rho f\|_q \ge \|f\|_p.$$

We now generalize the above reverse hypercontractivity result to more general operators, extending the noise operator T_ρ in two ways.

- Between two difference spaces: while T_ρ is the Markov operator associated with two correlated copies of the same probability space $(\Omega_1\times\Omega_1,\nu)$, we are interested in the Markov operator T associated with two correlated spaces $(\Omega_1\times\Omega_2,\nu')$, possibly $\Omega_1\neq\Omega_2$.
- Arbitrary distribution instead of diagonal distribution: ν samples x,y independently according to the marginal and output (x,x) with probability ρ and (x,y) with probability $1-\rho$. Since $\Omega_1\neq\Omega_2$, the former does not make sense. Instead, with probability ρ, ν' samples (x,y) according to another arbitrary distribution ν'', as long as the marginals of x and y are preserved.

This extension is based on simple observation that such an operator T can be expressed as $T=PT_\rho$ for some Markov operator $P\colon\mathcal{L}(\Omega_1)\to\mathcal{L}(\Omega_2)$ which shares the marginals with T. The idea of decomposition in terms of T_ρ was also used in [25] when analyzing general operators on the same space. The following lemma shows that any Markov operator does not decrease q-norm when $q\le 1$.

Lemma 4.2. *Let $(\Omega_1\times\Omega_2,\mu)$ be two correlated spaces, with the marginal distribution μ_i of Ω_i. Let P be the Markov operator associated with it. For any $q\le 1$ and $f\in\mathcal{L}_{(0,\infty)}(\Omega_1)$,*

$$\|Pf\|_q \ge \|f\|_q.$$

Proof. Since $x\mapsto x^q$ is concave,

$$\|Pf\|_q^q = \mathop{\mathbb{E}}_{y\sim\mu_2}[(Tf(y))^q] = \mathop{\mathbb{E}}_{y\sim\mu_2}\left[\left(\mathop{\mathbb{E}}_{x\sim\mu_1}[f(x)|y]\right)^q\right]$$

$$\geq \mathop{\mathbb{E}}_{y\sim\mu_2}\left[\mathop{\mathbb{E}}_{x\sim\mu_1}[f(x)^q|y]\right] = \mathop{\mathbb{E}}_{x\sim\mu_1}[f(x)^q] = \|f\|_q^q. \qquad \blacksquare$$

The following main lemma says that whenever T_ρ exhibits the reverse hypercontractive behavior for some p, q, the same conclusion holds for Markov operators with the same parameters.

Lemma 4.3 (Reverse hypercontractivity of two correlated spaces). *Let $(\Omega_1\times\Omega_2, \mu)$ be two correlated spaces, and with the marginal distribution μ_i of Ω_i. Let T be the Markov operator associated with it. Suppose that $T = \rho P + (1-\rho)J_{1,2}$ for $0 \leq \rho < 1$, where $J_{1,2}$ is the Markov operator associated with $(\Omega_1\times\Omega_2, \mu_1\otimes\mu_2)$ and P is the Markov operator associated with $(\Omega_1\times\Omega_2, \nu)$ for some ν with the same marginals as μ. Let $q < p < 1$ be such that $\|T_\rho f\|_q \geq \|f\|_p$ for any $f \in \mathcal{L}_{[0,\infty)}(\Omega_1)$. Then,*

$$\|Tf\|_q \geq \|f\|_p.$$

Proof. Note that $T_\rho = \rho I_1 + (1-\rho)J_1$, where I_1 is the identity operator, and J_1 is the Markov operator associated with $(\Omega_1^2, \mu_1^{\otimes 2})$. The following simple relationship holds between T and T_ρ,

$$PT_\rho = \rho P I_1 + (1-\rho)PJ_1 = \rho P + (1-\rho)J_{1,2} = T.$$

The fact that $T = PT_\rho$ implies

$$\|Tf\|_q = \|PT_\rho f\|_q \geq \|T_\rho f\|_q \geq \|f\|_p,$$

where the first inequality follows from Lemma 4.2. $\blacksquare$

Along the way to apply the above result to our setting, we introduce a basic intermediate problem which may be of independent interest.

Question 4.4. Let $(\Omega_1\times\Omega_2, \mu)$ be two correlated spaces. Given two (biased, not necessarily Boolean) hypercubes Ω_1^L and Ω_2^L, their subsets $S \subseteq \Omega_1^L, S' \subseteq \Omega_2^L$, and two random points $x \in \Omega_1^L, y \in \Omega_2^L$ such that each (x_i, y_i) is sampled from μ independently, what is the probability that $x \in S$ and $y \in S'$?

To answer this question, we use the following reverse Hölder inequality in a similar way to [24].

Theorem 4.5 ([17]). *Let f and g be nonnegative functions and supplse $\frac{1}{p} + \frac{1}{p'} = 1$, where $p < 1$. Then*

$$\mathbb{E}[fg] = ||fg||_1 \geq ||f||_p||g||_{p'}.$$

Using the above inequality and the standard two-function hypercontractivity induction [26], the following lemma shows that as long as μ contains nonzero copy of product distributions (equivalent to $T=\rho P+(1-\rho)J_{1,2}$ for $\rho<1$), the above probability is at least a positive number depending only on the measure of S and S', and ρ (but crucially it does not depend on L). Note that when f is an indicator function whose value is either 0 or 1, for any $p>0$, $||f||_p=(\mathbb{E}_x[f(x)^p])^{1/p}=(\mathbb{E}[f])^{1/p}$.

Lemma 4.6. *Let $(\Omega_1,\Omega_2,\mu),\rho,T,P$ be defined as Lemma 4.3. There exist $0<p,q<1$ such that for any $f\in\mathcal{L}_{[0,\infty)}(\Omega_1^L)$ and $g\in\mathcal{L}_{[0,\infty)}(\Omega_2^L)$,*

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu^{\otimes L}}[f(x)g(y)] = \mathop{\mathbb{E}}_{y\sim\mu_2^{\otimes L}}[g(y)T^{\otimes L}f(y)] \geq \|f\|_p\|g\|_q.$$

Proof. The equality holds by definition, so it only remains to prove the inequality. We first prove it for $L=1$, and do the induction on L. Invoke Theorem 4.1 to get $q'<0<p<1$ such that $\|T_\rho f\|_{q'}\geq\|f\|_p$. Let $0<q<1$ be such that $\frac{1}{q}+\frac{1}{q'}=1$. By the reverse Hölder inequality and Lemma 4.3,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}[f(x)g(y)] = \mathop{\mathbb{E}}_{y\sim\mu_2}[g(y)Tf(y)] \geq \|Tf\|_{q'}\|g\|_q \geq \|f\|_p\|g\|_q$$

as desired.

For $L>1$, we use the notation $x=(x',x_L)$ where $x'=(x_1,\ldots,x_{L-1})$, and similar notation for y. Note that $(x',y')\sim\mu^{\otimes L-1}$ and $(x_L,y_L)\sim\mu$. We also write f_{x_L} for the restriction of f in which the last coordinate is fixed to value x_L, and similarly for g.

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mu^{\otimes L}}[f(x)g(y)] &= \mathop{\mathbb{E}}_{(x_L,y_L)\sim\mu}\ \mathop{\mathbb{E}}_{(x',y')\sim\mu^{\otimes L-1}}[f_{x_L}(x')g_{y_L}(y')]\\ &\geq \mathop{\mathbb{E}}_{(x_L,y_L)\sim\mu}[\|f_{x_L}\|_{p,\mu_1^{\otimes L-1}}\|g_{y_L}\|_{q,\mu_2^{\otimes L-1}}]\end{aligned}$$

by induction. Let F,G be the function defined by $F(x_L)=\|f_{x_L}\|_p$, $G(y_L)=\|g_{y_L}\|_q$. Then

$$\mathop{\mathbb{E}}_{(x_L,y_L)\sim\mu}[F(x_L)G(y_L)] \geq \|F\|_{p,\mu_1}\|G\|_{q,\mu_2}$$

by the base case. Finally,

$$\|F\|_{p,\mu_1} = \mathop{\mathbb{E}}_{x_L\sim\mu_1}[|F(x_L)|^p]^{1/p} = \left(\mathop{\mathbb{E}}_{x_L\sim\mu_1}\ \mathop{\mathbb{E}}_{x'\sim\mu_1^{\otimes L-1}}[|f_{x_L}|^p]\right)^{1/p} = \|f\|_{p,\mu_1^{\otimes L}}$$

and similarly $\|G\|_{q,\mu_2}=\|g\|_{q,\mu_2^{\otimes L}}$. The induction is complete. ∎

By another induction on the number of functions, we can extend the answer to the previous question to $k>2$.

Question 4.7. Let (Ω^k, μ) be k correlated copies of the same space. Given a hypercube Ω^L, its subsets $S \subseteq \Omega^L$, and k random points $x_1, \dots, x_k \in \Omega^L$ such that each $((x_1)_i, \dots, (x_k)_i)$ is sampled from μ independently for $i \in [L]$, what is the probability that $x_j \in S$ for all $j \in [k]$?

Theorem 4.8. *Let (Ω^k, ν) be k correlated spaces with the same marginal σ for each copy of Ω. Suppose that ν is described by the following procedure to sample from Ω^k.*

- *With probability ρ $(0 \le \rho < 1)$, it samples from an arbitrary distribution on Ω^k, which has the marginal σ for each copy of Ω.*
- *With probability $1-\rho$, it samples from $\sigma^{\otimes k}$.*

Let $F_1, \dots, F_k \in \mathcal{L}_{[0,1]}(\Omega^L)$ such that $\mathbb{E}[F_i] \ge \epsilon > 0$ for all i. Then there exists $\zeta := \zeta(\rho, \epsilon, k) = \epsilon^{O_{\rho,k}(1)} > 0$ (independent of L) such that

$$\mathop{\mathbb{E}}_{x_1,\dots,x_k}\left[\prod_{1\le i\le k} F_i(x_i)\right] \ge \zeta,$$

where for each $1 \le j \le L$, $((x_1)_j, \dots, (x_k)_j)$ is sampled according to ν.

Proof. We proceed by the induction on k. For $k=1$, $\zeta = \epsilon$ works.

For $k > 1$, consider two correlated spaces $(\Omega \times \Omega^{k-1}, \nu)$ where the marginal of Ω is σ and the marginal of Ω^{k-1} is ν'. Note that the marginal of ν' on each copy of Ω is still σ. Invoke Lemma 4.6 to obtain $0 < p, q < 1$ be such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\nu^{\otimes L}}[F(x)G(y)] \ge \|F\|_{p,\sigma^{\otimes L}} \|G\|_{q,\nu'^{\otimes L}}$$

for any $F \in \mathcal{L}_{[0,\infty)}(\Omega^L)$ and $G \in \mathcal{L}_{[0,\infty)}(\Omega^{k-1})^L$.

$$\mathop{\mathbb{E}}_{x_1,\dots,x_k}\left[\prod_{1\le i\le k} F_i(x_i)\right] \ge \|F_1\|_{p,\sigma^{\otimes L}} \left\|\prod_{i=2}^{k} F_i(x_i)\right\|_{q,\nu'^{\otimes L}}.$$

Since $F_i \in \mathcal{L}_{[0,1]}(\Omega^L)$, $\|F_i\|_p \ge \epsilon^{1/p}$. Since ν' can be also described by the procedure in the statement of the theorem (except that it is on Ω^{k-1}), we obtain $\zeta(\rho, \epsilon, k-1)$ such that

$$\left\|\prod_{i=2}^{k} F_i(x_i)\right\|_{q,\nu'^{\otimes L}} \ge \left(\mathop{\mathbb{E}}_{x_2,\dots,x_k}\left[\prod_{i=2}^{k} F_i(x_i)\right]\right)^{1/q} \ge \zeta(\rho, \epsilon, k-1)^{1/q}.$$

Therefore, $\zeta(\rho, \epsilon, k) = \zeta(\rho, \epsilon, k-1)^{1/q} \epsilon^{1/p}$ completes the induction. Since p, q depend only on ρ, $\zeta(\rho, \epsilon, k) = \epsilon^{O_{\rho,k}(1)}$ in every step of induction. ∎

Remark 4.9. The same statement holds even when we replace Ω^k by the product of k different spaces $\Omega_1 \times \cdots \times \Omega_k$.

5. Hardness of rainbow coloring

Fix $Q, k \geq 2$. In this section, we show a reduction from Q-Hypergraph Label Cover to Qk-Hypergraph Coloring, proving Theorem 1.1.

5.1. Distributions

We first define the distribution for each block. Qk points $x_{q,i} \in [k]^d$ for $1 \leq q \leq Q$ and $1 \leq i \leq k$ are sampled by the following procedure.

- Sample $q' \in [Q]$ uniformly at random.
- Sample $x_{q',1}, \ldots, x_{q',k} \in [k]^d$ i.i.d.
- For $q \neq q'$, $1 \leq j \leq d$, sample a permutation $((x_{q,1})_j, \ldots, (x_{q,k})_j) \in \mathbb{S}_k$ uniformly at random.

There are several distributions involved.

Let $\Omega := [k]$ and ω be the uniform distribution on Ω. For any $1 \leq q \leq Q$, $1 \leq i \leq k$ and $1 \leq j \leq d$, the marginal of $(x_{q,i})_j$ follows (Ω, ω).

For any $1 \leq q \leq Q$ and $1 \leq i \leq k$, the marginal of $(x_{q,i})$ follows $(\Omega^d, \omega^{\otimes d})$. Let $\overline{\Omega} := \Omega^d$.

Let (Ω^k, μ) be the marginal distribution of $\big((x_{q,1})_j, \ldots, (x_{q,k})_j\big)$, which is the same for all q and i. Note that μ is not uniform – with probability $\frac{1}{Q}$ it is uniform on $[k]^k$, but with probability $\frac{Q-1}{Q}$ it samples from $k!$ permutations.

Let $(\Omega^{dk}, \overline{\mu})$ be the marginal distribution of $(x_{q,1}, \ldots, x_{q,k})$, which is the same for all q.

Finally, let $(\Omega^{Qkd}, \overline{\mu}')$ be the entire distribution of $(x_{q,i})_{q \in [Q], i \in [k]}$.

We first consider $(\Omega^{Qkd}, \overline{\mu}')$ as Qk correlated spaces $(\overline{\Omega}^{Qk}, \overline{\mu}')$, and bound $\rho(\overline{\Omega}^{Qk}; \overline{\mu}')$. Let $\overline{\Omega}_{q,i}$ denote the copy of $\overline{\Omega}$ associated with $x_{q,i}$, and $\overline{\Omega}'_{q,i}$ be the product of the other $Qk-1$ copies.

Fix some q and i. Note that $\overline{\mu}' = \frac{1}{Q}\alpha_q + \frac{Q-1}{Q}\beta_q$ where α_q denotes the distribution given $q' = q$ (so that each entry of $x_{q,1}, \ldots, x_{q,k}$ is sampled i.i.d.), and β_q denotes the distribution given $q' \neq q$. Since each entry of $x_{q,i}$ is sampled i.i.d. in α_q, $\rho(\overline{\Omega}_{q,i}, \overline{\Omega}'_{q,i}; \alpha_q) = 0$. Observed that, in both α_q and β_q, the marginal of $x_{q,i}$ is $\omega^{\otimes d}$. By Lemma 3.1, we conclude that $\rho(\overline{\Omega}_{q,i}, \overline{\Omega}'_{q,i}; \overline{\mu}') \leq$

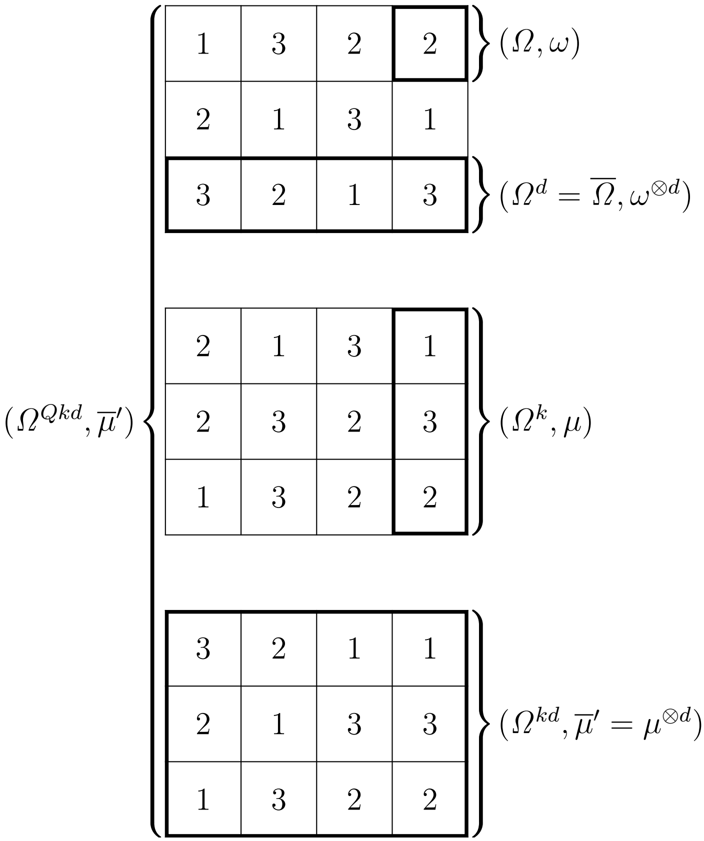

Figure 1. An example for $Q = k = 3$, $d = 4$. $q' = 2$ so that all columns of the first and third block are permutations

$\sqrt{\frac{Q-1}{Q}}$. Therefore we have

$$\rho((\overline{\Omega}_{q,i})_{q,i}; \overline{\mu}') = \max_{q,i} \rho(\overline{\Omega}_{q,i}, \overline{\Omega}'_{q,i}; \overline{\mu}') \le \sqrt{\frac{Q-1}{Q}}.$$

5.2. Reduction and completeness

We now describe the reduction from Q-Hypergraph Label Cover. Given a Q-uniform hypergraph $H = (V, E)$ with Q projections from $[R]$ to $[L]$ for each hyperedge (without loss of generality[2], we assume each projection is d-to-1 where $d=R/L$), the resulting instance of Qk-Hypergraph Coloring is $H'=(V',E')$ where $V'=V\times[k]^R$. Let $\mathsf{Cloud}(v):=\{v\}\times[k]^R$. The set E' consists of hyperedges generated by the following procedure.

- Sample a random hyperedge $e=(v_1,\dots,v_Q)\in E$ with associated projections $\pi_{e,v_1},\dots,\pi_{e,v_Q}$ from E.
- Sample $(x_{q,i})_{1\le q\le Q,1\le i\le k}\in\Omega^R$ in the following way. For each $1\le j\le L$, independently sample $((x_{q,i})_{\pi^{-1}_{e,v_q}(j)})_{q,i}$ from $(\Omega^{Qkd},\overline{\mu}')$.
- Add a hyperedge between Qk vertices $\{(v_q,x_{q,i})\}_{q,i}$ to E'. We say this hyperedge is *formed from* $e\in E$.

Given the reduction, completeness is easy to show.

Lemma 5.1. *If an instance of Q-Hypergraph Label Cover admits a labeling that strongly satisfies every hyperedge $e\in E$, there is a coloring $c\colon V'\to[k]$ such that every hyperedge $e'\in E'$ has at least $(Q-1)$ vertices of each color.*

Proof. Let $l\colon V\to[R]$ be a labeling that strongly satisfies every hyperedge $e\in E$. For any $v\in V, x\in[k]^R$, let $c(v,x)=x_{l(v)}$. For any hyperedge $e'=\{(v_q,x_{q,i})\}_{q,i}\in E'$, $c(v_q,x_{q,i})=(x_{q,i})_{l(v_q)}$, and all but one q satisfies $\{(x_{q,1})_{l(v_q)},\dots,(x_{q,k})_{l(v_q)}\}=[k]$. Therefore, the above strategy ensures that every hyperedge of E' contains at least $(Q-1)$ vertices of each color. ∎

5.3. Soundness

Lemma 5.2. *For any $\epsilon>0$, there exists $\eta:=\eta(\epsilon,Q,k)$ such that if $I\subseteq V'$ of measure ϵ induces less than $\epsilon^{O_{Q,k}(1)}$ fraction of hyperedges, the corresponding instance of Q-Hypergraph Label Cover admits a labeling that weakly satisfies a fraction η of hyperedges.*

As introduced in Section 2, the proof of soundness consists of the following five steps.

STEP 1. **Fixing a Good Hyperedge.** Let $I\subseteq V'$ be of measure ϵ. For each vertex $v\in V$, let $f_v\colon[k]^R\to\{0,1\}$ be the indicator function of $I\cap\mathsf{Cloud}(v)$. Call a vertex v *heavy* when $\mathbb{E}[f_v]\ge\frac{\epsilon}{2}$. By averaging, at least $\frac{\epsilon}{2}$ fraction of vertices are heavy. By Theorem 3.3, we can assume that the original

[2] We can assume that the number of labels from $[R]$ that project to a fixed label in $[L]$ is the same for all projections, since original Label Cover is also hard to approximate with this condition as shown in Theorem 1.17 of [31].

Q-Hypergraph Label Cover instance is $\frac{\epsilon}{2}$-weakly-dense. At least $\delta := \frac{(\frac{\epsilon}{2})^Q}{2^{Q+1}}$ fraction of hyperedges are induced by the heavy vertices.

Recall that we can require the original Q-Hypergraph Label Cover instance to be T-smooth for T that can be chosen arbitrarily large. Let J be a positive integer. The parameters J and T will be determined later as large constants depending on Q, k, and ϵ.

Fix f_v and $S \subseteq [R]$. Over a random hyperedge e containing v and the associated projection $\pi_{e,v}$, we bound the probability that $|S|$ is not shattered and $|\pi_{e,v}(S)| < J$. If $|S| \leq J$, by union bound over all pairs $i \neq j$, the probability that S is not shattered is at most $\frac{J^2}{T}$. If $|S| > J$, the probability that $|\pi_{e,v}(S)| < J$ is at most the probabilty that a fixed J-subset of S is not shattered, which is at most $\frac{J^2}{T}$. Since $\sum_S \|(f_v)_S\|_2^2 = \|f_v\|_2^2 \leq 1$, we have

$$\mathop{\mathbb{E}}_{e}[\|f_v^{\mathsf{bad}}\|_2^2] \leq \frac{J^2}{T},$$

where f_v^{bad} denotes the bad part of f_v under $\pi_{e,v}$ and J (we suppress the dependence on the projection $\pi_{e,v}$ and J for notational convenience). Therefore, $\mathbb{E}_e[\|f_v^{\mathsf{bad}}\|_2] \leq \left(\frac{J^2}{T}\right)^{1/2}$ and at least $1 - \left(\frac{J^2}{T}\right)^{1/4}$ fraction of hyperedges containing v satisfy $\|f_v^{\mathsf{bad}}\|_2 \leq \left(\frac{J^2}{T}\right)^{1/4}$. Call such hyperedges *good* for v.

By union bound, at least $1 - Q\left(\frac{J^2}{T}\right)^{1/4}$ fraction of hyperedges are good for every vertex they contain. By setting $Q\left(\frac{J^2}{T}\right)^{1/4} \leq \frac{\delta}{2}$, we can conclude that at least a fraction $\frac{\delta}{2}$ of hyperedges are induced by the heavy vertices and good for every vertex they contain.

Throughout the rest of the section, fix such a hyperedge $e = (v_1, \ldots, v_Q)$ and the associated projections $\pi_{e,v_1}, \ldots, \pi_{e,v_Q}$. For simplicity, let $f_q := f_{v_q}$ and $\pi_q := \pi_{e,v_q}$ for $q \in [Q]$. We now measure the fraction of hyperedges formed from e that are wholly contained within I. The fraction such hyperedges is

$$\mathop{\mathbb{E}}_{x_{q,i}}\left[\prod_{1 \leq q \leq Q, 1 \leq i \leq k} f_q(x_{q,i})\right]. \tag{1}$$

STEP 2. **Lower Bounding in Each Hypercube.** Fix any $q \in [Q]$. We prove that $\mathbb{E}\left[\prod_{1 \leq i \leq k} T_{1-\gamma} f_q(x_{q,i})\right] \geq \zeta$ for some $\zeta > 0$ and every $\gamma \in [0,1]$. The main tool in this part is a generalization of reverse hypercontractivity, which is discussed in Section 4. The final result is the following.

Theorem 5.3 (Restatement of Theorem 4.8). *Let (Ω^k, ν) be k correlated spaces with the same marginal σ for each copy of Ω. Suppose that ν is described by the following procedure to sample from Ω^k.*

- *With probability ρ $(0 \le \rho < 1)$, it samples from an arbitrary distribution on Ω^k, which has the same marginal σ for each copy of Ω.*
- *With probability $1-\rho$, it samples from $\sigma^{\otimes k}$.*

Let $F_1, \ldots, F_k \in \mathcal{L}_{[0,1]}(\Omega^L)$ such that $\mathbb{E}[F_i] \ge \epsilon > 0$ for all i. Then there exists $\zeta := \zeta(\rho, \epsilon, k) = \epsilon^{O_{\rho,k}(1)} > 0$ (independent of L) such that

$$\mathop{\mathbb{E}}_{x_1,\ldots,x_k}\left[\prod_{1\le i\le k} F_i(x_i)\right] \ge \zeta,$$

where for each $1 \le j \le L$, $\big((x_1)_j, \ldots, (x_k)_j\big)$ is sampled according to ν.

For each $1 \le j \le L$, $\big((\overline{x_{q,1}})_j, \ldots, (\overline{x_{q,k}})_j\big)$ is sampled according to $(\overline{\Omega}^k, \overline{\mu})$. $\overline{\mu}$ satisfies the requirement of Theorem 4.8 – with probability $\frac{1}{Q}$, it samples from $\omega^{\otimes kd}$, and with probability $\frac{Q-1}{Q}$, it samples from d permutations from $\mathbb{S}_k$ independently so that the marginal of each $(\overline{x_{q,i}})_j$ is $\omega^{\otimes d}$ for all i and j.

Therefore, we can apply Theorem 4.8 (setting $\Omega \leftarrow \overline{\Omega}$, $k \leftarrow k$, $\sigma \leftarrow \omega^{\otimes d}$, $\nu \leftarrow \overline{\mu}$, $\rho \leftarrow \frac{Q-1}{Q}$, $F_1 = \cdots = F_k \leftarrow \overline{f_q}$, $\epsilon \leftarrow \frac{\epsilon}{2}$) to conclude that there exists $\zeta := \zeta\big(\frac{Q-1}{Q}, \frac{\epsilon}{2}, k\big) = \epsilon^{O_{Q,k}(1)} > 0$ such that

$$\mathop{\mathbb{E}}_{x_{q,1},\ldots,x_{q,k}}\left[\prod_{1\le i\le k} f_q(x_{q,i})\right] = \mathop{\mathbb{E}}_{\overline{x_{q,1}},\ldots,\overline{x_{q,k}}}\left[\prod_{1\le i\le k} \overline{f_q}(\overline{x_{q,i}})\right] \ge \zeta.$$

The only properties of f_q used were $\mathbb{E}[f_q] \ge \frac{\epsilon}{2}$ and $f_q \in \mathcal{L}_{[0,1]}(L^R)$. For any $0 \le \gamma \le 1$, $T_{1-\gamma} f_q$ have the same properties, so we have the following lower bound for every $q \in [Q]$

$$\mathbb{E}\left[\prod_{1\le i\le k} T_{1-\gamma} f_q(x_{q,i})\right] \ge \zeta. \tag{2}$$

Step 3. **Smoothing Functions.** From unnoised functions to block noised functions, we use the following theorem from Mossel [23].

Theorem 5.4 ([23]). *Let $(\Omega_1 \times \cdots \times \Omega_K, \nu)$ be K correlated spaces with $\rho(\Omega_1, \ldots, \Omega_K; \nu) \le \rho < 1$. Consider K product spaces $((\Omega_1)^L \times \cdots \times (\Omega_K)^L, \nu^{\otimes L})$, and $F_i \in \mathcal{L}((\Omega_i)^L)$ for $i \in [K]$ such that $\mathsf{Var}[F_i] \le 1$. For every $\epsilon > 0$, there exists $\gamma := \gamma(\epsilon, \rho) > 0$ such that*

$$\left|\mathbb{E}\left[\prod_{1\le i\le K} F_i\right] - \mathbb{E}\left[\prod_{1\le i\le K} T_{1-\gamma} F_i\right]\right| \le K\epsilon.$$

Since $\rho\big(\overline{\Omega}^{Qk}, \overline{\mu}'\big) \leq \sqrt{\frac{Q-1}{Q}}$, we can apply the above theorem ($K \leftarrow Qk$, $\Omega_1 = \cdots = \Omega_K \leftarrow \overline{\Omega}$, $\nu \leftarrow \overline{\mu}'$, $\epsilon \leftarrow \frac{\zeta^Q}{4K}$, $F_{k(q-1)+i} \leftarrow \overline{f_q}$ for $q \in [Q]$ and $i \in [k]$) to have $\gamma := \gamma(Q,k,\zeta) \in (0,1)$ such that

$$(3) \quad \left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\leq q\leq Q, 1\leq i\leq k} f_q(x_{q,i}) \right] - \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\leq q\leq Q, 1\leq i\leq k} \overline{T}_{1-\gamma} f_q(x_{q,i}) \right] \right| \leq \frac{\zeta^Q}{4}.$$

From block noised functions to individual noised functions, we state the following general theorem inspired by Wenner [31]. The proof is in Appendix B.2.

Theorem 5.5. *Let $(\Omega_1^{d_1} \times \cdots \times \Omega_K^{d_K}, \nu)$ be joint probability spaces such that the marginal of each copy of Ω_i is ν_i, and the marginal of $\Omega_i^{d_i}$ is $\nu_i^{\otimes d_i}$. Fix $F_i \colon (\Omega_i^{d_i})^L \to \mathbb{R}$ for each $i = 1, \ldots, K$ with an associated projection $\pi_i \colon [d_i L] \to [L]$ such that $|\pi_i^{-1}(j)| = d_i$ for $1 \leq j \leq L$. For any $0 \leq \rho \leq 1$, the noise operator $T_\rho F_i$ and the block noise operator $\overline{T}_\rho F_i$ under π_i is defined as in Section 3. Fix a positive integer J and consider F_i^{bad} under π_i and J. Suppose $\max_{1\leq i\leq K} \|F_i\|_2 \leq 1$ and $\xi := \max_{1\leq i\leq K} \|F_i^{\mathsf{bad}}\|_2$. Then we have,*

$$\left| \mathop{\mathbb{E}}_{(x_1,\ldots,x_K)\sim\nu^{\otimes L}} \left[\prod_{1\leq i\leq K} \overline{T}_{1-\gamma} F_i(x_i) \right] - \mathop{\mathbb{E}}_{(x_1,\ldots,x_K)\sim\nu^{\otimes L}} \left[\prod_{1\leq i\leq K} T_{1-\gamma} F_i(x_i) \right] \right| \leq 2 \cdot 3^K ((1-\gamma)^J + \xi).$$

By applying the above theorem with $K \leftarrow Qk$, $L \leftarrow L$, $\Omega_1, \ldots, \Omega_K \leftarrow \Omega$, $d_1, \ldots, d_K \leftarrow d$, $\nu \leftarrow \overline{\mu}'$, $F_{k(q-1)+1} = \cdots = F_{k(q-1)+k} \leftarrow f_q$, $\pi_{k(q-1)+1} = \cdots = \pi_{k(q-1)+k} \leftarrow \pi_q$, $\xi \leftarrow (\frac{J^2}{T})^{1/4}$, we have

$$\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\leq q\leq Q, 1\leq i\leq k} \overline{T}_{1-\gamma} f_q(x_{q,i}) \right] - \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\leq q\leq Q, 1\leq i\leq k} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \leq 2 \cdot 3^{Qk} ((1-\gamma)^J + (\frac{J^2}{T})^{1/4}).$$

Fixing J and T to satisfy $2 \cdot 3^{Qk}((1-\gamma)^J + (\frac{J^2}{T})^{1/4}) \leq \frac{\zeta^Q}{4}$ as well as the previous constraint, and combining with (3), we can conclude that

$$(4) \quad \left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\leq q\leq Q, 1\leq i\leq k} f_q(x_{q,i}) \right] - \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\leq q\leq Q, 1\leq i\leq k} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \leq \frac{\zeta^Q}{2}.$$

In particular, if I induces less than $\frac{\zeta^Q}{4}$ fraction of hyperedges formed from e, from (4), we have

$$\mathop{\mathbb{E}}_{x_{q,i}}\left[\prod_{1\le q\le Q, 1\le i\le k} T_{1-\gamma} f_q(x_{q,i})\right] \le \frac{3\zeta^Q}{4}. \tag{5}$$

STEP 4. **Invariance.** We now want to show

$$\mathop{\mathbb{E}}_{x_{q,i}}\left[\prod_{1\le q\le Q, 1\le i\le k} T_{1-\gamma} f_q(x_{q,i})\right] \approx \prod_{1\le q\le Q} \mathop{\mathbb{E}}_{x_{q,i}}\left[\prod_{1\le i\le k} T_{1-\gamma} f_q(x_{q,i})\right],$$

unless f_q's share influential coordinates. Our invariance principle is similar to ones used in Wenner [31] and Chan [6]. With the goal of showing

$$\mathop{\mathbb{E}}_{x_1,\dots,x_K}\left[\prod_{1\le i\le K} F_i(x_i)\right] \approx \mathop{\mathbb{E}}_{x_1}[F_1(x_1)]\,\mathbb{E}\left[\prod_{2\le i\le K} F_i(x_i)\right],$$

one crucial property they used is that x_1 is independent of x_i for each $i = 2,\dots,K$ (even though any three x_i's are dependent).

Our $(x_{q,i})$ do not have such a property (any $x_{q,i}$ is dependent on $x_{q,i'}$ for $i \ne i'$), but it satisfies another property that any $x_{q,i}$ is independent of the joint distribution of $(x_{q',i'})_{q'\ne q, i'\in[k]}$ – everything not in the same hypercube. This property allows us to achieve the goal stated above. We formalize this intuition and prove the following general theorem, which will also be used in our other results. The proof appears in Appendix B.3.

Theorem 5.6. *Let $(\Omega_1^{k_1}\times\cdots\times\Omega_Q^{k_Q},\nu)$ be correlated spaces ($k_1,\dots,k_{Q-1}\ge 2$, $k_Q \ge 1$) where each copy of Ω_q has the same marginal and independent of $\prod_{q'\ne q}\Omega_{q'}^{k_q}$. Let $k_{\max} = \max_q k_q$ and $k_{\text{sum}} = \sum_q k_q$. For $1\le q\le Q$, let $F_q \in \mathcal{L}_{[0,1]}(\Omega_q^L)$. Suppose that for all $1\le q < Q$, $\sum_{1\le j\le L}\mathsf{Inf}_j[F_q] \le \Gamma$ and*

$$\sum_{1\le j\le L} \mathsf{Inf}_j[F_q](\mathsf{Inf}_j[F_{q+1}] + \cdots + \mathsf{Inf}_j[F_Q]) \le \tau.$$

Then,

$$\left|\mathop{\mathbb{E}}_{x_{q,i}}\left[\prod_{1\le q\le Q, 1\le i\le k_q} F_q(x_{q,i})\right] - \prod_{1\le q\le Q} \mathop{\mathbb{E}}_{x_{q,i}}\left[\prod_{1\le i\le k_q} F_q(x_{q,i})\right]\right| \le Q\cdot 2^{k_{\max}+1}\sqrt{\Gamma k_{\text{sum}}^2 \tau}.$$

By Lemma 1.13 of Wenner [31], there exists $\Gamma = O(\frac{1}{\gamma})$ such that

$$\sum_{1 \le j \le L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q] \le \sum_{1 \le j \le R} \mathsf{Inf}_j[T_{1-\gamma} f_q] \le \Gamma.$$

Fix τ to satisfy $Q \cdot 2^{k+1} \sqrt{\Gamma (Qk)^2 \tau} < \frac{\zeta^Q}{4}$. We have

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1 \le q \le Q, 1 \le i \le k} T_{1-\gamma} f_q(x_{q,i}) \right] - \prod_{1 \le q \le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1 \le i \le k} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \\
&\quad \ge \left| \prod_{1 \le q \le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1 \le i \le k} T_{1-\gamma} f_q(x_{q,i}) \right] \right| - \left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1 \le q \le Q, 1 \le i \le k} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \\
&\quad \ge \frac{\zeta^Q}{4} \quad \text{by (2) and (5).}
\end{aligned}$$

Thus, applying Theorem 5.6 with $Q \leftarrow Q$, $k_1 = \cdots = k_Q \leftarrow k, \Omega_1 = \cdots = \Omega_Q = \overline{\Omega}$, $\nu \leftarrow \overline{\mu}'$, $L \leftarrow L$, $F_q \leftarrow \overline{T_{1-\gamma} f_q}, \mathsf{Inf}_j[F_q] \leftarrow \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q]$, there exists $q \in \{1, \dots, Q-1\}$ such that

$$\sum_{1 \le j \le L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q](\overline{\mathsf{Inf}}_j[T_{1-\gamma} f_{q+1}] + \cdots + \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_Q]) > \tau. \tag{6}$$

STEP 5. **Decoding Strategy.** We use the standard strategy – each v_q samples a set $S \subseteq [R]$ according to $\|(f_q)_S\|_2^2$, and chooses a random element from S. For each $1 \le j \le L$, the probability that v chooses a label in $\pi^{-1}(j)$ is

$$\begin{aligned}
\sum_{S\colon S \cap \pi^{-1}(j) \neq \emptyset} \|(f_q)_S\|_2^2 \frac{|S \cap \pi^{-1}(j)|}{|S|} &\ge \sum_{S\colon S \cap \pi^{-1}(j) \neq \emptyset} \|(f_q)_S\|_2^2 \cdot \gamma (1-\gamma)^{\frac{|S|}{|S \cap \pi(j)|}} \\
&\ge \gamma \sum_{S\colon S \cap \pi^{-1}(j) \neq \emptyset} \|(f_q)_S\|_2^2 \cdot (1-\gamma)^{|S|} \\
&= \gamma \, \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q],
\end{aligned}$$

where the first inequality follows from the fact that $\alpha \ge \gamma (1-\gamma)^{1/\alpha}$ for $\alpha > 0$ and $0 < \gamma < 1$. Fix q to be the one obtained in Step 4 that satisfies (6). The probability that $\pi_q(l(v_q)) = \pi_{q'}(l(v_{q'}))$ for some $q < q' \le Q$ is at least

$$\gamma^2 \sum_{1 \le j \le L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q] \max_{q < q' \le Q} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_{q'}]$$

$$\geq \frac{\gamma^2}{Q} \sum_{1 \leq j \leq L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q](\overline{\mathsf{Inf}}_j[T_{1-\gamma} f_{q+1}] + \cdots + \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_Q])$$
$$\geq \frac{\gamma^2 \tau}{Q}.$$

Suppose that the total fraction of hyperedges (of E') wholly contained within I is less than $\frac{\delta}{4} \cdot \frac{\zeta^Q}{4} = \epsilon^{O_{Q,k}(1)}$. Since $\frac{\delta}{2}$ fraction of hyperedges (of E) are good, for at least $\frac{\delta}{2} - \frac{\delta}{4} = \frac{\delta}{4}$ fraction of hyperedges the above analysis works, and these edges are weakly satisfied by the above randomized strategy with probability $\frac{\gamma^2\tau}{Q}$. Setting the soundness parameter in Theorem 3.3 $\eta := \frac{\delta}{4} \cdot \frac{\gamma^2\tau}{Q}$ completes the proof of the soundness Lemma 5.2, and therefore also Theorem 1.1.

Dependencies between constants The above proof involves several constants that depend on each other. We summarize them in Table 2, in the order they are fixed in the proof.

Constants	How it is fixed	When it is fixed
Q, k, ϵ	Arbitrary $Q, k \geq 2$, $\epsilon > 0$	
δ	$\delta := \frac{(\frac{\epsilon}{2})^Q}{2^{Q+1}}$	STEP 1.
ζ	$\zeta := \zeta(Q, \epsilon, k)$ (by Theorem 4.8)	STEP 2.
γ	$\gamma := \gamma(Q, k, \zeta)$ (by Theorem 5.4)	STEP 3.
J, T	Large enough to satisfy $Q\left(\frac{J^2}{T}\right)^{1/4} \leq \frac{\delta}{2}$ $2 \cdot 3^{Qk}\left((1-\gamma)^J + \left(\frac{J^2}{T}\right)^{1/4}\right) \leq \frac{\zeta^Q}{4}$	STEP 3.
Γ	$\Gamma := O\left(\frac{1}{\gamma}\right)$ (by [31])	STEP 4.
τ	Small enough to satisfy $Q \cdot 2^{k+1}\sqrt{\Gamma(Qk)^2\tau} < \frac{\zeta^Q}{4}$	STEP 4.
η	$\eta := \frac{\delta}{4} \cdot \frac{\gamma^2\tau}{Q}$	STEP 5.

Table 2. List of the constants in the proof

Requirements for distributions. In the proof, we used the following three properties of the test distribution. We qualitatively describe them and how they are used in the proof. All the distributions in this work satisfy all the properties. Let $(\Omega_1 \times \cdots \times \Omega_K, \nu)$ be the test distribution for K points.

1. Let $\Omega_{i_1}, \ldots, \Omega_{i_k}$ correspond to k points queried in the same hypercube. We require the marginal distribution on $\Omega_{i_1} \times \cdots \times \Omega_{i_k}$ to have the *full support* – any $(x_{i_1}, \ldots, x_{i_k}) \in \Omega_{i_1} \times \cdots \times \Omega_{i_k}$ is sampled with nonzero probability. It is crucial in our application of the reverse hypercontractivity used in Step 2.
2. When $(x_1, \ldots, x_K)$ is sampled from ν, we require that for any $i \in [K]$, x_i is not always determined by the other $K-1$ points. This is used when bounding correlations and smoothing functions in Step 3.
3. When $(x_1, \ldots, x_K)$ is sampled from ν, we require that for any $i \in [K]$, x_i is completely independent from all the points not in the same hypercube. It is used in the application of the invariance principle in Step 4.

Acknowledgment. We thank Siu On Chan for sharing the latest version of his J. ACM paper [6] and explaining the underlying invariance principle.

References

[1] N. Alon, P. Kelsen, S. Mahajan and R. Hariharan: Approximate hypergraph coloring, *Nordic Journal of Computing* **3** (1996), 425–439.
[2] P. Austrin, V. Guruswami and J. Håstad: (2+ϵ)-SAT is NP-hard, in: *Proceedings of the 55th annual symposium on Foundations of Computer Science*, FOCS '14, 1–10, 2014.
[3] N. Bansal: Constructive algorithms for discrepancy minimization, in: *Proceedings of the 51st annual IEEE symposium on Foundations of Computer Science*, FOCS '10, 3–10. IEEE, 2010.
[4] N. Bansal and S. Khot: Inapproximability of hypergraph vertex cover and applications to scheduling problems, in: *Proceedings of the 37th International Colloquium on Automata, Languages and Programming*, ICALP '10, 250–261, 2010.
[5] B. Bollobás, D. Pritchard, T. Rothvoss and A. Scott: Cover-decomposition and polychromatic numbers, *SIAM Journal on Discrete Mathematics* **27** (2013), 240–256.
[6] S. O. Chan: Approximation resistance from pairwise independent subgroups, in: *Proceedings of the 45th annual ACM Symposium on Theory of Computing*, STOC '13, 447–456, 2013.
[7] M. Charikar, A. Newman and A. Nikolov: Tight hardness results for minimizing discrepancy, in: *Proceedings of the 22nd annual ACM-SIAM Symposium on Discrete Algorithms*, 1607–1614, 2011.
[8] H. Chen and A. M. Frieze: Coloring bipartite hypergraphs, in: *Proceedings of the 5th international conference on Integer Programming and Combinatorial Optimization*, IPCO 96, 345–358, 1996.

[9] I. Dinur and V. Guruswami: PCPs via low-degree long code and hardness for constrained hypergraph coloring, in: *Proceedings of the 54th annual symposium on Foundations of Computer Science*, FOCS 13, 340–349, 2013.

[10] I. Dinur, V. Guruswami, S. Khot and O. Regev: A new multilayered PCP and the hardness of hypergraph vertex cover, *SIAM Journal on Computing* **34** (2005), 1129–1146.

[11] I. Dinur, E. Mossel and O. Regev: Conditional hardness for approximate coloring, *SIAM Journal on Computing* **39** (2009), 843–873.

[12] I. Dinur, O. Regev and C. D. Smyth: The hardness of 3-Uniform hypergraph coloring, *Combinatorica* **25** (2005), 519–535.

[13] P. Gopalan, S. Khot and R. Saket: Hardness of reconstructing multivariate polynomials over finite fields, *SIAM Journal on Computing* **39** (2010), 2598–2621.

[14] V. Guruswami, J. Håstad, P. Harsha, S. Srinivasan and G. Varma: Superpolylogarithmic hypergraph coloring hardness via low-degree long codes, in: *Proceedings of the 46th annual ACM Symposium on Theory of Computing*, STOC '14, 2014.

[15] V. Guruswami, J. Håstad and M. Sudan: Hardness of approximate hypergraph coloring, *SIAM Journal on Computing* **31** (2002), 1663–1686.

[16] J. Håstad: On the NP-hardness of Max-Not-2, *SIAM Journal on Computing* **49** (2014), 179–193.

[17] G. H. Hardy, J. E. Littlewood and G. Pólya: *Inequalities*, Cambridge university press, 1952.

[18] J. Holmerin: Vertex cover on 4-regular hyper-graphs is hard to approximate within $2-\epsilon$, in: *Proceedings of the 34th annual ACM Symposium on Theory of Computing*, STOC '02, 544–552, 2002.

[19] S. Huang: Approximation resistance on satisfiable instances for predicates with few accepting inputs, in: *Proceedings of the 45th annual ACM symposium on Symposium on Theory of Computing*, STOC '13, 457–466, 2013.

[20] S. Khot: Hardness results for coloring 3-colorable 3-uniform hypergraphs, in: *Proceedings of the 43rd annual IEEE symposium on Foundations of Computer Science*, FOCS '02, 23–32. IEEE, 2002.

[21] S. Khot and R. Saket: Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs, in: *Proceedings of the 25th annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, 1607–1625, 2014.

[22] S. Lovett and R. Meka: Constructive discrepancy minimization by walking on the edges, in: *Proceedings of the 53rd annual IEEE symposium on Foundations of Computer Science*, FOCS 12, 61–67, 2012.

[23] E. Mossel: Gaussian bounds for noise correlation of functions, *Geometric and Functional Analysis* **19** (2010), 1713–1756.

[24] E. Mossel, R. O'Donnell, O. Regev, J. E. Steif and B. Sudakov: Non-interactive correlation distillation, inhomogeneous markov chains, and the reverse bonami-beckner inequality, *Israel Journal of Mathematics* **154** (2006), 299–336.

[25] E. Mossel, K. Oleszkiewicz and A. Sen: On reverse hypercontractivity, *Geometric and Functional Analysis* **23** (2013), 1062–1097.

[26] R. O'Donnell: *Analysis of Boolean Functions*, Cambridge University Press, 2014.

[27] R. O'Donnell and Y. Wu: Conditional hardness for satisfiable 3-CSPs, in: *Proceedings of the 41st annual ACM Symposium on Theory of Computing*, STOC '09, 493–502, 2009.

[28] S. Sachdeva and R. Saket: Optimal inapproximability for scheduling problems via structural hardness for hypergraph vertex cover, in: *Proceedings of the 28th annual IEEE Conference on Computational Complexity*, CCC '13, 219–229, 2013.
[29] R. Saket: Hardness of finding independent sets in 2-colorable hypergraphs and of satisfiable CSPs, in: *Proceedings of the 29th annual IEEE Conference on Computational Complexity*, CCC '14, 78–89, 2014.
[30] A. Samorodnitsky and L. Trevisan: Gowers uniformity, influence of variables, and PCPs, *SIAM Journal on Computing* **39** (2009), 323–360.
[31] C. Wenner: Circumventing d-to-1 for approximation resistance of satisfiable predicates strictly containing parity of width four, *Theory of Computing* **9** (2013), 703–757.

A. Variants of label cover

A.1. Hypergraph label cover

Hypergraph Label Cover is used in our main results on rainbow-colorable hypergraphs, including Theorem 1.1.

Theorem A.1 (Restatement of Theorem 3.3). *For any $Q \geq 2, T \geq 1$ and $\eta, \epsilon > 0$, given an instance of Q-Hypergraph Label Cover that is ϵ-weakly-dense and T-smooth, it is NP-hard to distinguish*

- *Completeness: There exists a labeling l that strongly satisfies every hyperedge.*
- *Soundness: No labeling l can weakly satisfy a fraction η of hyperedges.*

Proof. We reduce from T-smooth Label Cover first defined in Khot [20] to T-smooth Q-Hypergraph Label Cover using the technique of Gopalan et al. [13].

An instance of Label Cover consists of a biregular bipartite graph $G=(U \cup V, E)$ where each edge $e=(u,v)$ is associated with a projection $\pi_e\colon [R] \to [L]$ for some positive integers R and L. A labeling $l\colon U \cup V \to [R]$ satisfies e when $\pi_e(l(v)) = l(u)$. It is called T-smooth when for any $i \neq j$, $\Pr_e[\pi_e(i) = \pi_e(j)] \leq \frac{1}{T}$. The following theorem shows hardness of T-smooth Label Cover.

Theorem A.2 ([20]). *For any $T \geq 1$ and $\eta' > 0$, given an instance of Label Cover that is T-smooth, it is NP-hard to distinguish*

- *Completeness: There exists a labeling l that satisfies edge.*
- *Soundness: No labeling l can satisfy a fraction η' of hyperedges.*

We first claim that in Theorem [20], without loss of generality, we can assume that the degree d of $u \in U$ is large enough as a function of Q and ϵ, such that for any $\epsilon' \geq \frac{\epsilon}{2}$,

$$\frac{\binom{d\epsilon'}{Q}}{\binom{d}{Q}} = \prod_{i=0}^{Q-1} \frac{(d\epsilon' - i)}{d-i} \geq \frac{(\epsilon')^Q}{2}. \tag{7}$$

This is possible because in the construction of [20], the operations to increase T and reduce η' both increase the degree, so we can increase the degree while making T and η' even stronger for our purpose.

Given such an instance of Label Cover $G=(U_G \cup V_G, E_G)$, the corresponding instance of $H=(V_H, E_H)$ is produced by

- $V_H = V_G$.
- For $u \in U_G$ and Q distinct neighbors $v_1, \ldots, v_Q \in V_G$, we add a hyperedge $e = \{v_1, \ldots, v_Q\} \in E_H$ with the associated projections $\pi_{e,v_i} := \pi_{(u,v_i)}$. Say this hyperedge is *formed from* u. We can have the same hyperedges formed from different vertices.

Fix $v \in V_H$ and $i \neq j \in [R]$.

$$\Pr_{e \in E_H \colon v \in e}[\pi_{e,v}(i) = \pi_{e,v}(j)] = \Pr_{e=(u,v)\in E_G}[\pi_e(i) = \pi_e(j)] \leq \frac{1}{T},$$

so the resulting instance is also T-smooth.

For weak density, fix $I \subseteq V_H$ of measure ϵ, and for $u \in U_G$, let $\epsilon(u)$ be the fraction of neighbors of u contained in I. Biregularity of G implies $\epsilon = \mathbb{E}_u[\epsilon(u)]$. Let $\epsilon'(u) = \epsilon(u)$ if $\epsilon(u) \geq \frac{\epsilon}{2}$ and $\epsilon'(u) = 0$ otherwise. An averaging argument shows that $\mathbb{E}_u[\epsilon'(u)] \geq \frac{\mathbb{E}_u[\epsilon(u)]}{2} = \frac{\epsilon}{2}$. For any $u \in U_G$, whether $\epsilon'(u) = \epsilon(u) \geq \frac{\epsilon}{2}$ or $\epsilon'(u) = 0$, by (7), the fraction of hyperedges induced by I, out of the hyperedges formed from u, is at least

$$\frac{\binom{d\epsilon'(u)}{Q}}{\binom{d}{Q}} \geq \frac{\left(\epsilon'(u)\right)^Q}{2}.$$

Then the fraction of hyperedges induced by I is at least

$$\mathop{\mathbb{E}}_{u \in U_G}\left[\frac{\left(\epsilon'(u)\right)^Q}{2}\right] = \frac{1}{2}\mathop{\mathbb{E}}_{u \in U_G}\left[\left(\epsilon'(u)\right)^Q\right] \geq \frac{1}{2}\left(\mathop{\mathbb{E}}_{u \in U_G}[\epsilon'(u)]\right)^Q \geq \frac{\epsilon^Q}{2^{Q+1}}.$$

For completeness, given a labeling $l\colon U_G \cup V_G \to [R]$ that satisfies every edge of G, its projection to $V_G = V_H$ will strongly satisfy every hyperedge of H.

For soundness, let $l\colon V_H \to [R]$ be a labeling that weakly satisfies η fraction of hyperedges for some $\eta > 0$. Let $\eta(u)$ be the fraction of hyperedges satisfied by l formed from u, out of all hyperedges formed from u. Consider the following randomized strategy for G: V_G is labelled by l, and each $u \in U_G$ independently samples one of its neighbors v and set $l(u) \leftarrow \pi_{(u,v)}(l(v))$. The expected fraction of edges incident on u satisfied by this decoding strategy is (let $N(u)$ be the set of neighbors of u and $({}_{N(u)}P_Q)$ be the set of Q-tuples of the neighbors where Q vertices are pairwise distinct)

$$\begin{aligned}
&\mathop{\mathbb{E}}_{v_1 \in N(u)} \left[\Pr_{v_2 \in N(u)} [\pi_{(u,v_1)}\big(l(v_1)\big) = \pi_{(u,v_2)}(v_2)] \right] \\
&= \Pr_{(v_1,\dots,v_Q) \in N(u)^Q} \left[\pi_{(u,v_1)}\big(l(v_1)\big) = \pi_{(u,v_2)}(v_2)\right] \\
&\geq \Pr_{(v_1,\dots,v_Q) \in ({}_{N(u)}P_Q)} \left[\pi_{(u,v_1)}\big(l(v_1)\big) = \pi_{(u,v_2)}(v_2)\right] \\
&\geq \frac{1}{\binom{Q}{2}} \Pr_{(v_1,\dots,v_Q) \in ({}_{N(u)}P_Q)} [e := \{v_1, \dots, v_Q\} \text{ is weakly satisfied}] \\
&= \frac{1}{\binom{Q}{2}} \Pr_{\{v_1,\dots,v_Q\} \in \binom{N(u)}{Q}} [e := \{v_1, \dots, v_Q\} \text{ is weakly satisfied}] \\
&= \frac{\eta(u)}{\binom{Q}{2}}.
\end{aligned}$$

Overall, the strategy satisfies $\frac{\eta}{\binom{Q}{2}}$ fraction of edges of G in expectation. Setting $\eta' < \frac{\eta}{\binom{Q}{2}}$, we have contradiction, completing the proof of soundness. ∎

A.2. $(Q+1)$-Bipartite hypergraph label cover

Bipartite Hypergraph Label Cover is used in Theorem 1.6 for Q-out-of-$(2Q+1)$-SAT. An instance of $(Q+1)$-Bipartite Hypergraph Label Cover is based on a $(Q+1)$-uniform bipartite hypergraph $H = (U \cup V, E)$, where each hyperedge e contains one vertex from U and Q vertices from V. For every hyperedge $e = \{u, v_1, \dots, v_Q\}$ such that $u \in U$ and $v_q \in V$, each v_q is associated with a projection $\pi_{e,v_q} \colon [R] \to [L]$ for some positive integers R and L. A labeling $l \colon U \cup V \to [R]$ *strongly satisfies* $e = \{v_1, \dots, v_Q\}$ when $l(u) = \pi_{e,v_1}\big(l(v_1)\big) = \cdots = \pi_{e,v_Q}(l(v_Q))$ (we can imagine that $\pi_{e,u}$ is also defined as the identity). It *weakly satisfies* e when $\pi_{e,v_i}(l(v_i)) = \pi_{e,v_j}(l(v_j))$ for some $i \neq j$ or $\pi_{e,v_i}(l(v_i)) = l(u)$ for some i. As usual, the instance is T-smooth if for any $v \in V$ and $i \neq j$,

$$\Pr_{e \in E \colon v \in e} [\pi_{e,v}(i) = \pi_{e,v}(j)] \leq \frac{1}{T}.$$

Note that we do not need weak density for Q-out-of-$(2Q+1)$-SAT.

Theorem A.3. *For any $Q \geq 2, T \geq 1$ and $\eta > 0$, given an instance of $(Q+1)$-Bipartite Hypergraph Label Cover that is T-smooth, it is NP-hard to distinguish*

- *Completeness: There exists a labeling l that strongly satisfies every hyperedge.*
- *Soundness: No labeling l can weakly satisfy a fraction η of hyperedges.*

Proof. As in Theorem 3.3, we reduce from T-smooth Label Cover.

Given an instance of Label Cover $G = (U_G \cup V_G, E_G)$, the corresponding instance of $H = (U_H \cup V_H, E_H)$ is produced by

- $U_H = U_G, V_H = V_G$
- For $u \in U_G$ and Q distinct neighbors $v_1, \dots, v_Q \in V_G$, we add a hyperedge $e = \{u, v_1, \dots, v_Q\} \in E_H$ with the associated projections $\pi_{e,v_i} := \pi_{(u,v_i)}$. Say this hyperedge is *formed from* u.

Fix $v \in V_H$ and $i \neq j \in [R]$.

$$\Pr_{e \in E_H :\, v \in e}[\pi_{e,v}(i) = \pi_{e,v}(j)] = \Pr_{e=(u,v) \in E_G}[\pi_e(i) = \pi_e(j)] \leq \frac{1}{T},$$

so the resulting instance is also T-smooth.

For completeness, given a labeling $l \colon U_G \cup V_G \to [R]$ that satisfies every edge of G, it is easy to check that the same l will strongly satisfy every hyperedge of H.

For soundness, let $l \colon V_H \to [R]$ be a labeling that weakly satisfies η fraction of hyperedges for some $\eta > 0$. Let $\eta(u)$ be the fraction of hyperedges satisfied by l formed from u, out of all hyperedges formed from u. Consider the following randomized strategy for G:

- V_G is labeled by l.
- Each $u \in U_G$ is assigned $l(u)$ with probability half. With the remaining $1/2$ probability, it independently samples one of its neighbors v and sets $l(u) \leftarrow \pi_{(u,v)}(l(v))$.

Let $N(u)$ be the set of neighbors of u and $({}_{N(u)}P_Q)$ be the set of Q-tuples of the neighbors where Q vertices are pairwise distinct. The expected fraction of edges incident on u satisfied by this decoding strategy is

$$\frac{1}{2} \mathop{\mathbb{E}}_{v_1 \in N(u)}\left[\Pr_{v_2 \in N(u)}[\pi_{(u,v_1)}\big(l(v_1)\big) = \pi_{(u,v_2)}\big(l(v_2)\big)]\right] + \frac{1}{2} \Pr_{v \in N(u)}\big[\pi_{(u,v)}(l(v)) = l(u)\big]$$
$$= \frac{1}{2} \Pr_{(v_1,\dots,v_Q) \in N(u)^Q}\left[\pi_{(u,v_1)}\big(l(v_1)\big) = \pi_{(u,v_2)}(l(v_2)) \text{ or } \pi_{(u,v_1)}\big(l(v_1)\big) = l(u)\right]$$

$$\geq \frac{1}{2} \Pr_{(v_1,\ldots,v_Q)\in (_{N(u)}P_Q)} \left[\pi_{(u,v_1)}\big(l(v_1)\big) = \pi_{(u,v_2)}(l(v_2)) \text{ or } \pi_{(u,v_1)}\big(l(v_1)\big) = l(u)\right]$$

$$\geq \frac{1}{2\binom{Q}{2}} \Pr_{(v_1,\ldots,v_Q)\in (_{N(u)}P_Q)} [e := \{v_1,\ldots,v_Q\} \text{ is weakly satisfied}]$$

$$= \frac{1}{2\binom{Q}{2}} \Pr_{\{v_1,\ldots,v_Q\}\in \binom{N(u)}{Q}} [e := \{v_1,\ldots,v_Q\} \text{ is weakly satisfied}] = \frac{\eta(u)}{2\binom{Q}{2}}.$$

Overall, the strategy satisfies $\frac{\eta}{2\binom{Q}{2}}$ fraction of edges of G in expectation. Setting $\eta' < \frac{\eta}{2\binom{Q}{2}}$, we have contradiction, completing the proof of soundness. ∎

B. Proofs about influence, noise, and invariance

B.1. Influences

Lemma B.1 (Restatement of Lemma 3.2). *Let $(\Omega_1 \times \cdots \times \Omega_k, \mu)$ be k probability spaces and $(\Omega_1^L \times \cdots \times \Omega_k^L, \mu^{\otimes L})$ be its product space. Let $f_i\colon (\Omega_i)^L \to [-1,1]$, and $F\colon \Omega_1^L \times \cdots \times \Omega_k^L \to [-1,1]$ such that $F(x_1,\ldots,x_k) = \prod_{1\leq i\leq k} f_i(x_i)$. Then for $1\leq j\leq L$, $\mathsf{Inf}_j(F) \leq k\sum_{i=1}^k \mathsf{Inf}_j(f_i)$.*

Proof. We use $(x_i)_{-j} \in (\Omega_i)^{L-1}$ to denote x_i except the jth coordinate.

$$\begin{aligned}
&\mathsf{Inf}_j(F)\\
&= \mathop{\mathbb{E}}_{[(x_1)_{-j},\ldots,(x_k)_{-j}]} \mathop{\mathbb{E}}_{[(x_1)_j,\ldots,(x_k)_j,(x'_1)_j,\ldots,(x'_k)_j]} \left[(F(x_1,\ldots,x_k) - F(x'_1,\ldots,x'_k))^2\right]\\
&= \mathop{\mathbb{E}}_{[(x_1)_{-j},\ldots,(x_k)_{-j}]} \mathop{\mathbb{E}}_{[(x_1)_j,\ldots,(x_k)_j,(x'_1)_j,\ldots,(x'_k)_j]} \left[\left(\prod_i f_i(x_i) - \prod_i f_i(x'_i)\right)^2\right]\\
&\leq k\sum_i \mathop{\mathbb{E}}_{[(x_1)_{-j},\ldots,(x_k)_{-j}]} \mathop{\mathbb{E}}_{[(x_1)_j,\ldots,(x_k)_j,(x'_1)_j,\ldots,(x'_k)_j]} \left[(f_i(x_i) - f_i(x'_i))^2\right]\\
&= k\sum_i \mathop{\mathbb{E}}_{[(x_i)_{-j}]} \mathop{\mathbb{E}}_{[(x_i)_j,(x'_i)_j]} \left[(f_i(x_i) - f_i(x'_i))^2\right]\\
&= k\sum_i \mathsf{Inf}_j(f_i),
\end{aligned}$$

where the inequality follows from the fact that

$$\forall a_1,\ldots,a_k,b_1,\ldots,b_k \in [-1,1]\colon \left(\prod_i a_i - \prod_i b_i\right)^2 \leq k \cdot \sum_i (a_i - b_i)^2$$

proven in Lemma 4 of Samorodnitsky and Trevisan [30]. ∎

B.2. Block noise to individual noise

Theorem B.2 (Restatement of Theorem 5.5). *Let $(\Omega_1^{d_1}\times\cdots\times\Omega_K^{d_K},\nu)$ be joint probability spaces such that the marginal of each copy of Ω_i is ν_i, and the marginal of $\Omega_i^{d_i}$ is $\nu_i^{\otimes d_i}$. Fix $F_i\colon(\Omega_i^{d_i})^L\to\mathbb{R}$ for each $i=1,\dots,K$ with an associated projection $\pi_i\colon[d_iL]\to[L]$ such that $|\pi_i^{-1}(j)|=d_i$ for $1\le j\le L$. For any $0\le\rho\le1$, the noise operator $T_\rho F_i$ and the block noise operator $\overline{T}_\rho F_i$ under π_i is defined as in Section 3. Fix a positive integer J and consider F_i^{bad} under π_i and J. Suppose $\max_{1\le i\le K}\|F_i\|_2\le1$ and $\xi:=\max_{1\le i\le K}\|F_i^{\mathsf{bad}}\|_2$. Then we have,*

$$\left|\mathop{\mathbb{E}}_{(x_1,\dots,x_K)\sim\mu^{\otimes L}}\left[\prod_{1\le i\le K}\overline{T}_{1-\gamma}F_i(x_i)\right]-\mathop{\mathbb{E}}_{(x_1,\dots,x_K)\sim\mu^{\otimes L}}\left[\prod_{1\le i\le K}T_{1-\gamma}F_i(x_i)\right]\right|$$
$$\le 2\cdot3^K((1-\gamma)^J+\xi).$$

Proof. For each $1\le i\le K$, we decompose F_i as follows:

$$F_i^{\mathsf{shattered}}=\sum_{S\subseteq[d_iL]\colon S\text{ shattered under }\pi_i}(F_i)_S$$
$$F_i^{\mathsf{large}}=\sum_{S\subseteq[d_iL]\colon S\text{ not shattered and }|\pi_i(S)|\ge J}(F_i)_S$$
$$F_i^{\mathsf{bad}}=\sum_{S\subseteq[d_iL]\colon S\text{ not shattered and }|\pi_i(S)|<J}(F_i)_S.$$

Consider $C:=\{\mathsf{shattered},\mathsf{large},\mathsf{bad}\}^K$. Expanding $F_i=(F_i^{\mathsf{shattered}}+F_i^{\mathsf{large}}+F_i^{\mathsf{bad}})$, we have

$$\prod_{1\le i\le K}\overline{T}_{1-\gamma}F_i=\sum_{c\in C}\prod_{1\le i\le K}\overline{T}_{1-\gamma}F_i^{c_i}$$

and

$$\prod_{1\le i\le K}T_{1-\gamma}F_i=\sum_{c\in C}\prod_{1\le i\le K}T_{1-\gamma}F_i^{c_i}.$$

The quantity we want to bound can be also decomposable as

$$\left|\sum_{c\in C}\mathbb{E}\left[\prod_{1\le i\le K}\overline{T}_{1-\gamma}F_i^{c_i}-\prod_{1\le i\le K}T_{1-\gamma}F_i^{c_i}\right]\right|.$$

Since $\overline{T}_{1-\gamma}F_i^{\mathsf{shattered}}=T_{1-\gamma}F_i^{\mathsf{shattered}}$, the contribution of the case $c=\{\mathsf{shattered}\}^K$ is 0. We bound the other two cases of c.

- $c_{i'} = \mathsf{large}$ for some i':

$$\left|\mathbb{E}\left[\prod_{1\le i\le K}\overline{T}_{1-\gamma}F_i^{c_i}\right]\right| \le \left\|\overline{T}_{1-\gamma}F_{i'}^{\mathsf{large}}\right\|_2 \left\|\prod_{i\ne i'}\overline{T}_{1-\gamma}F_i^{c_i}\right\|_2 \le (1-\gamma)^J \left\|F_{i'}^{\mathsf{large}}\right\|_2 \le (1-\gamma)^J.$$

  Similarly, $\left|\mathbb{E}\left[\prod_{1\le i\le K} T_{1-\gamma}F_i^{c_i}\right]\right| \le (1-\gamma)^J$ and the contribution from such c is at most $2(1-\gamma)^J$.
- $c_{i'} = \mathsf{bad}$ for some i':

$$\left|\mathbb{E}\left[\prod_{1\le i\le K}\overline{T}_{1-\gamma}F_i^{c_i}\right]\right| \le \left\|\overline{T}_{1-\gamma}F_{i'}^{\mathsf{bad}}\right\|_2 \left\|\prod_{i\ne i'}\overline{T}_{1-\gamma}F_i^{c_i}\right\|_2 \le \xi.$$

  Similarly, $\left|\mathbb{E}\left[\prod_{1\le i\le K} T_{1-\gamma}F_i^{c_i}\right]\right| \le \xi$ and the contribution from such c is at most 2ξ.

Since there are at most 3^K choices for c, the total error is bounded by $2\cdot 3^K((1-\gamma)^J+\xi)$. ∎

B.3. Invariance

The following lemma is the basic building block that enables the induction used in proof of the main invariance principle (Theorem 5.6) used in our framework. It is essentially implied by a theorem stated in a more general setup by Wenner [31, Theorem 3.12]. For completeness, we present a proof below in simpler notation that fits for our purposes.

Lemma B.3. *Let $(\Omega_1^k \times \Omega_2, \nu)$ be $(k+1)$ correlated spaces $(k \ge 2)$ such that each copy of Ω_1 has the same marginal, and any one copy of Ω_1 and Ω_2 are independent. Let $F \in \mathcal{L}_{[0,1]}(\Omega_1^L)$, and $G \in \mathcal{L}(\Omega_2^L)$. Suppose that $\sum_{1\le j\le L}\mathsf{Inf}_j[F] \le \Gamma$ and*

$$\sum_{1\le j\le L}\mathsf{Inf}_j[F]\,\mathsf{Inf}_j[G] \le \tau.$$

Then,

$$\left|\mathop{\mathbb{E}}_{x_1,\dots,x_k,y}\left[\prod_{1\le i\le k}F(x_i)G(y)\right] - \mathop{\mathbb{E}}_{x_1,\dots,x_k,y}\left[\prod_{1\le i\le k}F(x_i)\right]\mathop{\mathbb{E}}_{y}[G(y)]\right| \le 2^{k+1}\sqrt{\Gamma\tau}.$$

Proof. Let ν' be the distribution where the marginals of Ω_1^k and Ω_2 are the same as those of ν, but Ω_1^k and Ω_2 are independent. Fix $j \in [L]$. Let $(x_1,\ldots,x_k,y)$ be sampled such that $((x_1)_{j'},\ldots,(x_k)_{j'},y_{j'}) \sim \nu$ for $j' < j$ and $((x_1)_{j'},\ldots,(x_k)_{j'},y_{j'}) \sim \nu'$ for $j' \geq j$. Let $(x'_1,\ldots,x'_k,y')$ be the same except that $((x'_1)_j,\ldots,(x'_k)_j,y_j) \sim \nu$. We want to bound

$$\left| \mathop{\mathbb{E}}_{x_1,\ldots,x_k,y} \left[\prod_{1\leq i\leq k} F(x_i)G(y) \right] - \mathop{\mathbb{E}}_{x'_1,\ldots,x'_k,y'} \left[\prod_{1\leq i\leq k} F(x'_i)G(y') \right] \right|,$$

since the LHS with $j=1$ and the RHS with $j=L$ are the two expectations we are interested in.

Decompose F into the following two parts.

$$F^{\mathsf{relevant}} = \sum_{S\colon j\in S} F_S$$
$$F^{\mathsf{not}} = \sum_{S\colon j\notin S} F_S.$$

Note that $\|F^{\mathsf{relevant}}\|_2^2 = \mathsf{Inf}_j[F]$. Decompose $G = G^{\mathsf{relevant}} + G^{\mathsf{not}}$ in the same way. Let $C = \{\mathsf{relevant}, \mathsf{not}\}^{k+1}$. The term we wanted to bound now becomes

$$\left| \sum_{c\in C} \left(\mathop{\mathbb{E}}_{x_1,\ldots,x_k,y} \left[\prod_{1\leq i\leq k} F^{c_i}(x_i)G^{c_{k+1}}(y) \right] - \mathop{\mathbb{E}}_{x'_1,\ldots,x'_k,y'} \left[\prod_{1\leq i\leq k} F^{c_i}(x'_i)G^{c_{k+1}}(y') \right] \right) \right|. \tag{8}$$

If $c_{k+1} = \mathsf{not}$ or $c_1 = \cdots = c_k = \mathsf{not}$, the contribution from c is zero because the marginals of $((x_1)_j,\cdots,(x_k)_j)$ and y_j are the same with those of $((x'_1)_j,\ldots,(x'_k)_j)$ and y'_j respectively. Furthermore, the same conclusion holds when $c_{k+1} = \mathsf{relevant}$ and exactly one of $c_1,\ldots,c_k$ is $\mathsf{relevant}$, since one copy of Ω_1 and Ω_2 are independent and $((x_i)_j,y_j)$ and $((x'_i)_j,y'_j)$ have the same distribution. Thus a $c\in C$ with nonzero contribution to (8) must satisfy $c_{i_1} = c_{i_2} = c_{k+1} = \mathsf{relevant}$ for some $i_1 \neq i_2$. For such c,

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{x_1,\ldots,x_k,y} \left[\prod_{1\leq i\leq k} F^{c_i}(x_i)G^{c_{k+1}}(y) \right] \right| \\ &\leq \left\| F^{\mathsf{relevant}}(x_{i_1})G^{\mathsf{relevant}}(y) \right\|_2 \left\| F^{\mathsf{relevant}}(x_{i_2}) \right\|_2 \left\| \prod_{i\neq i_1,i_2} F^{c_i} \right\|_\infty \end{aligned}$$

by Hölder inequality

$$= \left\| F^{\mathsf{relevant}} \right\|_2 \left\| G^{\mathsf{relevant}} \right\|_2 \left\| F^{\mathsf{relevant}} \right\|_2 \left\| \prod_{i \neq i_1, i_2} F^{c_i} \right\|_\infty$$

by independence

$$\leq \sqrt{\mathsf{Inf}_j[F]^2 \, \mathsf{Inf}_j[G]},$$

where the last inequality used the fact that $F^{\mathsf{not}}(x) = \mathbb{E}_{x'}[F(x') | x'_{[L]\setminus j} = x_{[L]\setminus j}] \in [0,1]$ and $F^{\mathsf{relevant}}(x) = F(x) - F^{\mathsf{not}}(x) \in [-1,1]$. There are at most 2^k choices for such c and

$$\left| \mathop{\mathbb{E}}_{x'_1,\ldots,x'_k,y} \left[\prod_{1 \leq i \leq k} F^{c_i}(x'_i) G^{c_{k+1}}(y') \right] \right| \leq \sqrt{\mathsf{Inf}_j[F]^2 \, \mathsf{Inf}_j[G]}$$

can be shown similarly, so

$$\left| \mathop{\mathbb{E}}_{x_1,\ldots,x_k,y} \left[\prod_{1 \leq i \leq k} F(x_i) G(y) \right] - \mathop{\mathbb{E}}_{x'_1,\ldots,x'_k,y'} \left[\prod_{1 \leq i \leq k} F(x'_i) G(y') \right] \right| \leq 2^{k+1} \sqrt{\mathsf{Inf}_j[F]^2 \, \mathsf{Inf}_j[G]}.$$

Summing over all $1 \leq j \leq J$, we conclude that

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{x_1,\ldots,x_k,y} \left[\prod_{1 \leq i \leq k} F(x_i) G(y) \right] - \mathop{\mathbb{E}}_{x_1,\ldots,x_k} \left[\prod_{1 \leq i \leq k} F(x_i) \right] \mathop{\mathbb{E}}_{y}[G(y)] \right| \\ &\leq 2^{k+1} \sum_{1 \leq j \leq L} \sqrt{\mathsf{Inf}_j[F]^2 \, \mathsf{Inf}_j[G]} \\ &\leq 2^{k+1} \sqrt{\sum_{1 \leq j \leq L} \mathsf{Inf}_j[F] \, \mathsf{Inf}_j[G]} \sqrt{\sum_{1 \leq j \leq L} \mathsf{Inf}_j[F]} \qquad \text{(by Cauchy-Schwartz)} \\ &\leq 2^{k+1} \sqrt{\Gamma \tau}. \end{aligned}$$

∎

Theorem B.4 (Restatement of Theorem 5.6). *Let $(\Omega_1^{k_1} \times \cdots \times \Omega_Q^{k_Q}, \nu)$ be correlated spaces ($k_1, \ldots, k_{Q-1} \geq 2$, $k_Q \geq 1$) where each copy of Ω_q has the same marginal and independent of $\prod_{q' \neq q} \Omega_{q'}^{k_q}$. Let $k_{\max} = \max_q k_q$ and*

$k_{\text{sum}} = \sum_q k_q$. *For* $1 \le q \le Q$, *let* $F_q \in \mathcal{L}_{[0,1]}(\Omega_q^L)$. *Suppose that for all* $1 \le q < Q$, $\sum_{1\le j\le L} \mathsf{Inf}_j[F_q] \le \Gamma$ *and*

$$\sum_{1\le j\le L} \mathsf{Inf}_j[F_q](\mathsf{Inf}_j[F_{q+1}] + \cdots + \mathsf{Inf}_j[F_Q]) \le \tau.$$

Then,

$$\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\le q\le Q, 1\le i\le k_q} F_q(x_{q,i}) \right] - \prod_{1\le q\le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\le i\le k_q} F_q(x_{q,i}) \right] \right| \le Q \cdot 2^{k_{\max}+1} \sqrt{\Gamma k_{\text{sum}}^2 \tau}.$$

Proof. We use induction on Q. When $Q = 2$, the application of Lemma B.3 (setting $F \leftarrow F_1$, $k \leftarrow k_1$, $\Omega_2 \leftarrow \Omega_2^{k_2}$, $G(x_{2,1}, \ldots, x_{2,k_2}) \leftarrow \prod_{1\le i\le k_2} F_2(x_{2,i})$) and applying Lemma 3.2 to have $\mathsf{Inf}_j[G] \le k_2^2 \mathsf{Inf}_j[F_2]$) implies the theorem.

Assuming the theorem holds for $Q-1$, the application of Lemma B.3 with

- $F \leftarrow F_1$, $k \leftarrow k_1$, $\Omega_2 \leftarrow \Omega_2^{k_2} \times \cdots \times \Omega_Q^{k_Q}$, $G(x_{q,i}) \leftarrow \prod_{2\le q\le Q, 1\le i\le k_2} F_q(x_{q,i})$
- $\mathsf{Inf}_j[G] \le k_{\text{sum}}^2(\mathsf{Inf}_j[F_2] + \cdots + \mathsf{Inf}_j[F_Q])$ by Lemma 3.2

gives

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\le q\le Q, 1\le i\le k_q} F_q(x_{q,i}) \right] - \prod_{1\le q\le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\le i\le k_q} F_q(x_{q,i}) \right] \right| \\
&\le \left| \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\le q\le Q, 1\le i\le k_q} F_q(x_{q,i}) \right] \right. \\
&\qquad \left. - \mathop{\mathbb{E}}_{x_{1,i}} \left[\prod_{1\le i\le k_1} F_1(x_{1,i}) \right] \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{2\le q\le Q, 1\le i\le k_q} F_q(x_{q,i}) \right] \right| \\
&\quad + \left| \prod_{1\le q\le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1\le i\le k_q} F_q(x_{q,i}) \right] \right. \\
&\qquad \left. - \mathop{\mathbb{E}}_{x_{1,i}} \left[\prod_{1\le i\le k_1} F_1(x_{1,i}) \right] \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{2\le q\le Q, 1\le i\le k_q} F_q(x_{q,i}) \right] \right| \\
&\le 2^{k_{\max}+1} \sqrt{\Gamma k_{\text{sum}}^2 \tau} + (Q-1) 2^{k_{\max}+1} \sqrt{\Gamma k_{\text{sum}}^2 \tau} \\
&= Q \cdot 2^{k_{\max}+1} \sqrt{\Gamma k_{\text{sum}}^2 \tau}.
\end{aligned}$$

∎

C. K-Hypergraph vertex cover

In this section, we prove the following two theorems, both implying that it is NP-hard to approximate K-Hypergraph Vertex Cover with in a factor of $K-1-\epsilon$.

Theorem C.1 (Restatement of Theorem 1.4). *For any $\epsilon > 0$ and $K \geq 3$, given a K-uniform hypergraph $H = (V, E)$, it is NP-hard to distinguish the following cases.*

- *Completeness: There is a vertex cover of measure $\frac{1}{K-1}$.*
- *Soundness: Every $I \subseteq V$ of measure ϵ induces at least a fraction $\epsilon^{O_K(1)}$ of hyperedges.*

Theorem C.2 (Restatement of Theorem 1.5). *For any $\epsilon > 0$ and $K \geq 3$, given a K-uniform hypergraph $H = (V, E)$, it is NP-hard to distinguish the following cases.*

- *Completeness: There exist $V^* \subseteq V$ of measure ϵ and a coloring $c\colon [V \setminus V^*] \to [K-1]$ such that for every hyperedge of the induced hypergraph on $V \setminus V^*$, $K-2$ colors appear once and the other color twice. Therefore, H has a vertex cover of size at most $\frac{1}{K-1} + \epsilon$.*
- *Soundness: There is no independent set of measure ϵ.*

The above two theorems are not comparable to each other. In the completeness case, Theorem 1.4 ensures a smaller vertex cover, while Theorem 1.5 guarantees richer structure. In the soundness case, Theorem 1.4 gives a stronger density. Since they differ only in the test distribution, we prove Theorem 1.5 in details and introduce the distribution for Theorem 1.4 at the end of this section.

C.1. Multilayered label cover

We reduce Multilayered Label Cover defined by Dinur et al. [10] with the smoothness property to K-Hypergraph Vertex Cover. An instance of Multilayered Label Cover with A layers is based on a graph $G = (V, E)$ where $V = V_1 \cup \cdots \cup V_A$ and $E = \cup_{1 \leq i < j \leq A} E_{i,j}$. Let $[R_i]$ be the label set of the variables in the V_i such that R_i divides R_j for all $i < j$. Any edge $e \in E_{i,j}$ is between $u \in V_i$ and $v \in V_j$, and associated with a projection $\pi_e\colon [R_j] \to [R_i]$. Given a labeling $l\colon V \to [R_A]$, an edge $e = (u, v)$ with $u \in V_i$ and $v \in V_j$ $(i < j)$ is satisfied when $\pi_e(l(v)) = l(u)$. The following are desired properties of an instance. Note that the definition of weak density here is not parameterized by ϵ.

- Weakly dense: for any $\epsilon > 0$ satisfying $\lceil \frac{4}{\epsilon} \rceil \leq A$, given $m = \lceil \frac{4}{\epsilon} \rceil$ layers $i_1 < \cdots < i_m$ and given any sets $I_{i_j} \subseteq V_{i_j}$ with $|I_{i_j}| \geq \epsilon |V_{i_j}|$, there exist $j < j'$ such that at least $\frac{\epsilon^3}{16}$ fraction of the edges between V_{i_j} and $V_{i_{j'}}$ are indeed between I_{i_j} and $I_{i_{j'}}$.
- T-smooth: for any $1 \leq i < j \leq A$, $v \in V_j$ and $a \neq b \in [R_j]$,

$$\Pr_{u \in V_i:\ (u,v) \in E_{i,j}} [\pi_{u,v}(a) = \pi_{u,v}(b)] \leq \frac{1}{T}.$$

Theorem C.3 ([20]). *For every $\eta > 0$, $A \geq 2$ and $T \geq 1$, given an instance of Multilayered Label Cover with A layers that is weakly dense and T-smooth, it is NP-hard to distinguish the following cases:*

- *Completeness: There exists a labeling l that satisfies every edge.*
- *Soundness: No labeling l can satisfy a fraction η of any $E_{i,j}$.*

C.2. Distribution

We first define the distribution of K points, one in a single cell and the other $K-1$ in a block of size d. Let $\Omega = \{*, 1, \ldots, K-1\}$ and $\overline{\Omega} = \Omega^d$. Let ω be the distribution on Ω such that $\omega(*) = \epsilon$ and $\omega(1) = \cdots = \omega(K-1) = \frac{1-\epsilon}{K-1}$. The K points $x \in \Omega$ and $y_1, \ldots, y_{K-1} \in \overline{\Omega}$ are sampled by the following procedure.

- Sample $x \sim \omega$.
- If $x = *$, sample $y_1, \ldots, y_{K-1} \sim \omega^{\otimes d}$ independently.
- If $x \neq *$, for each $1 \leq j \leq d$, sample $(y_1)_j, \ldots, (y_{K-1})_j \sim \mathbb{S}_{K-1}$ uniformly, and independently noise $(y_i)_j \leftarrow *$ with probability ϵ.

It is easy to see that the marginal distribution of each y_i is $\omega^{\otimes d}$. Let $(\Omega \times \overline{\Omega}^{K-1}, \overline{\mu}')$ denote the K correlated spaces corresponding to the above distribution, and let $\overline{\mu}$ denote the marginal distribution of $(y_1, \ldots, y_{K-1})$. Let $\overline{\Omega}_i$ $(1 \leq i \leq K-1)$ denote the copy of $\overline{\Omega}$ associated with y_i, and $\overline{\Omega}'_i$ be the product of the other $K-1$ spaces. With probability ϵ (when $x = *$), y_i is completely independent of the others. Even when $x \neq *$, y_i's marginal is $\omega^{\otimes d}$. By Lemma 3.1, we conclude that $\rho(\overline{\Omega}_i, \overline{\Omega}'_i; \overline{\mu}') \leq \sqrt{1-\epsilon}$.

However, bounding $\rho(\Omega, \overline{\Omega}^{K-1}; \overline{\mu}')$ (as the correlation between two spaces Ω and $\overline{\Omega}^{K-1}$) cannot be done in the same way. To get around this, we define the distribution $\overline{\mu}'_\beta$ be the same as $\overline{\mu}'$, but at the end each y_i is independently resampled with probability $1-\beta$. While we still use $\overline{\mu}'$ in the reduction, the fact that $\rho(\overline{\Omega}_i, \overline{\Omega}'_i; \overline{\mu}') \leq \sqrt{1-\epsilon}$ implies that our analysis, without much loss, can assume that each y_i is resampled as in $\overline{\mu}'_\beta$. In $\overline{\mu}'_\beta$, the same technique yields $\rho(\Omega, \overline{\Omega}^{K-1}; \overline{\mu}'_\beta) \le \sqrt{1-(1-\beta)^{K-1}}$, which allows the usual analysis to proceed.

C.3. Reduction and completeness

We now describe the reduction from Multilayered Label Cover with A layers. Given a $G = (\cup_{1 \le i \le A} V_i, \cup_{i<j} E_{i,j})$ with a projection $\pi_e \colon [R_j] \to [R_i]$ for each hyperedge $e = (u,v)$ $(u \in V_i, v \in V_j)$, the resulting instance for K-Hypergraph Vertex Cover is (V', E'), where $V' = \cup_{1 \le i \le A} V_i \times \Omega^{R_i}$. The weight of (v,x) $(v \in V_i)$ is $\prod_{1 \le j \le R_i} \omega(x_j)$, so that the sum of the weights of the vertices in $\mathsf{Cloud}(v)$ is 1. For $v \in V_i$, let $\mathsf{Cloud}(v) := \{v\} \times \Omega^{R_i}$. The set of hyperedges E' is described by the following procedure.

- Sample $1 \le a < b \le A$ uniformly and $e = (u,v) \in E_{i,j}$ such that $u \in V_i$, $v \in V_j$.
- Sample $x \in \Omega^{R_a}, y_1, \dots, y_{K-1} \in \Omega^{R_b}$ in the following way. For each $1 \le j \le R_a$, sample $x_j, ((y_i)_{\pi_e^{-1}(j)})_{i \in [K-1]}$ from $(\Omega \times \overline{\Omega}^{K-1}, \overline{\mu}')$.
- Add a hyperedge $((u,x),(v,y_1),\dots,(v,y_{K-1}))$ to E'. We say that this hyperedge is *formed from* e, and the weight of this hyperedge is the probability that it is sampled given that e is sampled in the first step.

Given the reduction, completeness is easy to show.

Lemma C.4. *If there is a labeling that satisfies every $e \in E$, there exist $V^* \subseteq V'$ of measure ϵ and $c \colon V' \backslash V^* \to [K-1]$ with the same measure for each color, such that in each hyperedge induced by $V' \setminus V^*$, $K-1$ colors appear once and the other color appears twice.*

Proof. Let $l \colon V \to [R_A]$ be a labeling that satisfies every edge in E. Let $V^* := \{(v,x) \colon (x)_{l(v)} = *\}$, and $c(v,x) = (x)_{l(v)}$. In each $\mathsf{Cloud}(v)$, V^* contains measure $\omega(*) = \epsilon$ and $c(i)$ contains $\omega(i) = \frac{1-\epsilon}{K-1}$. For each hyperedge $((u,x),(v,y_1),\dots,(v,y_{K-1}))$ induced by $V' \setminus V^*$, $\{(v,y_1)_{l(v)}, \dots, (v,y_{K-1})_{l(v)}\} = [K-1]$. ∎

C.4. Soundness

Unlike the previous reductions, the resulting instance is weighted – vertices and hyperedges can have different weights. The only reason is that (1) we used Multilyaered Label Cover and (2) and ω is not the uniform distribution. Once we fix an edge e of G, our hyperedge weights correspond to the above probability distribution and vertex weights correspond to its marginals. Therefore all the following probabilistic analysis works as in previous reductions.

Lemma C.5. *For any $\epsilon > 0$, there exists $\eta := \eta(\epsilon, K)$ such that if $I \subseteq V'$ of measure ϵ induces less than $\epsilon^{O_{Q,k}(1)}$ fraction of hyperedges, the corresponding instance of Multilayered Label Cover admits a labeling that satisfies η fraction of edges in $E_{a,b}$ for some $1 \leq a < b \leq A$.*

The proof is almost identical to the one presented in Section 5.3, with slightly more technical details dealing with noise.

STEP 1. *Fixing a good hyperedge.* Let $I \subseteq V'$ be of measure ϵ. Let f_v be the indicator function of $I \cap \mathsf{Cloud}(v)$. By averaging, $\frac{\epsilon}{2}$ fraction of vertices has $\mathbb{E}[f_v] \geq \frac{\epsilon}{2}$ – call these vertices *heavy*. Let $W_i \subseteq V_i$ be the set of heavy vertices in the ith layer.

By averaging, at least $\frac{\epsilon}{4}$ fraction of layers satisfy $|W_i| \geq \frac{\epsilon}{4}|V_i|$. Take $A = \lceil \frac{\epsilon}{16} \rceil$. By weak density, there exist $1 \leq a < b \leq A$ such that the fraction of edges in $E_{i,j}$ induced by W_a and W_b is at least $\frac{\epsilon^3}{1024}$. Let $L = R_a$ and $R = R_b$.

By the same argument as in Section 5.3, by adjusting the smoothness paramter T and an integer J, we can ensure that $\frac{\epsilon^3}{2048}$ fraction of edge $(u,v) \in E_{a,b}$ is good – both u and v are heavy and,

$$\|f_v^{\mathsf{bad}}\|_2 \leq \left(\frac{J^2}{T}\right)^{1/4}$$

under π_e and J.

Throughout the rest of the section, fix such an edge $e = (u,v)$ and the associated projections $\pi := \pi_e$. For simplicity, let $f := f_u$ and $g := f_v$. We now measure the weight of hyperedges induced by I, which is

$$\mathop{\mathbb{E}}_{x, y_1, \ldots, y_{K-1}} \left[f(x) \prod_{1 \leq i \leq K-1} g(y_i) \right]. \tag{9}$$

STEP 2. *Lower bounding in each hypercube.* For each $1 \leq j \leq L$, with probability ϵ, $(y_i)_{\pi^{-1}(j)}$ are sampled completely independently from $\overline{\Omega}$. By Theorem 4.8 (setting $\Omega \leftarrow \overline{\Omega}$, $k \leftarrow K-1$, $\sigma \leftarrow \omega^{\otimes d}$, $\nu \leftarrow \overline{\mu}$, $\rho \leftarrow 1-\epsilon$, $F_1 = \cdots = F_{K-1} \leftarrow g$, $\epsilon \leftarrow \frac{\epsilon}{2}$), there exists $\zeta = \zeta(\epsilon, K) > 0$ such that for every $\gamma \in [0,1]$,

$$\mathop{\mathbb{E}}_{y_1, \ldots, y_K \sim \overline{\mu}^{\otimes L}} \left[\prod_{1 \leq i \leq K-1} T_{1-\gamma} g(y_i) \right] \geq \zeta.$$

Note that $\overline{\mu}_\beta$ also satisfies the requirement of Theorem 4.8, so

$$\mathop{\mathbb{E}}_{y_1,\dots,y_K\sim(\overline{\mu}_\beta)^{\otimes L}}\left[\prod_{1\le i\le K-1} T_{1-\gamma} g(y_i)\right] \ge \zeta. \tag{10}$$

Let $\theta := \frac{\epsilon\zeta}{2}$ be the lower bound of $\mathbb{E}[f(x)]\,\mathbb{E}[\prod_i g(y_i)]$, which also holds for any noised versions of f, g and noised distributions.

Step 3. *Smoothing functions.* Due to the fact that $\rho(\Omega, \overline{\Omega}^{K-1}; \overline{\mu}')$ is not easily bounded, we insert the noise operator for $g(y_1), \dots, g(y_{K-1})$ first using $\rho(\overline{\Omega}_i, \overline{\Omega}'_i; \overline{\mu}') \le \sqrt{1-\epsilon}$ for $1 \le i \le K-1$. This follows from the following lemma from Mossel [23], which is indeed the main lemma for Theorem 5.4.

Lemma C.6 ([23]). *Let $(\Omega_1 \times \Omega_2, \nu)$ be two correlated spaces with $\rho(\Omega_1, \Omega_2; \nu) \le \rho < 1$, and the corresponding product spaces $((\Omega_1)^L \times (\Omega_2)^L, \nu^{\otimes L})$, and $F_i \in \mathcal{L}((\Omega_i)^L)$ for $i = 1, 2$ such that $\mathsf{Var}[F_i] \le 1$. For any $\epsilon > 0$, there exists $\gamma := \gamma(\epsilon, \rho) > 0$ such that*

$$|\,\mathbb{E}[F_1 F_2] - \mathbb{E}[F_1 T_{1-\gamma} F_2]| \le \epsilon.$$

Applying the above lemma to $(\overline{\Omega}_i, \overline{\Omega}'_i; \overline{\mu}')$ iteratively for $i = 1, \dots, K-1$, we have $\gamma_1 := \gamma_1(\epsilon, K, \theta)$ such that

$$\begin{aligned}
&\left|\mathop{\mathbb{E}}_{x,y_i\sim\overline{\mu}'^{\otimes L}}\left[f(x)\prod_{1\le i\le K-1} g(y_i)\right] - \mathop{\mathbb{E}}_{x,y_i\sim\overline{\mu}'^{\otimes L}}\left[f(x)\prod_{1\le i\le K-1} \overline{T}_{1-\gamma_1}\overline{T}_{1-\gamma_1} g(y_i)\right]\right| \\
&= \left|\mathop{\mathbb{E}}_{x,y_i\sim\overline{\mu}'^{\otimes L}}\left[f(x)\prod_{1\le i\le K-1} g(y_i)\right] - \mathop{\mathbb{E}}_{x,y_i\sim(\overline{\mu}'_{1-\gamma_1})^{\otimes L}}\left[f(x)\prod_{1\le i\le K-1} \overline{T}_{1-\gamma_1} g(y_i)\right]\right| \\
&\le \frac{\theta}{8}.
\end{aligned}$$

Let $\beta := 1-\gamma_1$, and use $\hat{\mathbb{E}}$ to denote the expectation over $(x, y_1, \dots, y_K) \sim (\overline{\mu}'_\beta)^{\otimes L}$ while $\mathbb{E}$ still denotes the expectation over $(x, y_1, \dots, y_K) \sim \overline{\mu}'^{\otimes L}$. This implies

$$\mathbb{E}\left[f(x)\prod_{1\le i\le K-1} \overline{T}_{1-\gamma_1}\overline{T}_{1-\gamma_1} g(y_i)\right] = \hat{\mathbb{E}}\left[f(x)\prod_{1\le i\le K-1} \overline{T}_{1-\gamma_1} g(y_i)\right].$$

Since $\rho(\Omega,\overline{\Omega}^{K-1};\overline{\mu}'_\beta)\le\sqrt{1-(1-\beta)^{K-1}}$, another application of Lemma C.6 will give γ_2 such that

$$\left|\hat{\mathbb{E}}\left[f(x)\prod_{1\le i\le K-1}\overline{T}_{1-\gamma_1}g(y_i)\right]-\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\prod_{1\le i\le K-1}\overline{T}_{1-\gamma_1}g(y_i)\right]\right|\le\frac{\theta}{8}.$$

By applying Theorem 5.5 ($K\leftarrow K$, $L\leftarrow L$, $\Omega_1,\dots,\Omega_K\leftarrow\Omega$, $\Omega_K=\Omega$, $d_1,\dots,d_{K-1}\leftarrow d$, $d_K=1$, $\nu\leftarrow\overline{\mu}'_\beta$, $F_1=\cdots=F_{K-1}\leftarrow g$, $F_K\leftarrow f$, $\pi_1=\cdots=\pi_{K-1}=\pi$, $\pi_K\leftarrow$ the identity, $\xi\leftarrow(\frac{J^2}{T})^{1/4}$), we have

$$\left|\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\prod_{1\le i\le K-1}T_{1-\gamma_1}g(y_i)\right]-\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\prod_{1\le i\le K-1}\overline{T}_{1-\gamma_1}g(y_i)\right]\right|$$
$$\le 2\cdot 3^K\left((1-\gamma_1)^J+\left(\frac{J^2}{T}\right)^{1/4}\right).$$

Fixing J and T to satisfy $2\cdot3^K((1-\gamma_1)^J+(\frac{J^2}{T})^{1/4})\le\frac{\theta}{8}$ as well as the previous constraint, we can conclude that

$$\left|\mathbb{E}\left[f(x)\prod_{1\le i\le K-1}g(y_i)\right]-\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\prod_{1\le i\le K-1}T_{1-\gamma_1}g(y_i)\right]\right|\le\frac{3\theta}{8}.\tag{11}$$

In particular, if I is independent, from (9) and (11)

$$\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\prod_{1\le i\le K-1}T_{1-\gamma_1}g(y_i)\right]\le\frac{\theta}{2}.\tag{12}$$

STEP 4. *Invariance.* The marginal of y_i (resp. x) is $\omega^{\otimes R}$ (resp. $\omega^{\otimes L}$) on both $\overline{\mu}'^{\otimes L}$ and $\overline{\mu}^{\otimes L}$. Therefore, the Efron-Stein decomposition of f and g as well as the notion of (block) influence remain the same between $\overline{\mu}'$ and $\overline{\mu}'_\beta$. Since g is noised, there exists $\Gamma=O(\frac{1}{\gamma_1})$ such that

$$\sum_{1\le j\le L}\overline{\mathsf{Inf}}_j[T_{1-\gamma_1}g]\le\Gamma.$$

Fix τ to satisfy $Q\cdot2^{K+1}\sqrt{\Gamma K^2\tau}<\frac{\theta}{4}$. From (10) and (12),

$$\left|\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\prod_{1\le i\le K-1}T_{1-\gamma_1}g(y_i)\right]-\hat{\mathbb{E}}\left[T_{1-\gamma_2}f(x)\right]\hat{\mathbb{E}}\left[\prod_{1\le i\le K-1}T_{1-\gamma_1}g(y_i)\right]\right|$$

$$\geq \hat{\mathbb{E}}[T_{1-\gamma_2} f(x)]\, \hat{\mathbb{E}}\left[\prod_{1\leq i\leq K-1} T_{1-\gamma_1} g(y_i)\right] - \hat{\mathbb{E}}\left[T_{1-\gamma_2} f(x) \prod_{1\leq i\leq K-1} T_{1-\gamma_1} g(y_i)\right]$$
$$\geq \frac{\theta}{2}.$$

Applying Theorem 5.6 ($Q \leftarrow 2$, $k_1 \leftarrow K-1$, $k_2 = 1$, $\Omega_1 = \overline{\Omega}$, $\Omega_2 \leftarrow \Omega$, $\nu \leftarrow \overline{\mu}'_\beta$, $L \leftarrow L$, $F_1 \leftarrow \overline{T_{1-\gamma_1} g}$, $F_2 \leftarrow T_{1-\gamma_2} f$, $\mathsf{Inf}_j[F_1] \leftarrow \overline{\mathsf{Inf}}_j[T_{1-\gamma_1} g]$),

$$\sum_{1\leq j\leq L} \overline{\mathsf{Inf}}_j[T_{1-\gamma_1} g]\, \mathsf{Inf}_j[T_{1-\gamma_2} f] > \tau.$$

STEP 5. *Decoding strategy.* We use the following standard strategy – v samples a set $S \subseteq [R]$ according to $\|g_S\|_2^2$, and chooses a random element from S. u also samples a set $S \subseteq [L]$ according to $\|f_S\|_2^2$, and chooses a random element from S. As shown in Section 5.3, for each $1 \leq j \leq L$, the probability that v chooses a label in $\pi^{-1}(j)$ is at least $\gamma_1 \overline{\mathsf{Inf}}_j[T_{1-\gamma_1} g]$, and the probability that u chooses j is at least $\gamma_2 \mathsf{Inf}[T_{1-\gamma_2} f]$.

The probability that $\pi_e(l(v)) = \pi(l(u))$ is at least

$$\gamma_1\gamma_2 \sum_{1\leq j\leq L} \overline{\mathsf{Inf}}_j[T_{1-\gamma_1} g]\, \mathsf{Inf}_j[T_{1-\gamma_2} f] \geq \gamma_1\gamma_2\tau.$$

Suppose that I is indepenent. For at least $\frac{\epsilon^3}{2048}$ fraction of edges (of $E_{a,b}$) the above analysis works, and these edges are satisfied by the above randomized strategy with probability $\gamma_1\gamma_2\tau$. Setting $\eta := \frac{\epsilon^3}{2048} \cdot \gamma_1\gamma_2\tau$ completes the proof of soundness.

C.5. Distribution for Theorem 1.4

For Theorem 1.4, we again define the distribution of K points, one in a single cell and the other $K-1$ in a block of size d. Let $\Omega = \{0,1\}$ and $\overline{\Omega} = \Omega^d$. Let ω be the $(1 - \frac{1}{K-1})$-biased distribution on Ω – $\omega(0) = \frac{1}{K-1}$ and $\omega(1) = 1 - \frac{1}{K-1}$. The K points $x \in \Omega$ and $y_1, \ldots, y_{K-1} \in \overline{\Omega}$ are sampled by the following procedure.

- Sample $x \sim \omega$.
- If $x = 0$, sample $y_1, \ldots, y_{K-1} \sim \omega^{\otimes d}$ independently.
- If $x = 1$, for each $1 \leq j \leq d$, sample $(y_1)_j, \ldots, (y_{K-1})_j \sim \mu$, where μ is the uniform distribution on $K-1$ bit strings with exactly $(K-2)$ 1's.

$\Pr[(y_i)_j = 1] = \frac{1}{K-1} \cdot \left(1 - \frac{1}{K-1}\right) + \left(1 - \frac{1}{K-1}\right)\left(\frac{K-2}{K-1}\right) = \left(1 - \frac{1}{K-1}\right)$ for all $i \in [K-1]$ and $j \in [d]$, and $(y_i)_1, \ldots, (y_i)_d$ are independent. Let $(\Omega \times \overline{\Omega}^{K-1}, \overline{\mu}')$ denote the K correlated spaces corresponding to the above distribution, and let $\overline{\mu}$ denote the marginal distribution of $(y_1, \ldots, y_{K-1})$. Let $\overline{\Omega}_i$ $(1 \leq i \leq K-1)$ denote the copy of $\overline{\Omega}$ associated with y_i, and $\overline{\Omega}'_i$ be the product of the other $K-1$ spaces. With probability $\frac{1}{K-1}$ (when $x = 0$), y_i is completely independent of the others. Even when $x = 1$, y_i's marginal is $\omega^{\otimes d}$. By Lemma 3.1, we conclude that $\rho(\overline{\Omega}_i, \overline{\Omega}'_i; \overline{\mu}') \leq \sqrt{\frac{K-2}{K-1}}$. Bounding $\rho(\Omega, \overline{\Omega}^{K-1}; \overline{\mu}')$ (as the correlation between two spaces Ω and $\overline{\Omega}^{K-1}$) can be done in the same way as the proof of Theorem 1.4 in this section: (1) define the distribution $\overline{\mu}'_\beta$ for the sake of analysis where each y_i is independently resampled with probability $1-\beta$ after sampled according to $\overline{\mu}'$, (2) show that analyzing $\overline{\mu}'_\beta$ instead of $\overline{\mu}'$ incurs little extra error, and (3) use the standard technique to prove $\rho(\Omega, \overline{\Omega}^{K-1}; \overline{\mu}'_\beta) \leq \sqrt{1 - (1-\beta)^{K-1}}$.

The fact that for each $1 \leq j \leq d$, at least one of $x, (y_1)_j, \ldots, (y_K)_j$ is 1 ensures completeness, and the bounded correlation ensures soundness. Furthermore, the fact that $y_1, \ldots, y_{K-1}$ become completely independent with probability $\frac{1}{K-1}$ (previously this was ϵ) implies $\zeta := \epsilon^{O_K(1)}$ and the same argument in Theorem 1.1 shows density in soundness.

D. Q-out-of-$(2Q+1)$-SAT

An instance of $(2Q+1)$-SAT is a tuple (V, Φ) consisting of the set of variables V and the set of clauses Φ. Each clause ϕ is described by $((v_1, z_1), \ldots, (v_{2Q+1}, z_{2Q+1}))$ where $v_q \in V$ and $z_q \in \{0,1\}$. To be consistent with the notation we used for hypergraph coloring, we use the unconventional notation where 0 denotes True and 1 denotes False. Let $f\colon V \to \{0,1\}$ be an assignment to variables. The number of literals of ϕ set to True by f is $\left|\{q\colon f(v_q) \oplus z_q = 0\}\right|$ where $\oplus$ denotes the sum over $\mathbb{Z}_2$.

D.1. Distribution

We first define the distribution of $2Q+1$ points, one in a single cell and the other $2Q$ in a block of size d. Let $\Omega = \{0,1\}$ and $\overline{\Omega} = \Omega^d$. Let ω be the uniform distribution on Ω. $2Q+1$ points $x_0 \in \Omega$ and $x_{q,i} \in \overline{\Omega}$ for $1 \leq q \leq Q$ and $1 \leq i \leq k$ are sampled by the following procedure.

- Sample $q' \in \{0, \ldots, Q\}$ uniformly at random.

- If $q'=0$,
  - Sample $x_0 \in \Omega$ uniformly independently.
  - For all $q \in [Q]$, sample $x_{q,1} \in \Omega^d$ independently and set $x_{q,2} = \mathbf{1}_d - x_{q,1}$, where $\mathbf{1}_d \in \Omega^d := (1,1,\dots,1)$.
- If $q'>0$,
  - For all $q \in [Q] \setminus \{q'\}$, sample $x_{q,1} \in \Omega^d$ independently and set $x_{q,2} = \mathbf{1}_d - x_{q,1}$.
  - Sample $x_0 \in \Omega$ independently. If $x_0 = 0$, sample $x_{q,1}, x_{q,2} \in \Omega^d$ independently. If $x_0 = 1$, sample $x_{q,1} \in \Omega^d$ independently and set $x_{q,2} = \mathbf{1}_d - x_{q,1}$.

Let $(\Omega \times \overline{\Omega}^{2Q}, \overline{\mu}')$ denote $2Q+1$ correlated spaces corresponding to the above distribution, and $\overline{\mu}$ denote the marginal distribution of $(x_{q,1}, x_{q,2})$, which is the same for all $q \in [Q]$. We bound $\rho(\Omega, \overline{\Omega}^{2Q}; \overline{\mu}')$.

Fix some $1 \le q \le Q$ and $1 \le i \le 2$. Let $\overline{\Omega}_{q,i}$ denote the copy of $\overline{\Omega}$ associated with $x_{q,i}$, and $\overline{\Omega}'_{q,i}$ be the product of the other $2Q$ copies. We have $\overline{\mu}' = \frac{1}{2(Q+1)}\alpha_q + \left(1 - \frac{1}{2(Q+1)}\right)\beta_q$ where α_q denotes the distribution given $q' = q$ and $x_0 = 0$ (so that $x_{q,1}, x_{q,2}$ are sampled i.i.d.), and β_q denotes the distribution $q' \ne q$ or $x_0 = 1$. Since each entry of $x_{q,i}$ is sampled i.i.d. in α_q, $\rho(\overline{\Omega}_{q,i}, \overline{\Omega}'_{q,i}; \alpha_q) = 0$. In both α_q and β_q, the marginal of $x_{q,i}$ is $\omega^{\otimes d}$. By Lemma 3.1, we conclude that $\rho(\overline{\Omega}_{q,i}, \overline{\Omega}'_{q,i}; \overline{\mu}') \le \sqrt{1 - \frac{1}{2(Q+1)}}$. Similarly, $\rho(\Omega, \overline{\Omega}^{2Q}; \overline{\mu}') \le \sqrt{1 - \frac{1}{Q+1}}$. Therefore we have

$$\rho(\Omega, (\overline{\Omega}_{q,i})_{q,i}; \overline{\mu}') \le \sqrt{1 - \frac{1}{2(Q+1)}}.$$

D.2. Reduction and completeness

We now describe the reduction from $(Q+1)$-Bipartite Hypergraph Label Cover. Given a $(Q+1)$-uniform hypergraph $H = (U \cup V, E)$ with Q projections from $[R]$ to $[L]$ for each hyperedge, the resulting instance for $(2Q+1)$-SAT is $(U' \cup V', \Phi)$ where $U' := (U \times \Omega^L)$ and $V' := (V \times \Omega^R)$. For $u \in U$ and $v \in V$, let $\mathsf{Cloud}(u) := \{u\} \times \Omega^L$ and $\mathsf{Cloud}(v) := \{v\} \times \Omega^R$. The clauses in Φ are described by the following procedure.

- Sample a random hyperedge $e = (u, v_1, \dots, v_Q)$ with associated projections $\pi_{e,v_1}, \dots, \pi_{e,v_Q}$ from E.
- Sample $x_0 \in \Omega^L, (x_{q,i})_{1 \le q \le Q, 1 \le i \le 2} \in \Omega^R$ in the following way. For each $1 \le j \le L$, sample $(x_0)_j, \left((x_{q,i})_{\pi^{-1}_{e,v_q}(j)}\right)_{q,i}$ from $(\Omega \times \overline{\Omega}^{2Q}, \overline{\mu}')$.

- Sample $z_0, (z_{q,i})_{1\le q\le Q, 1\le i\le 2} \in \Omega$ i.i.d.
- Add a clause

$$\big((u, x_0 \oplus z_0\, \mathbf{1}_L), z_0\big) \times \big((v_q, x_{q,i} \oplus z_{q,i}\, \mathbf{1}_R), z_{q,i}\big)_{1\le q\le Q, 1\le i\le 2}$$

to Φ. We say this clause is *formed from* $e \in E$.

Given the reduction, complteness is easy to show.

Lemma D.1. *If an instance of* $(Q+1)$*-Bipartite Hypergraph Label Cover admits a labeling that strongly satisfies every hyperedge* $e\in E$*, there is an assignment* $f\colon U'\cup V' \to \Omega$ *that sets at least* Q *literals to* 0 *(which denotes True in our convention) in every clause of* Φ*.*

Proof. Let $l\colon U\cup V \to [R]$ be a labeling that strongly satisfies every hyperedge $e\in E$. For any $u\in U, x\in \Omega^L$, let $f(u,x) = x_{l(u)}$. For any $v\in V, x\in \Omega^R$, let $f(v,x) = x_{l(v)}$. For any clause

$$\big((u, x_0 \oplus z_0\, \mathbf{1}_L), z_0\big) \times \big((v_q, x_{q,i} \oplus z_{q,i}\, \mathbf{1}_R), z_{q,i}\big)_{1\le q\le Q, 1\le i\le 2},$$

one of the following is true. Note that $f(u, x_0 \oplus z_0\, \mathbf{1}_L) \oplus z_0 = (x_0)_{l(u)}$ and $f(v_q, x_{q,i} \oplus z_{q,i}\, \mathbf{1}_R) \oplus z_{q,i} = (x_{q,i})_{l(v_q)}$.

- Each $q\in[Q]$ satisfies $(x_{q,1})_{l(v_q)} \ne (x_{q,2})_{l(v_q)}$.
- For some $q\in[Q]$, all $q'\in[Q]\setminus\{q\}$ satisfy $(x_{q',1})_{l(v'_q)} \ne (x_{q',2})_{l(v'_q)}$, and if $(x_0)_{l(u)} = 1$, q also satisfies $(x_{q,1})_{l(v_q)} \ne (x_{q,2})_{l(v_q)}$.

In any case, $(2Q+1)$-tuple $\big((x_0)_{l(u)}\big) \times \big((x_{q,i})_{l(v_q)}\big)_{q,i}$ contains at least Q zeros, which means that any clause has at least Q literals set True. ∎

D.3. Soundness

Lemma D.2. *There exist* $\epsilon, \eta > 0$*, only depending on* Q*, such that if there is an assignment that satisfies more than* $(1-\epsilon)$ *fraction of hyperedges, the corresponding instance of* Q*-Hypergraph Label Cover admits a labeling that weakly satisfies* η *fraction of hyperedges.*

The proof is almost identical to the one presented in Section 5.3. Let $g\colon U'\cup V' \to \Omega$ be any assignment. The fraction of clauses whose literals are all set to False is

$$\mathop{\mathbb{E}}_{u, v_1, \dots, v_Q}\ \mathop{\mathbb{E}}_{x_0, (x_q, i)}\ \mathop{\mathbb{E}}_{z_0, (z_{q,i})} \left[(g(u, x_0 \oplus \mathbf{1}_L\, z_0) \oplus z_0) \prod_{\substack{1\le q\le Q,\\ 1\le i\le 2}} (g(v_q, x_{q,i} \oplus \mathbf{1}_R\, z_{q,i}) \oplus (z_0)) \right]$$

$$= \mathop{\mathbb{E}}_{u,v_1,\ldots,v_Q} \mathop{\mathbb{E}}_{x_0,(x_q,i)} \left[\mathop{\mathbb{E}}_{z_0}[(g(u, x_0 \oplus \mathbf{1}_L z_0) \oplus z_0)] \prod_{\substack{1 \le q \le Q, \\ 1 \le i \le 2}} \mathop{\mathbb{E}}_{z_{q,i}} [g(v_q, x_{q,i} \oplus \mathbf{1}_R z_{q,i}) \oplus z_{q,i}] \right]$$

$$= \mathop{\mathbb{E}}_{u,v_1,\ldots,v_Q} \mathop{\mathbb{E}}_{x_0,(x_q,i)} \left[f(u, x_0) \prod_{\substack{1 \le q \le Q, \\ 1 \le i \le 2}} f(v, x_{q,i}) \right],$$

where we define

$$f(u,x) := \mathop{\mathbb{E}}_{z \in \Omega} \big[f(u, x \oplus \mathbf{1}_L z) \oplus z) \big] \qquad u \in U$$
$$f(v,x) := \mathop{\mathbb{E}}_{z \in \Omega} \big[f(v, x \oplus \mathbf{1}_R z) \oplus z) \big] \qquad v \in V.$$

For $u \in U$, let $f_u \in \mathcal{L}_{[0,1]}(\Omega^L)$ be the restriction of f to $\{u\} \times \Omega^L$, and define $f_v \in \mathcal{L}_{[0,1]}(\Omega^R)$ similarly for $v \in V$. Note that $\mathbb{E}[f_u] = \mathbb{E}[f_v] = \frac{1}{2}$.

STEP 1. *Fixing a good hyperedge.* Since $\mathbb{E}[f_u] = \mathbb{E}[f_v] = \frac{1}{2}$ for all $u \in U$, and $v \in V$, we do not need to define heavy vertices. By the same argument as in Section 5.3, by adjusting the smoothness paramter T and the integer J, we can ensure that $\delta := \frac{1}{2}$ fraction of hyperedges are good for every vertex they contain, i.e., the hyperedge $e = (u, v_1, \ldots, v_Q)$ satisfies for each $q \in [Q]$,

$$\| f_{v_q}^{\mathsf{bad}} \|_2 \le \left(\frac{J^2}{T} \right)^{1/4}$$

under π_{e,v_q} and J.

Throughout the rest of the section, fix such a hyperedge $e = (u, v_1, \ldots, v_Q)$ and the associated projections $\pi_{e,v_1}, \ldots, \pi_{e,v_Q}$. For simplicity, let $f_q := f_{v_q}$ and $\pi_q := \pi_{e,v_q}$ for $q \in [Q]$, and $f_{q+1} = f_u$. We now measure the fraction of clauses formed from e that are unsatisfied, which is

$$\mathop{\mathbb{E}}_{x_{q,i}} \left[f_u(x_0) \prod_{1 \le q \le Q, 1 \le i \le 2} f_q(x_{q,i}) \right]. \tag{13}$$

STEP 2. *Lower bounding in each hypercube.* Fix any $q \in [Q]$. For each $1 \le j \le L$, with probability $\frac{1}{2(Q+1)}$, $(x_{q,1})_{\pi_q^{-1}(j)}$ and $(x_{q,2})_{\pi_q^{-1}(j)}$ are sampled completely independently from $\overline{\Omega}$. By Theorem 4.8 (setting $\Omega \leftarrow \overline{\Omega}$, $k \leftarrow 2$,

$\sigma \leftarrow \omega^{\otimes d}$, $\nu \leftarrow \overline{\mu}$, $\rho \leftarrow \sqrt{\frac{2Q+1}{2(Q+1)}}$, $F_1 = F_2 \leftarrow \overline{f_q}$, $\epsilon \leftarrow \frac{1}{2}$), there exists $\zeta = \zeta(Q) > 0$ such that for every $\gamma \in [0,1]$,

$$\mathop{\mathbb{E}}_{x_{q,1},x_{q,2}} [T_{1-\gamma} f_q(x_{q,1}) T_{1-\gamma} f_q(x_{q,2})] \geq \zeta. \tag{14}$$

STEP 3. *Smoothing functions.* Since $\rho(\Omega, (\overline{\Omega}_{q,i})_{q,i}; \overline{\mu}') \leq \sqrt{1 - \frac{1}{2(Q+1)}}$, we can apply Theorem 5.4 ($K \leftarrow 2Q+1$, $\Omega_1 = \cdots = \Omega_{K-1} \leftarrow \overline{\Omega}$, $\Omega_K \leftarrow \Omega$, $\nu \leftarrow \overline{\mu}'$, $\epsilon \leftarrow \frac{\zeta^Q}{8K}$, $F_{2q-1} = F_{2q} \leftarrow \overline{f_q}$, $F_K \leftarrow f_u$) to have $\gamma := \gamma(Q, \zeta) \in (0,1)$ such that

$$\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[f_u(x_0) \prod_{\substack{1 \leq q \leq Q, \\ 1 \leq i \leq 2}} f_q(x_{q,i}) \right] - \mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{\substack{1 \leq q \leq Q, \\ 1 \leq i \leq 2}} \overline{T}_{1-\gamma} f_q(x_{q,i}) \right] \right| \leq \frac{\zeta^Q}{8}. \tag{15}$$

By applying Theorem 5.5 ($K \leftarrow 2Q+1$, $L \leftarrow L$, $\Omega_1, \ldots, \Omega_K \leftarrow \Omega$, $d_1, \ldots, d_{K-1} \leftarrow d$, $d_K = 1$, $\nu \leftarrow \overline{\mu}'$, $F_{2q-1} = F_{2q} \leftarrow f_q$, $F_K \leftarrow f_u$, $\pi_{2q-1} = \pi_{2q} \leftarrow \pi_q$, $\pi_K \leftarrow$ the identity, $\xi \leftarrow (\frac{J^2}{T})^{1/4}$), we have

$$\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{\substack{1 \leq q \leq Q, \\ 1 \leq i \leq 2}} \overline{T}_{1-\gamma} f_q(x_{q,i}) \right] - \mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{\substack{1 \leq q \leq Q, \\ 1 \leq i \leq 2}} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \leq 2 \cdot 3^{2Q+1} \left((1-\gamma)^J + \left(\frac{J^2}{T} \right)^{1/4} \right). \tag{16}$$

Fixing J and T to satisfy $2 \cdot 3^{2Q+1} \big((1-\gamma)^J + (\frac{J^2}{T})^{1/4} \big) \leq \frac{\zeta^Q}{8}$ as well as the previous constraint, we can conclude from (15) and (16) that

$$\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[f_u(x_0) \prod_{\substack{1 \leq q \leq Q, \\ 1 \leq i \leq 2}} f_q(x_{q,i}) \right] - \mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{\substack{1 \leq q \leq Q, \\ 1 \leq i \leq 2}} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \leq \frac{\zeta^Q}{4}.$$

In particular, if among the clauses formed from e, less than $\frac{\zeta^Q}{8}$ fraction of them are unsatisfied, from (13),

$$\mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{1 \leq q \leq Q, 1 \leq i \leq 2} T_{1-\gamma} f_q(x_{q,i}) \right] \leq \frac{3\zeta^Q}{8}. \tag{17}$$

STEP 4. *Invariance.* Since our functions are noised, there exists $\Gamma = O(\frac{1}{\gamma})$ such that

$$\sum_{1 \le j \le L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q] \le \Gamma.$$

Fix τ to satisfy $8Q \cdot \sqrt{\Gamma(2Q+1)^2\tau} < \frac{\zeta^Q}{8}$. We have

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{1 \le q \le Q, 1 \le i \le 2} T_{1-\gamma} f_q(x_{q,i}) \right] \right. \\ &\qquad \left. - \mathbb{E}[T_{1-\gamma} f_u] \prod_{1 \le q \le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1 \le i \le 2} T_{1-\gamma} f_q(x_{q,i}) \right] \right| \\ &\ge \mathbb{E}[T_{1-\gamma} f_u] \cdot \prod_{1 \le q \le Q} \mathop{\mathbb{E}}_{x_{q,i}} \left[\prod_{1 \le i \le 2} T_{1-\gamma} f_q(x_{q,i}) \right] \\ &\qquad - \mathop{\mathbb{E}}_{x_{q,i}} \left[T_{1-\gamma} f_u(x_0) \prod_{1 \le q \le Q, 1 \le i \le 2} T_{1-\gamma} f_q(x_{q,i}) \right] \\ &\ge \frac{1}{2}\zeta^Q - \frac{3\zeta^Q}{8} = \frac{\zeta^Q}{8} \quad \text{(using (14) and (17))}. \end{aligned}$$

Now, applying Theorem 5.6 ($Q \leftarrow Q+1$, $k_1 = \cdots = k_Q \leftarrow k$, $k_{Q+1} \leftarrow 1$, $\Omega_1 = \cdots = \Omega_Q = \overline{\Omega}$, $\Omega_{Q+1} \leftarrow \Omega$, $\nu \leftarrow \overline{\mu}'$, $L \leftarrow L$, $F_q \leftarrow \overline{T_{1-\gamma} f_q}$ for $q \in [Q]$, $F_{Q+1} \leftarrow T_{1-\gamma} f_u$, $\mathsf{Inf}_j[F_q] \leftarrow \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q]$ for $q \in [Q]$), there exists $q \in \{1, \ldots, Q\}$ such that

$$\sum_{1 \le j \le L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q] \big(\overline{\mathsf{Inf}}_j[T_{1-\gamma} f_{q+1}] + \cdots + \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_Q] + \mathsf{Inf}_j[f_u] \big) > \tau.$$

STEP 5. *Decoding strategy.* We use the standard strategy – each v_q samples a set $S \subseteq [R]$ according to $\|(f_q)_S\|_2^2$, and chooses a random element from S. u also samples a set $S \subseteq [L]$ according to $\|(f_u)_S\|_2^2$, and chooses a random element from S. As shown in Section 5.3, for each $1 \le j \le L$, the probability that v chooses a label in $\pi^{-1}(j)$ is at least $\gamma \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q]$, and the probability that u chooses j is at least $\gamma \mathsf{Inf}_j[T_{1-\gamma} f_u]$.

Fix q to be the one obtained from Theorem 5.6. The probability that $\pi_q\big(l(v_q)\big) = \pi_{q'}\big(l(v_{q'})\big)$ for some $q < q' \le Q$ or $\pi_q(l(v_q)) = l(u)$ is at least

$$\gamma^2 \sum_{1 \le j \le L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q] \max \left[\max_{q < q' \le Q} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_{q'}], \mathsf{Inf}_j[T_{1-\gamma} f_u] \right]$$

$$\geq \frac{\gamma^2}{Q+1} \sum_{1 \leq j \leq L} \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_q] \left(\overline{\mathsf{Inf}}_j[T_{1-\gamma} f_{q+1}] + \cdots + \overline{\mathsf{Inf}}_j[T_{1-\gamma} f_Q] + \mathsf{Inf}_j[T_{1-\gamma} f_u] \right)$$
$$\geq \frac{\gamma^2 \tau}{Q+1}.$$

If the total fraction of unsatisfied clauses is at most $\epsilon := \frac{1}{4} \cdot \frac{\zeta Q}{8}$, since at least $\frac{1}{2}$ fraction of hyperedges are good, at least $\frac{1}{4}$ fraction of hyperedges are weakly satisfied by the above randomized strategy with probability $\frac{\gamma^2 \tau}{Q+1}$. Setting $\eta := \frac{1}{4} \cdot \frac{\gamma^2 \tau}{Q+1}$ completes the proof of soundness.

Venkatesan Guruswami

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
guruswami@cmu.edu

Euiwoong Lee

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
euiwoonl@cs.cmu.edu