An Analytical Approach to Parallel Repetition via CSP Inverse Theorems

Amey Bhangale* Mark Braverman† Subhash Khot‡
Yang P. Liu§ Dor Minzer¶ Kunal Mittal‖

Abstract

Let $\mathcal{G}$ be a k-player game with value < 1, whose query distribution is such that no marginal on $k-1$ players admits a non-trivial Abelian embedding. We show that for every $n \geqslant N$, the value of the n-fold parallel repetition of $\mathcal{G}$ is

$$\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant \frac{1}{\underbrace{\log\log\cdots\log}_{C \text{ times}} n},$$

where $N = N(\mathcal{G})$ and $1 \leqslant C \leqslant k^{O(k)}$ are constants. As a consequence, we obtain a parallel repetition theorem for all 3-player games whose query distribution is pairwise-connected. Prior to our work, only inverse Ackermann decay bounds were known for such games [Ver96].

As additional special cases, we obtain a unified proof for all known parallel repetition theorems, albeit with weaker bounds:

1. A new analytic proof of parallel repetition for all 2-player games [Raz98, Hol09, DS14].
2. A new proof of parallel repetition for all k-player playerwise connected games [DHVY17, GHM+22].
3. Parallel repetition for all 3-player games (in particular 3-XOR games) whose query distribution has no non-trivial Abelian embedding into $(\mathbb{Z}, +)$ [BKM23c, BBK+25].
4. Parallel repetition for all 3-player games with binary inputs [HR20, GHM+21, GHM+22, GMRZ22].

*Department of Computer Science and Engineering, University of California, Riverside. Supported by the Hellman Fellowship award and NSF CAREER award 2440882.

†Department of Computer Science, Princeton University. Research supported in part by the NSF Alan T. Waterman Award, Grant No. 1933331.

‡Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. Research supported by NSF Award CCF-2515155, and the Simons Investigator Award.

§Department of Computer Science, Carnegie Mellon University.

¶Department of Mathematics, Massachusetts Institute of Technology. Research supported by NSF CCF award 2227876 and NSF CAREER award 2239160.

‖Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. Research supported by NSF Award CCF-2007462, and Simons Investigator Awards to Subhash Khot and Ran Raz.

Contents

1 Introduction **3**
- 1.1 Our Results . . . 4
- 1.2 Special Cases . . . 6
- 1.3 Organization . . . 7

2 Overview **7**
- 2.1 Setup for Parallel Repetition Proofs . . . 8
- 2.2 CSP Inverse Theorems . . . 8
- 2.3 Games with No Abelian Embeddings . . . 9
- 2.4 Games with No Marginal Abelian Embeddings . . . 10

3 Preliminaries **12**
- 3.1 Probability Distributions . . . 12
- 3.2 Noise Operators . . . 13

4 Abelian Embeddings, Inverse Theorems, and Random Restrictions **13**
- 4.1 Abelian Embeddings and Inverse Theorems . . . 13
- 4.2 Random Restrictions . . . 14
- 4.3 Noise Stability under Random Restrictions . . . 15
- 4.4 Coordinate-wise Connected Implies No Abelian Embeddings . . . 16

5 Product Pseudorandomness, Inverse Theorems, and Generalized Random Restrictions **17**
- 5.1 Product Pseudorandomness and Inverse Theorems . . . 17
- 5.2 Generalized Random Restrictions . . . 18
- 5.3 Product Pseudorandomness under Generalized Random Restrictions . . . 20
- 5.4 Conditional Generalized Restrictions . . . 24

6 Multiplayer Games **25**

7 Games with No Abelian Embeddings **26**
- 7.1 Strategy for a Single Copy of the Game . . . 27
- 7.2 Analysis of the First Term . . . 28
- 7.3 Analysis of the Second Term . . . 29
  - 7.3.1 Approximate Independence Under Random Restriction . . . 31
- 7.4 Some Remarks . . . 33

8 Pairwise Connected Games with No Marginal Abelian Embeddings **33**
- 8.1 Pseudorandom Partitions For Each Coordinate . . . 34
- 8.2 Embedding for a Single Copy of the Game . . . 35
  - 8.2.1 Approximate Independence Under Random Restriction . . . 40
- 8.3 Ensuring That Random Restrictions Don't Give Too Much Information . . . 41
- 8.4 Combining Together: Hard Coordinates under Random Restrictions . . . 45
- 8.5 Final Induction . . . 46

A Some Useful Lemmas **49**

B An Inductive Parallel Repetition Criterion **50**

1 Introduction

In a k-player game, a verifier samples a *question* $X = (X^1, \ldots, X^k)$ from a distribution Q over $\mathcal{X} := \mathcal{X}^1 \times \cdots \times \mathcal{X}^k$, where each $\mathcal{X}^i$ is a finite set, and gives X^i to the i^{th} player. Then, for each $i \in \{1, \ldots, k\}$, the i^{th} player answers with $A^i := f^i(X^i)$ for some function $f^i : \mathcal{X}^i \to \mathcal{A}^i$, and sends this to the verifier. The values $A^i \in \mathcal{A}^i$ are known as the *answers*, and let $\mathcal{A} := \mathcal{A}^1 \times \cdots \times \mathcal{A}^k$. The verifier accepts if $V((X^1, \ldots, X^k), (A^1, \ldots, A^k)) = 1$, where $V : \mathcal{X} \times \mathcal{A} \to \{0, 1\}$ is a predicate known to all players. This defines a game $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$, and we define the game's value, denoted $\mathrm{val}(\mathcal{G})$, as the maximum acceptance probability (with respect to the distribution Q) over all possible player strategies. See Definitions 6.1, 6.2 for more precise definitions.

A natural question that arises is: how does the value of a game behave under parallel repetition? The *n-fold parallel repetition* of $\mathcal{G}$, which we denote as $\mathcal{G}^{\otimes n}$, is the game where the verifier samples n questions $(X_1, \ldots, X_n) \sim Q^{\otimes n}$ independently, and sends the i^{th} coordinate of all of $(X_1, \ldots, X_n)$ to the i^{th} player simultaneously; each player then needs to respond n answers, one for each instance of the game. The players win the repeated game if they win each one of the n instances. See Definition 6.4 for a more precise definition.

The parallel repetition of two-player games is by now well understood. Originally proposed by [FRS94] as a means of amplifying the advantage in interactive protocols, it is now known that the value of a two-player game decays exponentially under parallel repetition whenever $\mathrm{val}(\mathcal{G}) < 1$. This exponential decay was first established by Raz [Raz98] using information-theoretic techniques. Subsequent works have simplified Raz's proof and strengthened the quantitative bounds [Hol09, BRR$^+$09, Rao11, RR12, DS14, BG15]. With the exception of [DS14], most of these works follow the information-theoretic framework introduced by Raz. The work of [DS14], in contrast, introduces an analytic approach applicable in the case where the game $\mathcal{G}$ is a projection game (which is the case in the majority of applications, especially those pertaining to probabilistically-checkable-proofs).

Although it might appear plausible that the naïve bound $\mathrm{val}(\mathcal{G}^{\otimes n}) \leq \mathrm{val}(\mathcal{G})^n$ should hold, this turns out to be false [For89, Fei91, FV02, Raz11]. This failure is also well understood: it is intimately connected to the geometry of high-dimensional Euclidean tilings (see [FKO07, KORW08, AK09, BM21]), and in particular to the existence of bodies of volume 1 and surface area $\Theta(\sqrt{n})$ that tile $\mathbb{R}^n$.

Parallel repetition of two-player games has found numerous applications across several domains, including interactive proofs [BGKW88], communication complexity [PRW97, BBCR13, BRWY13], quantum information [CHTW04, BBLV13], and hardness of approximation [FGL$^+$96, ABSS97, ALM$^+$98, AS98, BGS98, Fei98, Has01, Kho02a, Kho02b, GHS02, DGKR05, DRS05]. The reader is referred to the survey [Raz10] for more details.

Parallel repetition of k-player games for $k \geq 3$ is much more poorly understood. Even for $k = 3$, the best general bound on $\mathrm{val}(\mathcal{G}^{\otimes n})$ by [Ver96] is very weak: approximately $(\log^{*q} n)^{-1}$, where $q = |\mathrm{supp}(Q)|$ is the number of questions in $\mathcal{G}$, and $\log^{*q} n$ is defined recursively as the number of times one needs to apply $\log^{*(q-1)}$ starting from n to get down to a constant. This was proven by directly invoking the Density-Hales-Jewett theorem [FK91, Pol12]. More recently, there has been renewed interest in proving parallel repetition theorems for restricted classes of k-player games for $k \geq 3$. [DHVY17] used information-theoretic techniques to prove parallel repetition theorems with exponential decay, i.e., $\mathrm{val}(\mathcal{G}^{\otimes n}) \leq \exp(-\Omega(n))$ for connected games (see Definition 4.7), which are those where the following graph on question tuples is connected: draw an edge between $X, X' \in \mathrm{supp}(Q)$ if X and X' differ in a single coordinate. After this, a series of works [HR20, GHM$^+$21, GHM$^+$22, GMRZ22] established parallel repetition theorems for all 3-player games with binary inputs, i.e., $\mathcal{X}^1 = \mathcal{X}^2 = \mathcal{X}^3 = \{0, 1\}$, with polynomial decay, i.e., $\mathrm{val}(\mathcal{G}) \leq n^{-\Omega(1)}$. These works contained three key technical points (among many others)

that we mention. First, the work [GHM+22] proved polynomial decay for all k-player playerwise connected games, which are games where the projection of the connectivity graph (defined in Definition 4.8) to single coordinates are all connected graphs. Second, this was combined with previous works [HR20, GHM+21] studying a particular GHZ game, whose question distribution is supported on $\{(x_1, x_2, x_3) \in \{0,1\}^3 : x_1 + x_2 + x_3 \equiv 0 \pmod 2\}$, and is not playerwise connected (because the connectivity graph has no edges); more recently, parallel repetition with exponential decay was established for the GHZ game [BKM23c] and was even extended to all 3-XOR games [BBK+25, BBK+25] with a certain assumption on the underlying distribution. Third, the works [GHM+22, GMRZ22], proved a polynomial decay bound for games whose questions are supported on $\{(0,0,1),(0,1,0),(1,0,0)\} \subseteq \{0,1\}^3$; such games are intimately connected to the length-3 density Hales-Jewett problem, bounds for which have been significantly improved very recently [BKLM24c].

Proving improved bounds on the parallel repetition of multiplayer games can lead to several interesting applications. It is known that a strong parallel repetition theorem for a certain class of multiplayer games implies super-linear lower bounds for non-uniform Turing machines [MR21]. Also, parallel repetition of multiplayer games in the large alphabet regime is equivalent to many important problems in high-dimensional extremal combinatorics, like the density Hales-Jewett problem, and that of square free sets in finite fields [FV02, HHR16, Mit25]. Additionally, as mentioned in [DHVY17], it is suspected that the techniques used to prove multiplayer parallel repetition bounds may lead to an improved understanding of multipary communication complexity in the number-on-forehead (NOF) model, a problem intimately connected to circuit lower bounds.

The primary goal of this work is to present a new analytic framework for proving parallel repetition theorems. This framework utilizes recent inverse theorems for k-wise correlations over high-dimensional distributions (e.g., from [BKLM24a, BKLM24b]). While our main theorem requires a certain assumption on the input question distribution, it is sufficiently general and is able to reprove all currently known parallel repetition theorems, and much more. However, we get bounds that are weaker than the best known ones in many cases (a finite number of repeated logarithms), but still superior to the general bounds obtained by Verbitsky [Ver96].

1.1 Our Results

To state our main result, we need to introduce a few notions to describe the assumptions we make on the underlying input distribution. A distribution is pairwise connected if the support of its projection to any two coordinates forms a connected bipartite graph:

Definition 1.1. *(Pairwise-Connected) Let $k \in \mathbb{N}$, let $\Sigma_1, \dots, \Sigma_k$ be finite sets, and let $\mathcal{S} \subseteq \Sigma_1 \times \cdots \times \Sigma_k$. We say that $\mathcal{S}$ is* pairwise-connected *if for every $1 \leqslant i < j \leqslant k$, the bipartite projection graph $\mathcal{S}_{i,j}$ is connected: this has vertex set $\Sigma_i \cup \Sigma_j$, and edge set*

$$\left\{ (x_i, x_j) : \exists y \in \prod_{t \in [k] \setminus \{i,j\}} \Sigma_t,\ (x_i, x_j, y) \in \mathcal{S} \right\} \subseteq \Sigma_i \times \Sigma_j.$$

We say that a distribution μ over $\Sigma_1 \times \cdots \times \Sigma_k$ is pairwise-connected if $\mathrm{supp}(\mu)$ *is pairwise-connected.*

We define the notion of Abelian embeddings, and games/distributions that have no-Abelian-embeddings and no-marginal-Abelian-embeddings.

Definition 1.2. *(Abelian Embeddings) Let $k \in \mathbb{N}$, let $\Sigma_1, \dots, \Sigma_k$ be finite sets, and let $\mathcal{S} \subseteq \Sigma_1 \times \cdots \times \Sigma_k$. An Abelian embedding of $\mathcal{S}$ is a tuple $(G, \sigma_1, \dots, \sigma_k)$, where G is an Abelian group, and*

$\sigma_i : \Sigma_i \to G$, $i \in [k]$ *are mappings such that for each* $(a_1, \ldots, a_k) \in \mathcal{S}$, *it holds that* $\sum_{i=1}^{k} \sigma_i(a_i) = 0_G$. *Such an Abelian embedding is called non-trivial if not all of the maps* σ_i *are constant.*

We say that $\mathcal{S}$ *has no-Abelian-embeddings if it admits no non-trivial Abelian embedding. Similarly, we say that a distribution* μ *over* $\Sigma_1 \times \cdots \times \Sigma_k$ *has no-Abelian-embeddings if* $\mathrm{supp}(\mu)$ *has no-Abelian-embeddings.*

Definition 1.3. *(Marginal Abelian Embeddings) Let* $k \in \mathbb{N}$, *let* $\Sigma_1, \ldots, \Sigma_k$ *be finite sets, and let* μ *be a distribution over* $\Sigma_1 \times \cdots \times \Sigma_k$ *such that:*

1. μ *is a pairwise-connected distribution (see Definition 1.1).*
2. *There exists two distinct indices* $i_1, i_2 \in [k]$, *such that the* $(k-1)$*-marginals* μ_{-i_1} *and* μ_{-i_2} *admit no-Abelian-embeddings (see Definition 1.2).*

Then, we say that μ *is pairwise-connected with no-marginal-Abelian-embeddings.*

Our main theorem is a parallel repetition theorem for distributions with no-marginal-Abelian-embeddings. In the next section, we give several applications.

Theorem 1.4. *Let* $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ *be a* k*-player game with* $\mathrm{val}(\mathcal{G}) < 1$, *such that the distribution* Q *is pairwise-connected with no-marginal-Abelian-embeddings (see Definition 1.3). Then, there exists a constant* $C \in \mathbb{N}$, $C \leqslant k^{O(k)}$, *such that for every sufficiently large* $n \in \mathbb{N}$,

$$\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant \frac{1}{\log \log \cdots \log n},$$

where the number of logarithms is C.

To understand the above definitions and our main theorem better, we make a few simple remarks:

Remark 1. *Let* $\mathcal{S} \subseteq \Sigma_1 \times \cdots \times \Sigma_k$, *and let* μ *be a distribution with support* $\mathcal{S}$. *Then,*

1. *Suppose that* $\mathcal{S}$ *has no-Abelian-embeddings, then for every* $j \in [k]$, *the* projection *of* $\mathcal{S}$ *onto coordinate* j *equals* Σ_j. *Otherwise, a non-trivial Abelian embedding is obtained as follows: the map* σ_j *assigns arbitrary non-zero values to elements of* Σ_j *that never occur in* $\mathcal{S}$, *and everything else maps to zero.*
2. *Suppose that* μ *is not pairwise connected, then it admits a non-trivial Abelian embedding into any (non-trivial) Abelian group* G, *as follows: Suppose for some* $i, j \in [k]$, *there are partitions* $\Sigma_i = \Sigma_i^{(1)} \sqcup \Sigma_i^{(2)}$, $\Sigma_j = \Sigma_j^{(1)} \sqcup \Sigma_j^{(2)}$ *such that the graph* $\mathcal{S}_{i,j}$ *has edges contained in* $\left(\Sigma_i^{(1)} \times \Sigma_j^{(1)}\right) \cup \left(\Sigma_i^{(2)} \times \Sigma_j^{(2)}\right)$. *Then, the map* σ_i *assigns a value* $g \neq 0$ *to each element of* $\Sigma_i^{(1)}$, *and* σ_j *assigns value* $-g$ *to each element of* $\Sigma_j^{(1)}$; *everything else maps to zero.*
3. *Suppose that* μ *is such that some marginal* μ_T, *for* $T \subseteq [k]$, *has a non-trivial Abelian embedding. Then,* μ *also has a non-trivial Abelian embedding. This simply follows by extending the embedding by choosing* σ_j *as the zero map for* $j \notin T$.
4. *The two points above imply that if* μ *has no-Abelian embeddings, then it is pairwise-connected with no-marginal-Abelian-embeddings. Hence, Theorem 1.4 gives bounds for all games with no-Abelian-embeddings; this special case is also proven directly in Section 7.*
5. *Any set* $\mathcal{S}$ *that is connected (Definition 4.7), or even coordinate-wise connected (see Definition 4.8) has no-Abelian-embeddings. This follows from Lemma 4.9.*

6. *With the above, we get that for $k = 3$, Definition 1.3 is equivalent to μ being pairwise-connected. Hence, Theorem 1.4 gives bounds for all 3-player pairwise-connected games.*

7. *Note that the second condition in Definition 1.3 holds if μ_{-i} admits no-Abelian-embeddings for each $i \in [k]$. In this case, by the above observations, it also follows that μ is pairwise-connected, making that condition redundant in the definition.*

   A slightly more careful analysis also shows that pairwise-connectivity follows if three of the $(k-1)$-marginals μ_{-i} admit no-Abelian-embeddings (unlike two, as required in Definition 1.3).

We also show a simple (and naturally arising) game for which our theorem improves upon the state-of-the-art.

Example 1.5. *(Random 3-CNF Game; [GHM+22, Example 1.5]) Consider a random 3-CNF formula $\varphi = (C_1, C_2, \ldots, C_m)$, with m clauses over d variables, generated as follows: Sample each clause independently and uniformly from the set of all $(2d)^3 = 8d^3$ clauses; that is, each clause contains three variables, either negated or not, each chosen randomly.*

For a fixed formula φ, we consider the following 3-player game $\mathcal{G}_\varphi$. The verifier samples $r \in [m]$ uniformly at random, and gives variables corresponding to the literals in C_r to the three players respectively, with each player getting one of the three variables. The players answer back with a value for the variable they get, and the verifier accepts if these values satisfy the clause C_r.

It was shown in [GHM+22] that with high probability:

1. *If $m = \omega(d)$, the value of the game $\mathcal{G}_\varphi$ is close to 7/8, and hence < 1.*

2. *If $m = \omega(d^2 \log d)$, the game $\mathcal{G}_\varphi$ is connected, and its parallel repetition has exponential decay [DHVY17]. On the other hand, if $m = o(d^2)$, the game is not connected.*

3. *If $m = \omega(d^{1.5}\sqrt{\log d})$, the game $\mathcal{G}_\varphi$ is playerwise connected, and its parallel repetition has polynomial decay [GHM+22]. On the other hand, if $m = o(d^{1.5})$, the game is not playerwise connected.*

Our Theorem 1.4 can be used to close the gap above, and give an effective parallel repetition bound when $m = \omega(d \log d)$. This follows by observing that the game $\mathcal{G}_\varphi$ is pairwise connected when $m = \omega(d \log d)$: For any two players, the bipartite projection graph (Definition 1.1) is simply a bipartite graph on $[d] \times [d]$ obtained by choosing m edges uniformly and independently; when $m = \omega(d \log d)$, this is connected with high probability [Pal64].

1.2 Special Cases

In this section, we explain how our parallel repetition theorem (Theorem 1.4) applies in the setting of all previously known parallel repetition theorems with "effective bounds", i.e., not $\log^*$ type.

2-player games and k-player connected games. [DHVY17] studied parallel repetition for connected games—this generalizes the case of two-player games because we can assume without loss of generality that any two-player game is connected when proving parallel repetition theorems. By Lemma 4.9 and Remark 1, it follows that every connected game has no-Abelian-embeddings, and is hence pairwise-connected with no-marginal-Abelian-embeddings. Thus, Theorem 1.4 applies to such games.

Playerwise connected games. These games were studied by [GHM+22] as a step towards proving parallel repetition theorems for all 3-player games with binary inputs. It follows by Lemma 4.9 and Remark 1 that all playerwise connected games have no-Abelian-embeddings, and are hence pairwise-connected with no-marginal-Abelian-embeddings. Thus, Theorem 1.4 applies to such games.

Pairwise-connected 3-player games. [BBK+25], building off [BKM23c], established parallel repetition theorems for 3-XOR games whose question distributions have no non-trivial Abelian embeddings into the integers under addition, i.e., $(\mathbb{Z}, +)$. By Remark 1, all such games are pairwise-connected, and our Theorem 1.4 in fact applies to *all* 3-player games that are pairwise connected.

3-player games with binary inputs. The work [GHM+22] established parallel repetition theorems with polynomial decay for all 3-player games whose question distribution was binary, i.e., $\mathcal{X} = \{0,1\}^3$. Our Theorem 1.4 recovers this result, with weaker bounds. Indeed, by Remark 1, it suffices to argue that we can reduce to the case of pairwise-connected games. Towards this, let $\mathcal{S}$ be the support of Q, and without loss of generality consider the projection of $\mathcal{S}$ onto the first two coordinates. It has between 0 and 4 edges. The cases of 3 and 4 edges lead to connected graphs. 0 edges is impossible because $|\mathcal{S}| \geq 1$. In the case of one edge, both players 1 and 2 know their input deterministically, and thus the number of players can be reduced. In the case of 2 edges, either they share an endpoint or not. If they share an endpoint, again some player knows their input deterministically, and thus the number of players can be reduced. If the edges don't share an endpoint, they must either be $\{(0,0),(1,1)\}$ or $\{(0,1),(1,0)\}$. In the former case, players 1 and 2 have the same input, so they can be merged into a single player. In the latter case, they always have opposite inputs, so they can be merged into a single player again (eg. imagine always negating the input of player 2, so that now they have the same input as player 1). Thus, Theorem 1.4 establishes parallel repetition theorems for all 3-player games over binary alphabets.

1.3 Organization

In Section 2, we give an overview of our proofs. In Section 3, we establish some preliminaries. Then, in Section 4 and Section 5, we state the CSP inverse theorems relevant to this work, define generalized random restrictions, and prove several results regarding these. In Section 6, we formally define the notions of multiplayer games and parallel repetition. Then, in Section 7, we prove parallel repetition for games with no-Abelian-embeddings, an important special case of our main theorem. Finally, in Section 8, we prove our main theorem.

2 Overview

In this section, we give a proof overview for our main result: parallel repetition for all k-player games which are pairwise connected and such that any projection to $k-1$ coordinates have no-Abelian-embeddings. We start by giving the general setup for our parallel repetition proofs. Then, we discuss the main inverse theorems from prior works underlying our proofs; this is the only place where the structure of the support of the query distribution is used. Then, we overview the proof in the simplified case where the distribution itself has no-Abelian-embeddings. Finally, we discuss the additional pieces required to obtain the main result.

2.1 Setup for Parallel Repetition Proofs

The information-theoretic proofs of parallel repetition, such as the ones for two-player games [Raz98, Hol09], at a very high-level, take the following approach. Consider the game $\mathcal{G}^{\otimes n}$, and for each $i \in [n]$, let Win_i be the event that the i^{th} coordinate is won. Observe that for any permutation $i_1, i_2, \ldots, i_n$ of $[n]$, we can write

$$\Pr[\mathrm{Win}_1 \wedge \mathrm{Win}_2 \wedge \cdots \wedge \mathrm{Win}_n] = \prod_{k=1}^{n} \Pr[\mathrm{Win}_{i_k} \mid \mathrm{Win}_{i_1} \wedge \cdots \wedge \mathrm{Win}_{i_{k-1}}].$$

We know that $\Pr[\mathrm{Win}_i] \leq \mathrm{val}(\mathcal{G}) < 1$ for all i, where the probability is over the random questions. Now, let E be the event Win_1. From here, it would be very natural to prove that there is some $i \in \{2, \ldots, n\}$ such that $\Pr[\mathrm{Win}_i \mid E] < 1$ still, and even stronger, that

$$\Pr[\mathrm{Win}_i \mid E] \leq \mathrm{val}(\mathcal{G}) + o_n(1).$$

If this were true, we could condition on Win_i for another coordinate i, and continue inductively. Roughly speaking, the key lemma in these approaches takes the following form:

Claim: For any event E with $\Pr[E] \geq \alpha$, there is some coordinate $i \in [n]$ such that

$$\Pr[\mathrm{Win}_i \mid E] \leq 1 - \Omega_{\mathcal{G}}(1).$$

The rate of decay in the resulting parallel repetition theorem then depends on the smallest α (in terms of n) for which we can establish such a **Claim** (we include a more precise version of this discussion in Lemma B.1). For example, [Raz98, Hol09] establishes such a claim for 2-player games, as long as $\alpha \geq \exp(-\Omega(n))$, which leads to an exponential decay rate. In our theorems, one should think of $\alpha \approx \frac{1}{\log\log\cdots\log n}$, where the number of logarithms is some large constant depending on the game $\mathcal{G}$. The remainder of this overview is devoted to explaining how to establish the **Claim** in our setting, for this choice of α.

2.2 CSP Inverse Theorems

Our results rely on certain theorems about correlations of functions under k-ary n-dimensional distributions. Formally, let $\Sigma_1, \ldots, \Sigma_k$ be finite sets and let μ be a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$. Let $(f_i : \Sigma_i^n \to \mathbb{C})_{i \in [k]}$ be 1-bounded functions (that is, $|f_i(x)| \leq 1$ for all $x \in \Sigma_i^n$). The reader should think of the Σ_i and μ as being fixed while sending $n \to \infty$. A very general question considered by previous works was: under what conditions on μ and f_i can we guarantee that

$$\left|\mathbb{E}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[f_1(x_1) \ldots f_k(x_k)\right]\right| \leqslant o_n(1)? \tag{1}$$

This question, naturally, has applications to analyzing dictatorship tests, property testing, and additive combinatorics. A recent line of work has partially answered this question for increasingly more general distributions μ [BKM25a, BKM23b, BKM23a, BKM24, BKM25b]. For the purposes of this work, we only require the following theorems proven in [BKLM24b].

Projections and pairwise connectivity. Before stating the hypotheses on μ, we introduce some basic notions pertaining to the connectivity of μ. Towards this, given a distribution μ on $\Sigma_1 \times \cdots \times \Sigma_k$ and a subset $S \subseteq [k]$, we can naturally define the *projection* of μ onto S as the distribution on $\prod_{s \in S} \Sigma_s$ where we simply restrict the coordinates of μ onto S. Now, we say that μ is *pairwise connected* if the support of μ_S for any $|S| = 2$ forms a connected graph; see Definition 1.1 for a formal definition.

Distributions with no-Abelian-embeddings. For a formal definition of Abelian embeddings, the reader is referred to Definition 1.2, although it is not strictly necessary for this overview. [BKLM24b] proves that if μ has no-Abelian-embeddings, then (1) holds unless all the functions f_i have non-negligible noise stability, i.e., each of the f_i has non-negligible Fourier mass on low-degree terms. This is formally stated in Theorem 4.1.

Distributions with no-marginal-Abelian-embeddings. We say that a distribution μ has no-marginal-Abelian-embeddings if any projection onto $k-1$ coordinates has no-Abelian-embeddings; note that this is more general than the above definition (see Remark 1).[1] In this context, [BKLM24b] proved that if (1) fails for such μ, then each of the f_i satisfies the following property: there is a product function P_i such that f_iP_i has non-negligible Fourier mass on low-degree terms; here, a product function $P_i : \Sigma_i^n \to \mathbb{C}$ takes the form $P_i(x_1, x_2, \ldots, x_n) = P_{i,1}(x_1) \cdot P_{i,2}(x_2) \cdots P_{i,n}(x_n)$, i.e., it is the product of functions depending only on a single coordinate. Later, we call such functions not *product pseudorandom* (see Definition 5.1). The precise statement of this result is given in Theorem 5.2.

2.3 Games with No Abelian Embeddings

In this section, we overview a simplified version of Theorem 1.4, for games whose query distribution Q has no-Abelian-embeddings; this appears as Theorem 7.1. Recall from the discussion in Section 2.1 that our goal is to prove that given some event E with $\Pr[E] \geq \alpha$ for α not too small, there is some coordinate $i \in [n]$ with $\Pr[\text{Win}_i \mid E] \leq \text{val}(\mathcal{G}) + o(1)$. It turns out that we can assume that E is a product event $E = E^1 \times \cdots \times E^k$, i.e., the product of events E^j that depend only on player j.[2]

A common strategy used in prior works to establish such a claim is to use an *embedding strategy*: prove that the players can obtain value close to $\Pr[\text{Win}_i \mid E]$ on using some honest strategy on a single copy of the game. Then, this quantity must then be bounded by $\text{val}(\mathcal{G})$, as desired. Here, we shall consider the following randomized strategy for a single coordinate of the game $\mathcal{G}$. Let $\tilde{X} \sim Q$ be the query in a single copy of the game $\mathcal{G}$; the players think of this as the input in coordinate i of the game $\mathcal{G}^{\otimes n}$, and do the following: each player $j \in [k]$ randomly fills out the remaining questions for coordinates $i' \neq i$ conditioned on event E^j, and then outputs based off of their strategy for $\mathcal{G}^{\otimes n}$ in coordinate i.

We want to show that the success probability of the above strategy is close to $\Pr[\text{Win}_i \mid E]$. For this, we first try to analytically express the probability $\Pr[\text{Win}_i \mid E]$. Towards this, we define functions $F^{j,\tilde{X}^j} : (\mathcal{X}^j)^{\otimes(n-1)} \to \{0,1\}$ as follows. To define $F^{j,\tilde{X}^j}(X)$, think of X as being assignments to the $n-1$ coordinates of player j's input other than i. Then, $F^{j,\tilde{X}^j}(X) = 1$ if filling in the i^{th} coordinate with $\tilde{X}^j$ gives an input in E^j, and otherwise $F^{j,\tilde{X}^j}(X) = 0$. Also, for answers $\tilde{A} \in \mathcal{A}$, define the functions $F^{j,\tilde{X}^j,\tilde{A}^j} : (\mathcal{X}^j)^{\otimes(n-1)} \to \{0,1\}$ as follows: $F^{j,\tilde{X}^j,\tilde{A}^j}(X) = 1$ if any only if $F^{j,\tilde{X}^j}(X) = 1$ and when X is filled in with $\tilde{X}^j$ in coordinate i, player j's strategy outputs $\tilde{A}^j$. Note that these are precisely the functions defined in Definition 7.8 and Definition 8.1. Then, the quantity $\Pr[\text{Win}_i \mid E]$, which is winning probability of coordinate i when conditioned on E,

[1] Everything we say going forward will work with the slightly more general Definition 1.3 as well.

[2] In Section 2.1, E was the event of winning a few coordinates; roughly speaking, we can choose to work with the event E which fixes some winning questions and answers in these coordinates, and this is a product event.

roughly[3] equals

$$\mathop{\mathbb{E}}_{\tilde{X}\sim Q}\sum_{\tilde{A}:V(\tilde{X},\tilde{A})=1}\left[\frac{\mathbb{E}_{(X^1,\ldots,X^k)\sim Q^{\otimes(n-1)}}\left[F^{1,\tilde{X}^1,\tilde{A}^1}(X^1)\cdot F^{2,\tilde{X}^2,\tilde{A}^2}(X^2)\cdots F^{k,\tilde{X}^k,\tilde{A}^k}(X^k)\right]}{\mathbb{E}_{(X^1,\ldots,X^k)\sim Q^{\otimes(n-1)}}\left[F^{1,\tilde{X}^1}(X^1)\cdot F^{2,\tilde{X}^2}(X^2)\cdots F^{k,\tilde{X}^k}(X^k)\right]}\right]. \tag{2}$$

Observe that for fixed $\tilde{X}$ and $\tilde{A}$, the expressions in the numerator and denominator are all k-wise correlations over $Q^{\otimes(n-1)}$, where Q by assumption is a distribution with no-Abelian-embeddings. Thus, by our above discussion (also see Theorem 4.1), the expressions in the numerator and denominator behave as if the k players are acting independently/honestly (that is, according to the honest strategy described above), as long as all the functions $F^{j,\tilde{X}^j}$ and $F^{j,\tilde{X}^j,\tilde{A}^j}$ have low noise stability, i.e., almost all their mass is on high-degree Fourier coefficients (except the constant term). Hence, if this high-degree property holds for all the functions, we are done.

Thus, we have reduced the problem to ensuring that each $F^{j,\tilde{X}^j}$ and $F^{j,\tilde{X}^j,\tilde{A}^j}$ has low noise stability; note that the number of such functions is a constant (depending on $\mathcal{G}$). To achieve this, we apply random restrictions to the overall game. Precisely, Lemma 4.6 states that taking a random restriction with a randomly chosen probability makes any small set of functions high-degree, with high probability. Thus, in our proof, we first apply such a random restriction to the game (note that a random restriction of a game is still a valid game with many coordinates) to make the relevant functions $F^{j,\tilde{X}^j}$ and $F^{j,\tilde{X}^j,\tilde{A}^j}$ have low noise stability, and then use the inverse theorem to argue that after this the players are acting as if they have no communication, and thus $\Pr[\mathrm{Win}_i \mid E] \le \mathrm{val}(\mathcal{G}) + o(1)$ as desired.

As a technical step, we need to prove that the questions to the players in coordinate i have distribution essentially the same as Q, even when conditioning on E and a typical random restriction as above. Very roughly speaking, this holds because the random restriction leaves a large number of coordinates free. We carry out this step by an information-theoretic argument, as in [Raz98, Hol09]; see Lemma 7.6 for formal details. We note however that the use of information theory in this step is only for simplicity, and not necessary at all—in the proof of our actual main theorem, this step becomes more far more complicated, and we use an analytic argument with no information theory.

2.4 Games with No Marginal Abelian Embeddings

We now describe the modifications needed to extend our the discussion in the above section to our main theorem (Theorem 1.4) about games with no-marginal-Abelian-embeddings. All the discussion up through (2) proceeds identically. In the case of no-Abelian-embeddings, Theorem 4.1 said that the expressions in the numerator and denominator of (2) behave as if the k players are independent as long as the function $F^{j,\tilde{X}^j}$ and $F^{j,\tilde{X}^j,\tilde{A}^j}$ have low noise stability. However, in the more general setting, this is no longer true and what we require instead is that the functions are *product pseudorandom* (see Definition 5.1), i.e., the product of $F^{j,\tilde{X}^j}$ and any product-function has low noise stability. This is much more difficult to ensure and leads to technical complications.

Fortunately, previous work [BKLM24a, BKLM24c] introduced a type of random restriction that can handle functions that are not product pseudorandom. In this work, we call this a *generalized random restriction* (see Definition 5.7), which takes the following form: some coordinates are randomly restricted, and the remaining coordinates are partitioned into sets $T_1, T_2, \ldots, T_m \subseteq [n]$. Now, we force that all coordinates in a set T_i take the same value. Note that this naturally takes an n-dimensional function and turns it into an m-dimensional function. Furthermore, the restriction

[3] For this to be exact, $\tilde{X}$ must also be drawn conditioned on E.

and these sets $T_1, \ldots, T_m$ are randomly chosen so as to ensure that the overall distribution on $Q^{\otimes n}$ is approximately preserved.

Ultimately, we prove that if the functions $F^{j,\tilde{X}^j}$ or $F^{j,\tilde{X}^j,\tilde{A}^j}$ are not product pseudorandom, then we can apply such a generalized random restriction to increase some ℓ_2 energy potential function, which cannot happen forever. In this way, we ensure that the relevant functions are all product pseudorandom eventually.[4] With this, we carry out a proof that very roughly resembles the no-Abelian-embeddings case, however, it turns out that making the argument actually work is much more technical, for reasons described next.

Generalized Random Restrictions. Working with generalized random restrictions is far more technically challenging than usual random restrictions. They only preserve the original distribution approximately, and hence extra care is needed to analyze the error terms for each of the steps in the proof.

Moreover, it turns out that in this case, to be able to describe the honest strategy for one copy of the game, we need to be able to sample generalized random restrictions conditioned on the event E. A priori it is not even clear what this means, however the natural definition, via Bayes' rule, works for us. This leads to our notion of conditional generalized random restrictions (see Definition 5.11).

Making sure random restrictions do not reveal too much information. A major technical hurdle that arises with the use of generalized random restrictions is that the restrictions needed depend on the set of functions we wish to make product pseudorandom.

As mentioned in the no-Abelian-embeddings case, for technical reasons, we still need to prove that the questions to the players in coordinate i—the coordinate we are embedding into—have distribution essentially the same as Q, even when conditioning on E and a typical generalized random restriction. Since the generalized random restriction here may depend both on the coordinate i being considered and the event E, this may not even be true.[5] We get around this hurdle by observing that whenever this is false, it must be because the generalized random restriction, along with the question in coordinate i, *split* the mass of E unevenly. Hence, if we perform the generalized random restriction and fix question in coordinate i randomly, we get an ℓ_2 increment with respect to the mass of the set E, which cannot happen forever; see Section 8.3 for formal details of this step.

Hard coordinates after a generalized random restriction. Due to the aforementioned step, we don't end up proving that $\Pr[\mathrm{Win}_i \mid E]$ itself is small; instead, we only prove that after performing a generalized random restriction, we get many coordinates that are hard to win. Finally obtaining a parallel repetition bound from this statement requires carrying out a very careful induction,

[4] For some intuition, we demonstrate a simple example where generalized restrictions are useful to make functions product-pseudorandom. Suppose $f : \{0,1\}^n \to \{-1,1\}$ is a character over $\mathbb{F}_2$, given by $f(x) = (-1)^{\sum_{i \in S} x_i}$, for some $S \subseteq [n]$. If $S = [n]$, any usual random restriction of f, down to any number of coordinates, is still a character (possibly with a -1 sign), and hence not product-pseudorandom. On the other hand, there always exists a generalized random restriction, down to $m = \lfloor n^{1/3} \rfloor$ coordinates, that makes f product-pseudorandom: If $|S| \leqslant n/2$, we do a usual random restriction and fix the input x_i randomly for each $i \in S$; the function after the restriction is a constant. If $|S| > n/2$, we randomly pick pairs of coordinates $T_1 = \{i_1, j_1\}, T_2 = \{i_2, j_2\}, \ldots, T_m = \{i_m, j_m\}$ inside S, force both coordinates inside each T_i to take the same value, and randomly restrict the coordinates in $S \setminus \cup_{i \in [m]} T_i$; after any such restriction, the function becomes a constant, since coordinates inside each T_i *cancel out*. Moreover, for $m = \lfloor n^{1/3} \rfloor$, as T_i's were chosen randomly, the overall distribution is approximately preserved on average.

[5] For example, suppose the inputs to the players are binary, and that E says that the inputs in the first $n/2$ coordinates sum to zero (mod 2), and also that the inputs in the second $n/2$ coordinates sum to zero (mod 2). Now, for each coordinate i, the corresponding generalized random restriction may randomly fix the inputs to all coordinates except i in the block of size $n/2$ that i lies in.

where we alternate between steps of finding a hard coordinate and performing a generalized random restriction; see Section 8.5 for formal details of this step.

3 Preliminaries

Let $\mathbb{N} = \{1, 2, \ldots\}$ denote the set of natural numbers. For $n \in \mathbb{N}$, we use $[n]$ to denote the set $\{1, 2, \ldots, n\}$. For any set $S \subseteq [n]$, we use $\overline{S}$ to denote its complement $[n] \setminus S$.

We use $\log(\cdot)$ to denote the function $\log_2(\cdot)$, and $\exp(\cdot)$ to denote the function $2^{(\cdot)}$.

3.1 Probability Distributions

We will use calligraphic letters to denote sets, capital letters to denote random variables and small letters to denote values.

Let P be a distribution (over an underlying *finite* set Ω, which is usually clear from context). We use $\mathrm{supp}(P) = \{\omega \in \Omega : P[\omega] > 0\}$ to denote the *support* of the distribution P. For a random variable X, we use P_X to denote the distribution of X, that is, $P_X[x] = P[X = x]$. For random variables X and Y, we use P_{XY} to denote the joint distribution of X and Y.

For an event E with $P[E] > 0$, we use $P|E$ to denote the conditional probability distribution P conditioned on E. Similarly, we use $(P|E)_X = P_{X|E}$ to denote the distribution of X conditioned on the event E, given by

$$P_{X|E}[x] = \frac{P\left[X = x \wedge E\right]}{P[E]}.$$

Suppose R is a random variable, and r is such that $P_R[r] > 0$. We will frequently use the shorthand $P_{X|r}$ to denote the distribution $P_{X|R=r}$.

Let P_X and Q_X be distributions over set $\mathcal{X}$. The ℓ_1-distance between P_X and Q_X is defined as

$$\left\|P_X - Q_X\right\|_1 = \sum_{x \in \mathcal{X}} \left|P_X[x] - Q_X[x]\right|.$$

Next, we state a useful lemma that was used in previous works on two-player parallel repetition. Informally, it says that if we condition a product distribution P on some event E with nonnegligible probability, then the marginals of P and its conditioning are very close on average.

Lemma 3.1. *[Raz98, Hol09] Let $V = (V_1, \ldots, V_n)$ be a random variable, and let F be an event in some finite probability space (Ω, P). Suppose that $P_V = P_{V_1} \times \cdots \times P_{V_n}$ is a product distribution. Then, we have*

$$\frac{1}{n} \sum_{i=1}^{n} \left\|P_{V_i|F} - P_{V_i}\right\|_1 \leqslant \sqrt{\frac{2}{n} \log_2 \frac{1}{P[F]}}.$$

Proof Sketch. It holds that

$$D\left(P_{V|F} || P_V\right) = \sum_v P[V = v|F] \cdot \log_2 \left(\frac{P[V = v|F]}{P[V = v]}\right) \leqslant \log_2 \left(\frac{1}{P[F]}\right),$$

where we use D to denote the relative entropy (or the KL divergence). Also, we have

$$D\left(P_{V|F} || P_V\right) \geqslant \sum_{i=1}^{n} D\left(P_{V_i|F} || P_{V_i}\right) \geqslant \frac{1}{2} \cdot \sum_{i=1}^{n} \left\|P_{V_i|F} - P_{V_i}\right\|_1^2 \geqslant \frac{1}{2n} \cdot \left(\sum_{i=1}^{n} \left\|P_{V_i|F} - P_{V_i}\right\|_1\right)^2.$$

The first inequality above holds as relative entropy is super-additive when the second distribution is a product distribution; the second inequality follows from Pinsker's inequality. □

3.2 Noise Operators

Let (Σ, ν) be a probability space, with Σ a finite set. Define an inner product on the space $L^2(\Sigma, \nu)$ as follows: for $f, g : \Sigma \to \mathbb{C}$, define

$$\langle f, g \rangle_\nu = \mathop{\mathbb{E}}_{x \sim \nu} \left[f(x) \overline{g(x)} \right].$$

Definition 3.2. *(Noise Distribution) For a parameter $\rho \in [0, 1]$, and $y \in \Sigma$, the distribution of inputs ρ-correlated to y, denoted $y' \sim \mathrm{T}_\rho y$, is defined as follows: take $y' = y$ with probability ρ, and else sample $y' \sim \nu$.*

We also view T_ρ as a map on the space of functions $L^2(\Sigma, \nu)$, given by

$$(\mathrm{T}_\rho g)(y) = \mathop{\mathbb{E}}_{y' \sim \mathrm{T}_\rho y} \left[g(y') \right].$$

We consider the tensorization $\mathrm{T}_\rho^{\otimes n}$ of this operator, which acts on $L^2(\Sigma^n, \nu^{\otimes n})$ by applying T_ρ independently across the n coordinates; for ease of notation, we drop the $\otimes n$ superscript and simply call this operator T_ρ as well.

Definition 3.3. *(Noise Stability) For a function $g : \Sigma^n \to \mathbb{C}$, define its ρ-noise-stability as*

$$\mathrm{Stab}_\rho^\nu[g] = \langle g, \mathrm{T}_\rho g \rangle_{\nu^{\otimes n}}.$$

This satisfies $\mathrm{Stab}_\rho^\nu[g] = \left\| \mathrm{T}_{\sqrt{\rho}} g \right\|_2^2$; in particular, $\mathrm{Stab}_\rho^\nu[g]$ is always a non-negative real number. We often drop the superscript ν when it is clear from context.

Definition 3.4. *(Expectation) For a function $f \in L^2(\Sigma^n, \nu^{\otimes n})$, we use $\nu(f)$ to denote its expectation $\mathbb{E}_{x \sim \nu^{\otimes n}}[f(x)]$.*

4 Abelian Embeddings, Inverse Theorems, and Random Restrictions

4.1 Abelian Embeddings and Inverse Theorems

In this section we state a result of Bhangale, Khot, Liu and Minzer [BKLM24b] which says that products of high-degree[6] functions behave pseudorandomly under distributions with no-Abelian-embeddings:

Theorem 4.1. *(Inverse Theorem under No-Abelian-Embeddings; [BKLM24b, Theorem 1])*

Let $k \in \mathbb{N}$ and let $\Sigma_1, \ldots, \Sigma_k$ be finite sets. Let μ be a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$ with no-Abelian-embeddings (see Definition 1.2). Then, for every sufficiently large $n \in \mathbb{N}$ and every $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$, such that the following holds:

If 1-bounded functions $f_i : \Sigma_i^n \to \mathbb{C}$, $i \in [k]$, satisfy

$$\left| \mathop{\mathbb{E}}_{(x_1, \ldots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^{k} f_i(x_i) \right] \right| \geqslant \epsilon,$$

[6] We note that the condition $\mathrm{Stab}_{1-\delta}[f] \leqslant \delta$ in the theorem serves as a convenient proxy for f to have high Fourier degree. The reader is referred to [BKLM24b] for more details.

then $\mathrm{Stab}^{\mu_i}_{1-\delta}[f_i] \geqslant \delta$ *for each* $i \in [k]$*, where* μ_i *is the marginal of* μ *on the* i^{th} *coordinate.*

Quantitatively, when $\epsilon \leqslant o_n(1)$*, we can take* $\delta = \exp(-\exp(\cdots \exp(\epsilon^{-1})))$*, where the number of exponentials is at most* $k^{O(k)}$*.*[7]

Corollary 4.2. *Let* $k \in \mathbb{N}$ *and let* $\Sigma_1, \ldots, \Sigma_k$ *be finite sets. Let* μ *be a distribution over* $\Sigma_1 \times \cdots \times \Sigma_k$ *with no-Abelian-embeddings (see Definition 1.2). Then, for every sufficiently large* $n \in \mathbb{N}$ *and every* $\delta > 0$*, there exists* $\epsilon = \epsilon(\delta) > 0$*, such that the following holds:*

Let $f_i : \Sigma_i^n \to [0,1]$*,* $i \in [k]$ *be functions, and for each* i*, let* μ_i *be the marginal of* μ *on the* i^{th} *coordinate. If* $\mathrm{Stab}^{\mu_i}_{1-\delta}[f_i - \mu_i(f_i)] < \delta$ *for each* $i \in [k]$*, it holds that*

$$\left| \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[\prod_{i=1}^{k} f_i(x_i) \right] - \prod_{i=1}^{k} \mu_i(f_i) \right| \leqslant \epsilon.$$

Quantitatively, when $\delta \leqslant o_n(1)$*, we can take* $\epsilon = \frac{1}{\log\log\cdots\log\left(\frac{1}{\delta}\right)}$*, where the number of logarithms at most* $k^{O(k)}$*.*

Proof. Let $\delta > 0$, and let $\epsilon > 0$ be such that $(\frac{\epsilon}{k}, \delta)$ satisfy Theorem 4.1. Then, we have

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[\prod_{i=1}^{k} f_i(x_i) \right] - \prod_{i=1}^{k} \mu_i(f_i) \right| \\ &\qquad = \left| \sum_{t=1}^{k} \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[\prod_{i=1}^{t-1} \mu_i(f_i) \cdot (f_t(x_t) - \mu_t(f_t)) \cdot \prod_{i=t+1}^{k} f_i(x_i) \right] \right| \leqslant k \cdot \frac{\epsilon}{k} = \epsilon, \end{aligned}$$

where we first used the triangle inequality and then used Theorem 4.1 for each term. Note that $f_t(x_t) - \mu_t(f_t)$ is 1-bounded for each $t \in [k]$.

The bound on ϵ is now $\frac{k}{\log\log\cdots\log\left(\frac{1}{\delta}\right)}$, and for $\delta \leqslant o_n(1)$, this is as desired (by possibly increasing the number of logarithms by 1). □

The above corollary is very useful in calculating expectations containing high-degree functions; however, in most applications, the functions we care about are not high-degree. Next, we introduce the notion of random restrictions, which helps make arbitrary functions high-degree.

4.2 Random Restrictions

In this section we define the notions of a *restriction* and *random restriction*. A restriction takes a subset of coordinates $I \subseteq [n]$ and sets them to some fixed values.

Definition 4.3. *(Restriction) Let* Σ *be a finite alphabet and let* $n \in \mathbb{N}$*. A restriction on* Σ^n *is a tuple* $\rho = (I, z)$ *where* $I \subseteq [n]$ *and* $z \in \Sigma^I$*.*

For a function $f : \Sigma^n \to \mathbb{C}$ *and a restriction* $\rho = (I, z)$ *we define the restricted function* $f_{I\to z} : \Sigma^{\bar{I}} \to \mathbb{C}$ *as* $f_{I\to z}(y) = f(y, z)$*; we shall also use* f_ρ *to denote this function.*

The notation $I \to z$ above indicates that the variables in the set I are fixed to value z. A random restriction is a restriction where the subset of coordinates I is chosen randomly and the values they are fixed to are chosen from the base distribution.

[7] The quantitative dependence in [BKLM24b] is $\delta = \exp(-\exp(\cdots \exp(\epsilon^{-O_\alpha(1)})))$, where the number of exponentials is $C \leqslant k^{O(k)}$, and $\alpha \in (0,1]$ is such that $\min_{x\in\mathrm{supp}(\mu)} \mu[x] \geqslant \alpha$. For our purposes, α is a constant independent of n, and hence when $\epsilon \leqslant o_n(1)$, the choice $\delta = \exp(-\exp(\cdots \exp(\epsilon^{-1})))$ works, where the number of exponentials is $C + 1 \leqslant k^{O(k)}$.

Definition 4.4. *(p-random restriction) Let $p \in [0,1]$, $n \in \mathbb{N}$, and let (Σ, ν) be a probability space.*

We use $I \sim_p [n]$ to denote a random set $I \subseteq [n]$, where each $i \in [n]$ is included in I with probability p independently.

A p-random restriction on $(\Sigma^n, \nu^{\otimes n})$ is a random tuple $\rho = (I, z)$, where $I \sim_{1-p} [n]$, and $z \sim \nu^{\otimes I}$. For a function $f \in L^2(\Sigma^n, \nu^{\otimes n})$, its p-random restriction is then the function $f_{I \to z} : \Sigma^{\bar{I}} \to \mathbb{C}$. We shall also sometimes think $f_{I \to z} : \Sigma^n \to \mathbb{C}$ by ignoring the input coordinates in the set I.

4.3 Noise Stability under Random Restrictions

Next, we prove a simple lemma about the average noise stability of functions under random restrictions.

Lemma 4.5. *(Noise Stability under Random Restrictions) Let $f \in L^2(\Sigma^n, \nu^{\otimes n})$. Then, for $p, \delta \in [0,1]$, we have*

$$\mathop{\mathbb{E}}_{I \sim_{1-p}[n],\ z \sim \nu^{\otimes I}} \left[\mathrm{Stab}_{1-\delta}\left[f_{I \to z} - \nu(f_{I \to z})\right]\right] = \mathrm{Stab}_{1-p\delta}[f] - \mathrm{Stab}_{1-p}[f].$$

Proof. For $\rho = 1 - \delta$, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \left[\mathrm{Stab}_{1-\delta}\left[f_{I \to z} - \nu(f_{I \to z})\right]\right] \\
&\quad = \mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \left\langle f_{I \to z} - \nu(f_{I \to z}),\ \mathrm{T}_\rho^{\otimes \bar{I}} \left(f_{I \to z} - \nu(f_{I \to z})\right)\right\rangle_{\nu^{\otimes \bar{I}}} \\
&\quad = \mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \left[\left\langle f_{I \to z},\ \mathrm{T}_\rho^{\otimes \bar{I}} f_{I \to z}\right\rangle_{\nu^{\otimes \bar{I}}} - \left|\nu(f_{I \to z})\right|^2\right] \\
&\quad = \mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \left[\left\langle f_{I \to z},\ \mathrm{T}_\rho^{\otimes \bar{I}} f_{I \to z}\right\rangle_{\nu^{\otimes \bar{I}}}\right] - \mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \left[\left|\nu(f_{I \to z})\right|^2\right].
\end{aligned}$$

Note that the second term in the above expression is the same as the first term under the substitution $\rho = 0$; hence, it suffices to show that the first term equals $\mathrm{Stab}_{1-p(1-\rho)}[f]$. This is done as follows:

$$\begin{aligned}
\mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \left[\left\langle f_{I \to z},\ \mathrm{T}_\rho^{\otimes \bar{I}} f_{I \to z}\right\rangle_{\nu^{\otimes \bar{I}}}\right] &= \mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \mathop{\mathbb{E}}_{z \sim \nu^{\otimes I}} \mathop{\mathbb{E}}_{x \sim \nu^{\otimes \bar{I}}} \mathop{\mathbb{E}}_{y \sim \mathrm{T}_\rho^{\otimes \bar{I}} x} f(x,z) \cdot \overline{f(y,z)} \\
&= \mathop{\mathbb{E}}_{I \sim_{1-p}[n]} \left\langle f,\ \mathrm{T}_\rho^{\otimes \bar{I}} \mathrm{T}_1^{\otimes I} f\right\rangle_{\nu^{\otimes n}} \\
&= \left\langle f,\ \left(p \cdot \mathrm{T}_\rho + (1-p) \cdot \mathrm{T}_1\right)^{\otimes n} f\right\rangle_{\nu^{\otimes n}} \\
&= \left\langle f,\ \mathrm{T}_{p\rho+1-p}^{\otimes n}\ f\right\rangle_{\nu^{\otimes n}} \\
&= \mathrm{Stab}_{1-p(1-\rho)}[f].
\end{aligned}$$

□

With the above, we show that random restrictions of arbitrary functions are essentially high-degree, when the restriction parameter is chosen appropriately.

Lemma 4.6. *Let $f \in L^2(\Sigma^n, \nu^{\otimes n})$, and let $\delta \in (0,1)$, $T \in \mathbb{N}$. Let $p \in [0,1]$ be chosen uniformly at random from the set $\left\{1, \delta, \delta^2, \ldots, \delta^{T-1}\right\}$, and let $I \sim_{1-p} [n]$, $z \sim \nu^{\otimes I}$. Then, for every $\eta > 0$, we have*

$$\Pr\left[\mathrm{Stab}_{1-\delta}\left[f_{I \to z} - \nu(f_{I \to z})\right] \geqslant \eta \cdot \mathrm{Var}[f]\right] \leqslant \frac{1}{\eta T}.$$

Proof. By Lemma 4.5, we have

$$\begin{aligned}\mathop{\mathbb{E}}_{p,I,z}\left[\mathrm{Stab}_{1-\delta}\left[f_{I\to z}-\nu(f_{I\to z})\right]\right] &= \mathop{\mathbb{E}}_{p}\left[\mathrm{Stab}_{1-p\delta}[f]-\mathrm{Stab}_{1-p}[f]\right]\\ &= \frac{1}{T}\sum_{t=0}^{T-1}\left(\mathrm{Stab}_{1-\delta^{t+1}}[f]-\mathrm{Stab}_{1-\delta^{t}}[f]\right)\\ &= \frac{1}{T}\left(\mathrm{Stab}_{1-\delta^{T}}[f]-\mathrm{Stab}_{0}[f]\right)\\ &\leqslant \frac{1}{T}\left(\mathrm{Stab}_{1}[f]-\mathrm{Stab}_{0}[f]\right)=\frac{\mathrm{Var}[f]}{T}.\end{aligned}$$

The result follows by Markov's inequality. □

4.4 Coordinate-wise Connected Implies No Abelian Embeddings

In this subsection, we show that any *sufficiently connected* set $\mathcal{S}$ has no-Abelian-embeddings. We start by defining connected and coordinate-wise connected subsets $\mathcal{S}\subseteq\Sigma_1\times\cdots\times\Sigma_k$, which are notions that have been studied in previous works [DHVY17, GHM+22].

Definition 4.7. *(Connectivity Sets) Let $k\in\mathbb{N}$, let $\Sigma_1,\ldots,\Sigma_k$ be finite sets, and let $\mathcal{S}\subseteq\Sigma_1\times\cdots\times\Sigma_k$. Define the (simple, undirected) connection graph $\mathcal{H}(\mathcal{S})$ as follows: The vertex set is $\mathcal{S}$, and there is an edge between $x,y\in\mathcal{S}$ if and only if they differ in exactly one coordinate; that is, there exists $j\in[k]$ such that $x_{-j}=y_{-j}$ and $x_j\neq y_j$.*[8]

We say that $\mathcal{S}$ is connected *if the graph $\mathcal{H}(\mathcal{S})$ is connected.*

Next, we define an even weaker[9] notion of connectivity, where we only require the projection of the above graph with respect to each of the k coordinates to be connected.

Definition 4.8. *(Coordinate-wise Connected Sets) Let $k\in\mathbb{N}$, let $\Sigma_1,\ldots,\Sigma_k$ be finite sets, and let $\mathcal{S}\subseteq\Sigma_1\times\cdots\times\Sigma_k$. For $j\in[k]$, define the (simple, undirected) connection graph $\mathcal{H}_j(\mathcal{S})$ as follows: The vertex set is Σ_j, and there is an edge between $x_j,x_j'\in\Sigma_j$ if and only if there exists $x_{-j}\in\Sigma_{-j}$ such that both $(x_j,x_{-j})\in\mathcal{S}$ and $(x_j',x_{-j})\in\mathcal{S}$.*

We say that $\mathcal{S}$ is connected with respect to coordinate $j\in[k]$ if $\mathcal{H}_j(\mathcal{S})$ is connected. We say that $\mathcal{S}$ is coordinate-wise connected *if $\mathcal{S}$ is connected with respect to each $j\in[k]$.*

We show that coordinate-wise connected sets have no-Abelian-embeddings:

Lemma 4.9. *Let $k\in\mathbb{N}$, let $\Sigma_1,\ldots,\Sigma_k$ be finite sets, and let $\mathcal{S}\subseteq\Sigma_1\times\cdots\times\Sigma_k$. Let $(G,\sigma_1,\ldots,\sigma_k)$ be an Abelian embedding of $\mathcal{S}$, and let $j\in[k]$ be such that $\mathcal{S}$ is connected with respect to coordinate $j\in[k]$. Then, the map σ_j is constant.*

In particular, if $\mathcal{S}$ is coordinate-wise connected, then it has no-Abelian-embeddings.

Proof. Let $(G,\sigma_1,\ldots,\sigma_k)$ be an Abelian embedding of $\mathcal{S}$, and let $j\in[k]$ be such that $\mathcal{S}$ is connected with respect to coordinate $j\in[k]$.

Consider any edge $\left\{x_j,x_j'\right\}$ in the graph $\mathcal{H}_j(\mathcal{S})$; there exists $x_{-j}\in\Sigma_{-j}$ such that $x=(x_j,x_{-j})\in\mathcal{S}$ and $x'=(x_j',x_{-j})\in\mathcal{S}$. By the definition of Abelian embeddings, we have

$$\sum_{i=1}^{k}\sigma_i(x_i)=0_G=\sum_{i=1}^{k}\sigma_i(x_i'),$$

[8] By x_{-j} we mean $(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_k)\in\Sigma_{-j}=\prod_{i\in[k],i\neq j}\Sigma_i$.

[9] Observe that any set $\mathcal{S}$ that is connected must also be coordinate-wise connected, while the converse is not necessarily true.

and hence $\sigma_j(x_j) = \sigma_j(x'_j)$. The lemma now follows by observing that $\mathcal{H}_j(\mathcal{S})$ is connected. □

5 Product Pseudorandomness, Inverse Theorems, and Generalized Random Restrictions

5.1 Product Pseudorandomness and Inverse Theorems

We define the crucial notion of a product pseudorandom function. We say that a function is product-pseudorandom if its random restriction down to n' coordinates has nontrivial correlation to a product function with small probability.

Definition 5.1. *((n', γ)-product pseudorandomess) Let (Σ, μ) be a probability space, $n \in \mathbb{N}$. For $n' \leqslant n$ and $\gamma > 0$, we say that a function $f : \Sigma^n \to \mathbb{C}$ is (n', γ)-product pseudorandom if for any $\delta \in \left[\frac{n'}{n}, 1\right]$, the probability that a random restriction of f down to $\bar{I} \sim_\delta [n]$ is γ-correlated to a product function is less than γ. Precisely,*

$$\Pr_{I \sim_{1-\delta}[n], z \sim \mu^{\otimes I}} \left[\exists \left\{P_i : \Sigma \to \mathbb{C},\ \|P_i\|_\infty \leqslant 1\right\}_{i \in \bar{I}} \text{ with } \left| \mathop{\mathbb{E}}_{x \sim \mu^{\otimes \bar{I}}} \left[f_{I \to z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geqslant \gamma \right] < \gamma.$$

A key input to our analysis is an inverse theorem of [BKLM24a, BKLM24b] which analyzes the correlations of product-pseudorandom functions over certain pairwise-connected distributions.

Theorem 5.2. *(Inverse Theorem; [BKLM24b, Lemma 1.4])*
Let $k \in \mathbb{N}$, let $\Sigma_1, \ldots, \Sigma_k$ be finite sets, and let μ be a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$ such that:

1. *μ is a pairwise-connected distribution (see Definition 1.1).*
2. *The the marginal of μ on coordinates $\{1, 2, \ldots, k-1\}$, denoted μ_{-k}, admits no-Abelian-embeddings (see Definition 1.2).*

Then, for every sufficiently large $n \in \mathbb{N}$ and every $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$, such that the following holds: If 1-bounded functions $f_i : \Sigma_i^n \to \mathbb{C}$, $i \in [k]$, satisfy

$$\left| \mathop{\mathbb{E}}_{(x_1, \ldots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^{k} f_i(x_i) \right] \right| \geqslant \epsilon,$$

then f_1 is not $(\delta n, \delta)$-product pseudorandom (over the probability space $(\Sigma_1^n, \mu_1^{\otimes n})$, where μ_1 denotes the marginal distribution of μ on coordinate 1).

Quantitatively, when $\epsilon \leqslant o_n(1)$, we can take $\delta = \exp(-\exp(\cdots \exp(\epsilon^{-1})))$, where the number of exponentials is at most $k^{O(k)}$.[7]

Note that in the above theorem, by symmetry, if μ_{-k} has no-Abelian-embeddings, then each of $f_1, \ldots, f_{k-1}$ are not product-pseudorandom. Thus, if any two μ_{-i} and μ_{-j} have no-Abelian embeddings, then all the functions $f_1, \ldots, f_k$ are not product-pseudorandom. This motivates the notion of connectivity/pseudorandomness defined in Definition 1.3.

Corollary 5.3. *Let $k \in \mathbb{N}$, let $\Sigma_1, \ldots, \Sigma_k$ be finite sets, and let μ be a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$ that is pairwise-connected with no-marginal-Abelian-embeddings (see Definition 1.3). Then, for every sufficiently large $n \in \mathbb{N}$ and every $\delta > 0$, there exists $\epsilon = \epsilon(\delta) > 0$, such that the following holds:*

Let $f_i : \Sigma_i^n \to [0,1]$, $i \in [k]$ be functions such that for each $i \in [k]$, the function $f_i - \mu_i(f_i)$ is $(\delta n, \delta)$-product pseudorandom (over the space $(\Sigma_i^n, \mu_i^{\otimes n})$, where μ_i denotes the marginal distribution of μ on coordinate i). Then,

$$\left| \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[\prod_{i=1}^{k} f_i(x_i) \right] - \prod_{i=1}^{k} \mu_i(f_i) \right| \leqslant \epsilon.$$

Quantitatively, when $\delta \leqslant o_n(1)$, we can take $\epsilon = \frac{1}{\log\log\cdots\log\left(\frac{1}{\delta}\right)}$, where the number of logarithms at most $k^{O(k)}$.

Proof. Let $\delta > 0$, and let $\epsilon > 0$ be such that $(\frac{\epsilon}{k}, \delta)$ satisfy Theorem 5.2. Then, we have

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[\prod_{i=1}^{k} f_i(x_i) \right] - \prod_{i=1}^{k} \mu_i(f_i) \right| \\
&\qquad = \left| \sum_{t=1}^{k} \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes n}} \left[\prod_{i=1}^{t-1} \mu_i(f_i) \cdot (f_t(x_t) - \mu_t(f_t)) \cdot \prod_{i=t+1}^{k} f_i(x_i) \right] \right| \leqslant k \cdot \frac{\epsilon}{k} = \epsilon,
\end{aligned}$$

where we first used the triangle inequality and then used Theorem 4.1 for each term. Note that $f_t(x_t) - \mu_t(f_t)$ is 1-bounded for each $t \in [k]$; also note that since there exist two distinct indices $i_1, i_2 \in [k]$, such that μ_{-i_1} and μ_{-i_2} admit no-Abelian-embeddings, the theorem is applicable to each term.

The bound on ϵ is now $\frac{k}{\log\log\cdots\log\left(\frac{1}{\delta}\right)}$, and for $\delta \leqslant o_n(1)$, this is as desired (by possibly increasing the number of logarithms by 1). □

Similar to how Corollary 4.2 is only applicable to high-degree functions, the above corollary is only applicable to functions that are product-pseudorandom. In most applications, the functions we care about are not product-pseudorandom, and we use a suitable generalization of the notion of random restrictions to make arbitrary functions product-pseudorandom. In contrast to Lemma 4.6, the random restriction here shall depend on the functions we are working with.

5.2 Generalized Random Restrictions

In this section we introduce the key notion of a generalized random restriction which appeared in the prior works [BKLM24a, BKLM24c]. To start, we define a specific type of restriction of a function/distribution which enforces that certain subsets of coordinates all take the same value.

Definition 5.4. *Let Σ be a finite alphabet, let $n \in \mathbb{N}$, and let $f : \Sigma^n \to \mathbb{C}$ be a function.*

Let $T \subseteq [n]$. We define the function $f_{=T} : \Sigma^{n-|T|+1} \to \mathbb{C}$ as follows: for $y \in \Sigma$ and $z \in \Sigma^{[n]\setminus T}$, let $x \in \Sigma^n$ be the vector with $x_i = z_i$ for $i \in [n] \setminus T$, and let $x_i = y$ for each $i \in T$; then, $f_{=T}(y,z) := f(x)$.

For disjoint sets $T_1, \ldots, T_m \subseteq [n]$, we write $f_{=T_1,\ldots,T_m} = \left(\ldots(f_{=T_1})_{=T_2}\ldots\right)_{=T_m}$ as the naturally defined function on $n - \sum_{i=1}^{m} |T_i| + m$ coordinates.

With this definition in hand, we are ready to define our notion of generalized restrictions, which both enforces that certain subsets of coordinates take the same value, while fixing the values of other coordinates (i.e., a restriction as in Definition 4.3).

Definition 5.5. *(Generalized Restrictions) Let Σ be a finite alphabet, and $n \in \mathbb{N}$. A generalized restriction ρ on Σ^n is a tuple $\rho = (T_1, T_2, \ldots, T_m, I, z)$ where the sets $T_1, T_2, \ldots, T_m, I \subseteq [n]$ form a disjoint partition of $[n]$, and $z \in \Sigma^I$.*

For a generalized restriction ρ, we call m (also denoted $m(\rho)$) as the number of free variables/coordinates *in ρ, and define the event*

$$E_\rho = \{x \in \Sigma^n : x_i = x_j \ \forall k \in [m],\ i, j \in T_k \text{ and } x_i = z_i \ \forall i \in I\}.$$

For a function $f : \Sigma^n \to \mathbb{C}$, and a generalized restriction $\rho = (T_1, T_2, \ldots, T_m, I, z)$, we define the restricted function $f_\rho : \Sigma^m \to \mathbb{C}$ as $f_\rho(y) = f(x)$, where

$$x_i = \begin{cases} y_j, & i \in T_j,\ j \in [m] \\ z_i, & i \in I \end{cases}.$$

Note that the above is the same as the function $(f_{I \to z})_{=T_1, T_2, \ldots, T_m}$.

Thus, the restriction ρ sets all input variables in each T_j to be equal, and sets variables in coordinates I to value z. Note that this generalizes the usual notions of restrictions (see Definition 4.3) which is the same as above with each T_j of size 1. We observe that generalized restrictions are closed under composition:

Observation 5.6. *(Composition of generalized restrictions) Let Σ be a finite alphabet, let $n \in \mathbb{N}$, and let $f : \Sigma^n \to \mathbb{C}$ be a function. Let $\rho = (T_1, \ldots, T_m, I, z)$ be a generalized random restriction on Σ^n with $m \leqslant n$ free coordinates, and let $\rho' = (S_1, \ldots, S_k, J, w)$ be a generalized random restriction on Σ^m with $k \leqslant m$ free coordinates.*

Then, the function $(f_\rho)_{\rho'} : \Sigma^k \to \mathbb{C}$ equals the function $f_{\rho' \circ \rho}$, where $\rho' \circ \rho$ is a generalized random restriction on on Σ^n with k free coordinates, defined as $\rho' \circ \rho = (U_1, \ldots, U_k, K, y)$, where:

1. *For each $i \in [k]$, it holds that $U_i = \bigcup_{j \in S_i} T_j$.*
2. *$K = \bigcup_{j \in J} T_j \cup I$.*
3. *$y \in \Sigma^K$ is given as follows: for $i \in I$, it holds that $y_i = z_i$; for $i \in T_j$, for some $j \in J$, it holds that $y_i = w_j$.*

Proof. This is verified by simply following the definitions. □

Finally, we define a generalized random restriction. Formally, this is a distribution over generalized restrictions such that the overall original distribution $\mu^{\otimes n}$ is preserved on average.

Definition 5.7. *(Generalized Random Restriction) Let (Σ, μ) be a probability space, and let $n \in \mathbb{N}$. Let $1 \leqslant m \leqslant n$, and let $\epsilon \in [0, 1]$. An (m, ϵ)-generalized random restriction on Σ^n is a distribution $\mathcal{R}$ over generalized restrictions on Σ^n (see Definition 5.5) satisfying:*

1. *Each $\rho \in \operatorname{supp}(\mathcal{R})$ has at least m free coordinates.*
2. *The distribution obtained by first sampling $\rho \sim \mathcal{R}$ and then sampling $x \in \Sigma^n$ conditioned on the restriction ρ, is close (in ℓ_1-norm) to the distribution $\mu^{\otimes n}$. Formally,*

$$\left\| \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}} \left[\mu^{\otimes n} | E_\rho \right] - \mu^{\otimes n} \right\|_1 \leqslant \epsilon.$$

Note that property 2 justifies the use of the term *random restriction*. Property 2 additionally implies that $|\mathbb{E}_{\rho \sim \mathcal{R}} [\mu(f_\rho)] - \mu(f)| \leqslant \epsilon$ for every 1-bounded $f : \Sigma^n \to \mathbb{C}$.

5.3 Product Pseudorandomness under Generalized Random Restrictions

In this subsection, we show that arbitrary functions can be made product-pseudorandom under suitable generalized random restrictions. We start by showing an increment lemma, which we later iterate. The proof of this lemma follows Lemma 7.1 in [BKLM24c].

Lemma 5.8. *(Increment Lemma) Let (Σ, μ) be a probability space with $|\Sigma| \geqslant 2$, let $c = \frac{1}{100|\Sigma|^2}$, and let $n \in \mathbb{N}$ be sufficiently large.*

For any $\gamma \in [n^{-c}, 1]$, the following holds: Let $g : \Sigma^n \to \mathbb{C}$ be a 1-bounded function such that $g - \mu(g)$ is not $(\sqrt{n}, \gamma)$-product pseudorandom. Then, there exists a $(n^c, n^{-0.01})$-generalized random restriction $\mathcal{R}$ on Σ^n (see Definition 5.7) such that

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}} |\mu(g_\rho)|^2 \geqslant |\mu(g)|^2 + \frac{\gamma^3}{16}.$$

Proof. Let $\delta = 1/\sqrt{n}$ and $f = g - \mu(g)$; then, we know that f is 2-bounded, and not $(\delta n, \gamma)$-product pseudorandom. That is, for some $\delta' \geqslant \delta$,

$$\Pr_{I \sim_{1-\delta'} [n], z \sim \mu^{\otimes I}} \left[\exists \{P_i : \Sigma \to \mathbb{C},\ \|P_i\|_\infty \leqslant 1\}_{i \in \bar{I}} \text{ with } \left| \mathop{\mathbb{E}}_{x \sim \mu^{\otimes \bar{I}}} \left[f_{I \to z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geqslant \gamma \right] \geqslant \gamma.$$

This implies that with probability at least $\gamma/2$ over $I \sim_{1-\delta'} [n]$, it holds that

$$\Pr_{z \sim \mu^{\otimes I}} \left[\exists \{P_i : \Sigma \to \mathbb{C},\ \|P_i\|_\infty \leqslant 1\}_{i \in \bar{I}} \text{ with } \left| \mathop{\mathbb{E}}_{x \sim \mu^{\otimes \bar{I}}} \left[f_{I \to z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geqslant \gamma \right] \geqslant \frac{\gamma}{2}.$$

By a Chernoff bound (Fact A.4), we know $\Pr_{I \sim_{1-\delta'} [n]}[|I| \geqslant (1 - \delta/2)\, n] \leqslant e^{-\delta n/8} < \gamma/2$. Hence, there exists $I \subseteq [n]$, with $|I| \leqslant (1 - \delta/2)\, n$, and such that

$$\Pr_{z \sim \mu^{\otimes I}} \left[\exists \{P_i : \Sigma \to \mathbb{C},\ \|P_i\|_\infty \leqslant 1\}_{i \in \bar{I}} \text{ with } \left| \mathop{\mathbb{E}}_{x \sim \mu^{\otimes \bar{I}}} \left[f_{I \to z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geqslant \gamma \right] \geqslant \frac{\gamma}{2}.$$

We fix such an I, and relabel $\bar{I}$ as $[m]$, with $m \geqslant \delta n/2$. Also, we define the set $\mathcal{G}$ as the set of all z for which the condition in the above equation holds; that is

$$\mathcal{G} = \left\{ z \in \Sigma^I : \exists \{P_i : \Sigma \to \mathbb{C},\ \|P_i\|_\infty \leqslant 1\}_{i \in \bar{I}} \text{ with } \left| \mathop{\mathbb{E}}_{x \sim \mu^{\otimes \bar{I}}} \left[f_{I \to z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geqslant \gamma \right\}.$$

The required distribution $\mathcal{R}$ is now defined as follows: Let $z \sim \mu^{\otimes I}$ be chosen randomly. Consider the following cases:

- If $z \notin \mathcal{G}$: then all the coordinates in $[m]$ are kept alive; formally, output the restriction $\rho = (T_1, \ldots, T_m, I, z)$ where each $T_i = \{i\} \subseteq [m] = \bar{I}$.
- Suppose $z \in \mathcal{G}$: Let $\{P_i : \Sigma \to \mathbb{C},\ \|P_i\|_\infty \leqslant 1\}_{i \in [m]}$ be such that

$$\left| \mathop{\mathbb{E}}_{x \sim \mu^{\otimes m}} \left[f_{I \to z}(x) \prod_{i \in [m]} P_i(x_i) \right] \right| \geqslant \gamma.$$

Let $v_1, \dots, v_m : \Sigma \to \mathbb{R}/\mathbb{Z}$ be such that $P_j(x) = e^{2\pi i v_j(x)}$ for all $j \in [m], x \in \Sigma$.[10]

Let $s = |\Sigma| \geqslant 2$ and $r = \lceil m^{\frac{1}{16s^2}} \rceil$. By the pigeonhole principle, we can find disjoint sets $S_1, S_2, \dots, S_r \subseteq [m]$, each of size $|S_i| = \lceil \sqrt{m}/2 \rceil$, such that for each $i \in [r]$ and $j, j' \in S_i$, we have $\left\|v_j - v_{j'}\right\|_\infty \leqslant m^{-\frac{1}{2s}}$. Consider arbitrary indices $a_1 \in S_1,\ a_2 \in S_2, \dots, a_r \in S_r$, and for each $i \in [r]$, let $1 \leqslant k_i \leqslant m^{\frac{1}{4s}}$ be such that $\|k_i v_{a_i}\|_\infty \leqslant m^{-\frac{1}{8s^2}}$; note that such a k_i exists by applying the pigeonhole principle on the vectors $v_{a_i}, 2v_{a_i}, 3v_{a_i}, \dots, \lfloor m^{\frac{1}{4s}} \rfloor \cdot v_{a_i}$.

Now, for each $i \in [r]$, let T_i be a uniformly random subset of S_i of size k_i. We perform the following (generalized) random restriction: For each $i \in [r]$, force $x_j = x_{j'}$ for each $j, j' \in T_i$, and then do a uniform random restriction on coordinates $J := [m] \setminus (T_1 \cup T_2 \cup \dots T_r)$, given by $u \sim \mu^{\otimes J}$.

Finally the random restriction is given by $\rho = (T_1, \dots, T_r,\ I \cup J,\ (z, u))$.

We show that $\mathcal{R}$ satisfies the three conditions of the lemma statement.

1. By definition, each $\rho \in \mathrm{supp}(\mathcal{R})$ contains at least $r \geqslant m^{\frac{1}{16s^2}} \geqslant \left(\frac{\delta n}{2}\right)^{\frac{1}{16s^2}}$ free coordinates. This is at least n^c.

2. It suffices to show that for every choice of $z \in \mu^{\otimes I}$, it holds that
$$\left\|\mathbb{E}\left[\mu^{\otimes n} \mid E_\rho, z\right] - \left(\mu^{\otimes n} \mid z\right)\right\|_1 \leqslant n^{-0.01}.$$
In the first case when $z \notin \mathcal{G}$, this holds with error zero.

   In the second case $z \in \mathcal{G}$, denoting $\rho' = (T_1, \dots, T_r, J, u)$ as the generalized random restriction on $\Sigma^m = \Sigma^{\bar{I}}$, we can bound the expression on the left hand side using Lemma A.1 as
$$\left\|\mathop{\mathbb{E}}_{\rho}\left[\mu^{\otimes m} \mid E_{\rho'}\right] - \mu^{\otimes m}\right\|_1 \leqslant \sum_{i=1}^{r} \frac{Ck_i}{\sqrt{|S_i|}} \leqslant \sqrt{2}C \cdot r \cdot m^{\frac{1}{4s} - \frac{1}{4}} \leqslant O\left(m^{\frac{1}{16s^2} + \frac{1}{4s} - \frac{1}{4}}\right) \leqslant n^{-0.01}. \quad (3)$$
   For the last inequality, we used that $s \geqslant 2$ and $m \geqslant \delta n/2 = \sqrt{n}/2$.

It remains to show the third condition, which is the ℓ_2 increment. For this, we consider the second case of the generalized random restriction (which occurs with probability $\gamma/2$) and consider some fixed $z \in \mathcal{G}$. Let $P = P_1 \cdot P_2 \cdots P_m$, and $Q = (P_{J \to u})_{=T_1, \dots, T_r}$. First, we show that for every choice of $T_1, \dots, T_r, u$, the restricted function Q is nearly constant: for any $x \in \Sigma^r$:
$$\begin{aligned}\left|Q(x) - \prod_{j \in J} P_j(u_j)\right| &= \left|\left(\prod_{j \in J} P_j(u_j)\right) \cdot \left[\exp\left(2\pi i \sum_{t=1}^{r} \sum_{j \in T_t} v_j(x_t)\right) - 1\right]\right| \\ &\leqslant \left|\sum_{t=1}^{r} \sum_{j \in T_t} v_j(x_t)\right| \leqslant \sum_{t=1}^{r} \left\|\sum_{j \in T_t} v_j\right\|_\infty.\end{aligned}$$

Now, for any $i \in [r]$, we have
$$\left\|\sum_{j \in T_i} v_j\right\|_\infty \leqslant \|k_i v_{a_i}\|_\infty + \sum_{j \in T_i} \|v_j - v_{a_i}\|_\infty \leqslant m^{-\frac{1}{8s^2}} + k_i \cdot m^{-\frac{1}{2s}} \leqslant 2 \cdot m^{-\frac{1}{8s^2}}.$$

[10] It is without loss of generality that $|P_j(x)| = 1$ for each $j \in [m], x \in \Sigma$. This is because there always exist such functions $P_1, \dots, P_m$ maximizing the quantity $\left|\mathbb{E}_{x \sim \mu^{\otimes m}}\left[f_{I \to z}(x) \prod_{i \in [m]} P_i(x_i)\right]\right|$, since the expression is *multilinear* in the $P_i(x_i)$'s.

This implies that $\left|Q(x) - \prod_{j\in J} P_j(u_j)\right| \leqslant 2 \cdot r \cdot m^{-\frac{1}{8s^2}} \leqslant 2 \cdot m^{-\frac{1}{16s^2}}$.

Now, using the above and Equation 3, along with the fact that f is 2-bounded, we get

$$\begin{aligned}
\gamma &\leqslant \left| \mathop{\mathbb{E}}_{x\sim\mu^{\otimes m}} [f_{I\to z}(x)P(x)] \right| \\
&\leqslant \left| \mathop{\mathbb{E}}_{T_1,\dots,T_r,u} \mathop{\mathbb{E}}_{x\sim\mu^{\otimes r}} \left[(f_{I\to z, J\to u})_{=T_1,\dots T_r}(x) \cdot (P_{J\to u})_{=T_1,\dots,T_r}(x) \right] \right| + 2\cdot n^{-0.01} \\
&\leqslant \mathop{\mathbb{E}}_{T_1,\dots,T_r,u} \left| \mathop{\mathbb{E}}_{x\sim\mu^{\otimes r}} [(f_{I\to z, J\to u})_{=T_1,\dots T_r}(x)] \right| + 4\cdot m^{-\frac{1}{16s^2}} + 2\cdot n^{-0.01} \\
&= \mathop{\mathbb{E}}_{T_1,\dots,T_r,u} |\mu(f_\rho)| + 4\cdot m^{-\frac{1}{16s^2}} + 2\cdot n^{-0.01} \\
&= \mathop{\mathbb{E}}_{T_1,\dots,T_r,u} |\mu(g_\rho) - \mu(g)| + 4\cdot m^{-\frac{1}{16s^2}} + 2\cdot n^{-0.01} \\
&\leqslant \mathop{\mathbb{E}}_{T_1,\dots,T_r,u} |\mu(g_\rho) - \mu(g)| + 6\cdot n^{-\frac{1}{50s^2}}.
\end{aligned}$$

Using that $\gamma \geqslant n^{-c} \geqslant 12\cdot n^{-\frac{1}{50s^2}}$, and Cauchy-Schwarz, we get

$$\mathop{\mathbb{E}}_{T_1,\dots,T_r,u} |\mu(g_\rho) - \mu(g)|^2 \geqslant \left(\frac{\gamma}{2}\right)^2 = \frac{\gamma^2}{4}.$$

Now, averaging over z, and observing that $\Pr[z\in\mathcal{G}] \geqslant \gamma/2$, we get

$$\begin{aligned}
\frac{\gamma^3}{8} &\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} |\mu(g_\rho) - \mu(g)|^2 \\
&= \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} |\mu(g_\rho)|^2 + |\mu(g)|^2 - \overline{\mu(g)} \cdot \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} [\mu(g_\rho)] - \mu(g) \cdot \overline{\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} [\mu(g_\rho)]} \\
&\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} |\mu(g_\rho)|^2 + |\mu(g)|^2 - 2\,|\mu(g)|^2 + 2\cdot n^{-0.01} \\
&\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} |\mu(g_\rho)|^2 - |\mu(g)|^2 + \frac{\gamma^3}{16}.
\end{aligned}$$

□

Now, we iterate the increment lemma, and show that any small collection of functions can be made product pseudorandom under generalized random restrictions.

Proposition 5.9. *(Uniformization) Let (Σ,μ) be a probability space with $|\Sigma| \geqslant 2$, let $c = \frac{1}{100|\Sigma|^2}$, and let $r, n \in \mathbb{N}$. Let $g_1,\dots,g_r : \Sigma^n \to \mathbb{C}$ be a collection of 1-bounded functions.*

Let $0 < \delta, \gamma \leqslant 1$ be such that $\frac{1}{\delta\gamma^3} \leqslant \frac{1}{200r\log(1/c)} \cdot \log\log n$, and let $T = \left\lceil \frac{50r}{\delta\gamma^3} \right\rceil \in \mathbb{N}$. Then, there exists a $(n^{c^T},\ n^{-c^T})$-generalized random restriction $\mathcal{R}$ on Σ^n (see Definition 5.7), such that the following holds with probability at least $1-\delta$ over $\rho\sim\mathcal{R}$:

Suppose $m(\rho)$ denotes the number of free coordinates in ρ. Then, for every $i\in[r]$, the function $(g_i)_\rho : \Sigma^{m(\rho)} \to \mathbb{C}$ is such that $(g_i)_\rho - \mu((g_i)_\rho)$ is $(\sqrt{m(\rho)},\gamma)$-product pseudorandom.

Proof. For any generalized restriction ρ, we say that a restriction ρ is bad for $i\in[r]$ if the function $(g_i)_\rho : \Sigma^{m(\rho)} \to \mathbb{C}$ is such that $(g_i)_\rho - \mu((g_i)_\rho)$ is not $(\sqrt{m(\rho)},\gamma)$-product pseudorandom. We say ρ is bad if it is bad for some $i\in[r]$, and say that ρ is good otherwise.

Let $\mathcal{R}^{(0)}$ be the random restriction that does nothing. For $t = 1,2,\dots,T$, define a generalized random restriction $\mathcal{R}^{(t)}$ as follows:

1. Choose $\rho \sim \mathcal{R}^{(t-1)}$.
2. If ρ is good, output the restriction ρ.
3. Else, consider an arbitrary $i \in [r]$ such that ρ is bad for i. Let $\mathcal{R}_\rho$ be the $(m(\rho)^c, m(\rho)^{-0.01})$-generalized random restriction on $\Sigma^{m(\rho)}$ obtained by applying Lemma 5.8 to the function $(g_i)_\rho$. Choose $\rho' \sim \mathcal{R}_\rho$ and output $\rho' \circ \rho$.

By induction, it is easily verified that for each $t = 0, 1, \ldots, T$, the random restriction $\mathcal{R}^{(t)}$ is a $(m^{(t)}, \epsilon^{(t)})$-generalized random restriction on Σ^n, with

$$m^{(t)} = \left(m^{(t-1)}\right)^c = n^{c^t} \geqslant n^{c^T} \geqslant 2^{\sqrt{\log n}}.$$

and

$$\epsilon^{(t)} = \epsilon^{(t-1)} + \left(m^{(t-1)}\right)^{-0.01} \leqslant t \cdot n^{-\frac{1}{100}c^{t-1}} \leqslant T \cdot n^{-\frac{1}{100}c^{T-1}} \leqslant T \cdot n^{-4c^T} \leqslant n^{-c^T} =: \epsilon.$$

In the last inequality above, we used that $\frac{100r}{\delta\gamma^3} \leqslant n^{c^T}$, which follows from the choice of parameters. Note that this also implies that $\gamma \geqslant n^{-c^T}$, and this is at least $m(\rho)^{-c}$ whenever Lemma 5.8 is applied above in Step 3, as needed in the assumption of the lemma.

Now, suppose at some point we found some ρ that is bad for $i \in [r]$, and applied Step 3 above to get $\rho' \sim \mathcal{R}_\rho$; then it holds:

$$\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_\rho}\left[\left|\mu\left((g_i)_{\rho' \circ \rho}\right)\right|^2\right] \geqslant \left|\mu\left((g_i)_\rho\right)\right|^2 + \frac{\gamma^3}{16},$$

and for each $j \neq i$,

$$\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_\rho}\left[\left|\mu\left((g_j)_{\rho' \circ \rho}\right)\right|^2\right] \geqslant \left|\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_\rho}\left[\mu\left((g_j)_{\rho' \circ \rho}\right)\right]\right|^2 \geqslant \left|\mu\left((g_j)_\rho\right)\right|^2 - 2\epsilon,$$

where we used Cauchy-Schwarz, the second property in Defintiion 5.7, and Lemma A.2. In particular, it holds

$$\sum_{j=1}^r \mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_\rho}\left[\left|\mu\left((g_j)_{\rho' \circ \rho}\right)\right|^2\right] \geqslant \sum_{j=1}^r \left|\mu\left((g_j)_\rho\right)\right|^2 + \left(\frac{\gamma^3}{16} - 2(r-1)\epsilon\right) \geqslant \sum_{j=1}^r \left|\mu\left((g_j)_\rho\right)\right|^2 + \frac{\gamma^3}{32}.$$

In the last inequality, we used $\gamma^3 \geqslant 64r\epsilon$, which follows from the inequality $\frac{100r}{\delta\gamma^3} \leqslant n^{c^T}$ that we used earlier. With the above, we get that for every $t = 1, \ldots, T$,

$$\sum_{j=1}^r \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}^{(t)}}\left[\mu((g_j)_\rho)^2\right] \geqslant \sum_{j=1}^r \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}^{(t-1)}}\left[\mu((g_j)_\rho)^2\right] + \frac{\gamma^3}{32} \cdot \Pr_{\rho \sim \mathcal{R}^{(t-1)}}\left[\rho \text{ is bad}\right].$$

Suppose, for the sake of contradiction that $\Pr_{\rho \sim \mathcal{R}^{(t-1)}}[\rho \text{ is bad}] \geqslant \delta$ for each $t \in [T]$. Then, we have

$$r \geqslant \sum_{j=1}^r \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}^{(T)}}\left[\mu((g_j)_\rho)^2\right] \geqslant \sum_{j=1}^r \mu(g_j)^2 + \frac{T\delta\gamma^3}{32} \geqslant 0 + \frac{50}{32} \cdot r > r,$$

which is a contradiction. Hence, for some $t = 0, 1, \ldots, T-1$, the generalized random restriction $\mathcal{R}^{(t)}$ satisfies the statement of the lemma. □

Corollary 5.10. *Let (Σ, μ) be a probability space with $|\Sigma| \geqslant 2$, and let $n \in \mathbb{N}$. Let $g_1, \ldots, g_r : \Sigma^n \to \mathbb{C}$ be a collection of 1-bounded functions, with $r = O(1)$.*

Let $0 < \gamma \leqslant 1$ be such that $\frac{1}{\gamma} \leqslant o(\log \log n)^{1/4}$. Then, there exists a $(\frac{1}{\eta}, \eta)$-generalized random restriction $\mathcal{R}$ on Σ^n, with $\eta = n^{-\exp(-1/\gamma^4)}$, such that the following holds with probability at least $1 - \gamma$ over $\rho \sim \mathcal{R}$: For every $i \in [r]$, the function $(g_i)_\rho : \Sigma^{m(\rho)} \to \mathbb{C}$ is such that $(g_i)_\rho - \mu((g_i)_\rho)$ is $(\sqrt{m(\rho)}, \gamma)$-product pseudorandom.

The constant in the exp *depends only on $r, |\Sigma|$.*

Proof. This follows from Proposition 5.9 by choosing $\delta = \gamma$. □

5.4 Conditional Generalized Restrictions

We shall also be interested in a conditional notion of generalized restrictions, defined as follows:

Definition 5.11. *Let (Σ, μ) be a probability space, and let $n \in \mathbb{N}$. Let $\mathcal{R}$ be a (m, ϵ)-generalized random restriction (see Definition 5.7) over Σ^n, and let $E \subseteq \Sigma^n$ be an event such that $\Pr[E] > \epsilon$.*

We define the corresponding conditional generalized restriction, denoted $\mathcal{R}|E$, as the distribution over generalized random restrictions ρ given by:

$$(\mathcal{R}|E)[\rho] = \frac{\Pr[E|E_\rho] \cdot \mathcal{R}[\rho]}{\mathbb{E}_{\rho' \sim \mathcal{R}}\left[\Pr\left[E|E_{\rho'}\right]\right]}.$$

Note that the denominator in the above expression satisfies $\mathbb{E}_{\rho' \sim \mathcal{R}}\left[\Pr\left[E|E_{\rho'}\right]\right] \geqslant \Pr[E] - \epsilon > 0$ by the assumption $\Pr[E] > \epsilon$.

These satisfy the following properties, which informally say that as long as $\Pr[E]$ is much larger than ϵ, then conditional generalized restrictions preserve probabilities and distributions on average.

Lemma 5.12. *Under the setting of Definition 5.11, we have*

1. *For any ρ, it holds that*

$$\left|(\mathcal{R}|E)[\rho] - \frac{\Pr[E|E_\rho] \cdot \mathcal{R}[\rho]}{\Pr[E]}\right| \leqslant (\mathcal{R}|E)[\rho] \cdot \frac{\epsilon}{\Pr[E]}.$$

2. *The distribution obtained by first sampling $\rho \sim \mathcal{R}|E$ and then sampling x from $\mu^{\otimes n}|E$ conditioned on the restriction ρ, is close (in ℓ_1-norm) to the distribution $\mu^{\otimes n}|E$. Formally,*

$$\left\| \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}|E}\left[\mu^{\otimes n}|E_\rho, E\right] - \left(\mu^{\otimes n}|E\right) \right\|_1 \leqslant \frac{2\epsilon}{\Pr[E]}.$$

Proof. By Definition 5.7, it follows that $\left|\mathbb{E}_{\rho' \sim \mathcal{R}}\left[\Pr\left[E|E_{\rho'}\right]\right] - \Pr[E]\right| \leqslant \epsilon$. Hence, for every ρ, it holds that

$$\left|(\mathcal{R}|E)[\rho] - \frac{\Pr[E|E_\rho] \cdot \mathcal{R}[\rho]}{\Pr[E]}\right| = (\mathcal{R}|E)[\rho] \cdot \frac{\left|\mathbb{E}_{\rho' \sim \mathcal{R}}\left[\Pr\left[E|E_{\rho'}\right]\right] - \Pr[E]\right|}{\Pr[E]} \leqslant (\mathcal{R}|E)[\rho] \cdot \frac{\epsilon}{\Pr[E]}.$$

Using this, we also have

$$\begin{aligned}
\left\| \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}|E} \left[\mu^{\otimes n}|E_\rho, E\right] - \left(\mu^{\otimes n}|E\right) \right\|_1 &= \sum_{x\in\Sigma^n} \left| \left(\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}|E} \left[\mu^{\otimes n}|E_\rho, E\right] \right)[x] - \Pr[x|E] \right| \\
&= \sum_{x\in E} \left| \sum_\rho (\mathcal{R}|E)[\rho] \cdot \Pr[x|E, E_\rho] - \Pr[x|E] \right| \\
&\leqslant \sum_{x\in E} \left| \sum_\rho \frac{\Pr[E|E_\rho]\cdot\mathcal{R}[\rho]}{\Pr[E]} \cdot \Pr[x|E, E_\rho] - \Pr[x|E] \right| + \frac{\epsilon}{\Pr[E]} \\
&= \sum_{x\in E} \left| \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}} \frac{\Pr[x, E|E_\rho]}{\Pr[E]} - \Pr[x|E] \right| + \frac{\epsilon}{\Pr[E]} \\
&= \frac{\sum_{x\in E} \left|\mathbb{E}_{\rho\sim\mathcal{R}} \Pr[x|E_\rho] - \Pr[x]\right|}{\Pr[E]} + \frac{\epsilon}{\Pr[E]} \\
&\leqslant \frac{2\epsilon}{\Pr[E]}.
\end{aligned}$$

□

6 Multiplayer Games

We formally define notions associated to multiplayer games, and establish some notation.

Definition 6.1. *(Multiplayer Game) A k-player game $\mathcal{G}$ is a tuple $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$, where the question set $\mathcal{X} = \mathcal{X}^1 \times \cdots \times \mathcal{X}^k$ and the answer set $\mathcal{A} = \mathcal{A}^1 \times \cdots \times \mathcal{A}^k$ are finite sets, Q is a probability distribution over $\mathcal{X}$, and $V : \mathcal{X} \times \mathcal{A} \to \{0, 1\}$ is a predicate.*

The game $\mathcal{G}$ proceeds as follows: A verifier samples questions $X = (X^1, \ldots, X^k) \sim Q$; then, for each $j \in [k]$, the verifier sends the question $X^j \in \mathcal{X}^j$ to the j^{th} player, to which the player responds back with answer $A^j \in \mathcal{A}^j$. Finally, the verifier declares that the players win if and only if $V\left(X^1, \ldots, X^k, A^1, \ldots, A^k\right) = 1$.

Definition 6.2. *(Game Value) Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ be a k-player game.*

For a sequence $\left(f^j : \mathcal{X}^j \to \mathcal{A}^j\right)_{j\in[k]}$ of functions, define the function $f = f^1 \times \cdots \times f^k : \mathcal{X} \to \mathcal{A}$ by $f(x^1, \ldots, x^k) = \left(f^1(x^1), \ldots, f^k(x^k)\right)$. We use the term product functions *to denote functions f defined in this manner, and the functions $(f^j)_{j\in[k]}$ are called* player strategies.

The value $\mathrm{val}(\mathcal{G})$ of the game $\mathcal{G}$ is defined as

$$\mathrm{val}(\mathcal{G}) = \max_{f=f^1\times\cdots\times f^k} \Pr_{X\sim Q} \left[V(X, f(X)) = 1\right],$$

where the maximum is over all product functions $f = f^1 \times \cdots \times f^k$.

Fact 6.3. *The value of the game is unchanged even if we allow the player strategies to be randomized; that is, we allow the strategies to depend on some additional shared and private randomness. This is because there always exists an optimal fixed value for the randomness.*

Next, we define the parallel repetition of a k-player game, which corresponds to playing n independent copies of the game in parallel.

Definition 6.4. *(Parallel Repetition) Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ be a k-player game. We define its n-fold repetition as $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$. The sets $\mathcal{X}^{\otimes n}$, $\mathcal{A}^{\otimes n}$ are defined to be the n-fold product*

of the sets $\mathcal{X}, \mathcal{A}$ with themselves respectively.[11] *The distribution $Q^{\otimes n}$ is the n-fold product of the distribution Q with itself, that is, $Q^{\otimes n}[x] = \prod_{i=1}^{n} Q[x_i]$ for each $x \in \mathcal{X}^{\otimes n}$. The predicate $V^{\otimes n}$ is defined as $V^{\otimes n}(x, a) = \bigwedge_{i=1}^{n} V(x_i, a_i)$.*

Notation: We use subscripts to denote the coordinates in the parallel repetition, and superscripts to denote the players. That is, for any $S \subseteq [n], T \subseteq [k]$, we shall use x_S^T to denote the questions in coordinates S, that the players in set T receive. For example, for $i \in [n]$ and $j \in [k]$, we will use x_i^j to refer to the question to the j^{th} player in the i^{th} repetition of the game. Similarly, x_i will refer to the vector of questions to the k players in the i^{th} repetition, and x^j will refer to the vector of questions received by the j^{th} player over all repetitions. We use x^{-j} to refer to the questions to all players except the j^{th} player, and use x_{-i} to refer to the questions in all coordinates except the i^{th} coordinate.

7 Games with No Abelian Embeddings

Theorem 7.1. *Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ be a k-player game such that the distribution Q has no-Abelian-embeddings (see Definition 1.2), and such that $\mathrm{val}(\mathcal{G}) < 1$. Then, there exists a constant $C \in \mathbb{N}$, $C \leqslant k^{O(k)}$, such that for every sufficiently large $n \in \mathbb{N}$,*

$$\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant \frac{1}{\log \log \cdots \log n},$$

where the number of logarithms is C.

In this section, we shall prove this theorem. Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ be a k-player game such that Q has no-Abelian-embeddings, and such that $\mathrm{val}(\mathcal{G}) < 1$. For some sufficiently large n, consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, P = Q^{\otimes n}, V^{\otimes n})$, and fix an optimal strategy for the k players in this game. Let $X = (X^1, \ldots, X^k)$ be the random variable denoting the questions to the k players in the game $\mathcal{G}^{\otimes n}$, and let $A = (A^1, \ldots, A^k)$ be the random variable denoting the answers of the players using these strategies. For each $i \in [n]$, let Win_i (resp. Lose_i) be the event that $V(X_i, A_i) = 1$ (resp. $V(X_i, A_i) = 0$); that is, the players win (resp. lose) the i^{th} coordinate of the game.

We introduce some parameters that will be useful:

$$\delta = (\log n)^{-1/3},\ T = \left\lceil \frac{1}{\delta^2} \right\rceil,\ \epsilon = \frac{1}{\underbrace{\log \log \cdots \log}_{C \text{ times}} n},\ \alpha = \sqrt{\epsilon},$$

where $2 \leqslant C \leqslant k^{O(k)}$ is a constant so that (ϵ, δ) satisfy Corollary 4.2 with respect to the probability space $(\mathcal{X}, Q)$; recall that the distribution Q has no-Abelian-embeddings.

Let $E = E^1 \times \cdots \times E^k \subseteq (\mathcal{X}^1)^{\otimes n} \times \cdots \times (\mathcal{X}^k)^{\otimes n} = \mathcal{X}^{\otimes n}$ be an arbitrary product event with $\Pr_{Q^{\otimes n}}[E] \geqslant \alpha$. We prove the following lemma:

Lemma 7.2. *For parameters chosen as above and $\Pr[E] \geq \alpha$, it holds that*

$$\mathop{\mathbb{E}}_{i \sim [n]} [\Pr[\mathrm{Win}_i \mid E]] \leqslant \mathrm{val}(\mathcal{G}) + o_n(1),$$

where the expectation is over $i \in [n]$ chosen uniformly at random.

Assuming this, the main theorem follows directly:

[11] We use the notation $\mathcal{X}^{\otimes n}$ instead of the usual $\mathcal{X}^n$ so as to avoid confusion with the sets $\mathcal{X}^1, \ldots, \mathcal{X}^k$.

Proof of Theorem 7.1. This follows by combining a standard inductive parallel repetition proof (see Lemma B.1) and Lemma 7.2, and using the fact that $\mathrm{val}(\mathcal{G}) < 1$; note that the theorem works with the constant $C + 1 \leqslant k^{O(k)}$. □

The remainder of this section is devoted to proving Lemma 7.2.

7.1 Strategy for a Single Copy of the Game

The above lemma is proven via an *embedding argument*, where we try to embed a single copy of the game $\mathcal{G}$ into the i^{th} coordinate of the game $\mathcal{G}^{\otimes n}$ (while conditioning on E). Formally, we construct the following randomized strategy for the game $\mathcal{G}$:

1. The verifier samples $\tilde{X} \sim Q$, and for each $j \in [k]$, gives player j the input $\tilde{X}^j$.
2. Using shared randomness, the players sample:
   (a) $i \sim [n]$ uniformly at random.
   (b) $p \in (0, 1]$ uniformly at random from the set $\left\{1, \delta, \delta^2, \ldots, \delta^{T-1}\right\}$.
   (c) $I \sim_p [n] \setminus \{i\}$; let $I' = [n] \setminus (I \cup \{i\})$.
   (d) $Z \sim P_{X_{I'}|E} = Q^{\otimes I'} \,|\, E$.
3. For each $j \in [k]$, player j does the following:
   We say that $\tilde{X}^j, Z^j$ are consistent with E^j if
   $$\left\{x^j \in (\mathcal{X}^j)^{\otimes n} : x^j \in E^j,\ x^j_i = \tilde{X}^j,\ x^j_{I'} = Z^j\right\} \neq \emptyset.$$
   (a) If $\tilde{X}^j, Z^j$ are not consistent with E^j, output an arbitrary answer from $\mathcal{A}^j$; for example, we may assume the output is the first element of $\mathcal{A}^j$ under some ordering.
   (b) Else, using private randomness, output $\tilde{A}^j \sim P_{A^j_i \,|\, X^j \in E^j,\ X^j_i = \tilde{X}^j,\ X^j_{I'} = Z^j}$.
4. Let $\tilde{A} = (\tilde{A}^1, \tilde{A}^2, \ldots, \tilde{A}^k)$; the players win if and only if $V(\tilde{X}, \tilde{A}) = 1$.

Let L be the event that $V(\tilde{X}, \tilde{A}) = 0$; that is, the players lose the game $\mathcal{G}$ when using the above strategy. We shall prove:

Lemma 7.3.

$$\Pr[L] \leqslant \mathop{\mathbb{E}}_{i \sim [n]} [\Pr[\mathrm{Lose}_i \,|\, E]] + \frac{4\,|\mathcal{A}|}{\alpha} \cdot \left(\frac{k}{\delta T} + \epsilon\right) + \sqrt{\frac{2}{\delta^T n} \cdot \log_2\left(\frac{1}{\alpha}\right)}.$$

Assuming this, we can easily complete the proof of Lemma 7.2:

Proof of Lemma 7.2. It must hold that $\Pr[L] \geqslant 1 - \mathrm{val}(\mathcal{G})$. Hence, by Lemma 7.3, we get

$$\mathop{\mathbb{E}}_{i \sim [n]} [\Pr[\mathrm{Win}_i \,|\, E]] \leqslant \mathrm{val}(\mathcal{G}) + \frac{4\,|\mathcal{A}|}{\alpha} \cdot \left(\frac{k}{\delta T} + \epsilon\right) + \sqrt{\frac{2}{\delta^T n} \cdot \log_2\left(\frac{1}{\alpha}\right)}.$$

By our choice of parameters, we have

1. $\frac{4|\mathcal{A}|k}{\alpha\delta T} \leqslant O\left(\frac{\delta}{\alpha}\right) \leqslant O\left(\frac{\epsilon}{\alpha}\right) = O\left(\sqrt{\epsilon}\right) \leqslant o_n(1)$.

2. $\frac{4|\mathcal{A}|\epsilon}{\alpha} \leqslant O\left(\frac{\epsilon}{\alpha}\right) = O\left(\sqrt{\epsilon}\right) \leqslant o_n(1)$.

3. $\delta^T \geqslant \frac{1}{\sqrt{n}}$, and so $\frac{1}{\delta^T n}\log_2(1/\alpha) \leqslant \frac{\log_2 n}{\sqrt{n}} \leqslant o_n(1)$.

Hence, $\mathbb{E}_{i\sim[n]}\left[\Pr[\mathrm{Win}_i \mid E]\right] \leqslant \mathrm{val}(\mathcal{G}) + o_n(1)$, as desired. □

Now, we focus on proving Lemma 7.3. We break up the losing probability into two terms, and bound them individually, as follows:

Lemma 7.4.

$$\mathop{\mathbb{E}}_{i,p,I,Z}\left\|P_{X_i|E,X_{I'}=Z} - Q\right\|_1 \leqslant \sqrt{\frac{2}{\delta^T n}\cdot\log_2\left(\frac{1}{\alpha}\right)}.$$

Lemma 7.5. *For every fixed $i \in [n]$, it holds that*

$$\mathop{\mathbb{E}}_{p,I,Z}\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}}\left[\Pr[L \mid \tilde{X}, i, p, I, Z]\right] \leqslant \Pr\left[\mathrm{Lose}_i|E\right] + \frac{4\,|\mathcal{A}|}{\alpha}\cdot\left(\frac{k}{\delta T} + \epsilon\right).$$

Assuming these, the lemma follows easily:

Proof of Lemma 7.3. We can write the losing probability as

$$\begin{aligned}
\Pr[L] &= \mathop{\mathbb{E}}_{\tilde{X}\sim Q}\ \mathop{\mathbb{E}}_{i,p,I,Z}\left[\Pr[L \mid \tilde{X}, i, p, I, Z]\right] \\
&= \mathop{\mathbb{E}}_{i,p,I,Z}\ \mathop{\mathbb{E}}_{\tilde{X}\sim Q}\left[\Pr[L \mid \tilde{X}, i, p, I, Z]\right] \\
&\leqslant \mathop{\mathbb{E}}_{i,p,I,Z}\left\|P_{X_i|E,X_{I'}=Z} - Q\right\|_1 + \mathop{\mathbb{E}}_{i,p,I,Z}\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}}\left[\Pr[L \mid \tilde{X}, i, p, I, Z]\right].
\end{aligned}$$

Now the result follows by Lemma 7.4 and Lemma 7.5. □

We complete our proof by proving Lemma 7.4 in Section 7.2 and Lemma 7.5 in Section 7.3.

7.2 Analysis of the First Term

We wish to analyze the following term, i.e., how much the random restriction I' and event E can affect the marginal distribution of the i-th question:

$$\mathop{\mathbb{E}}_{i,p,I,Z}\left\|P_{X_i|E,X_{I'}=Z} - Q\right\|_1.$$

To analyze this, we first fix a value of p, and show the following using Lemma 3.1.

Lemma 7.6. *Conditioned on any fixed $p \in (0,1]$, it holds that*

$$\mathop{\mathbb{E}}_{i\sim[n]}\ \mathop{\mathbb{E}}_{I\sim_p[n]\setminus\{i\}}\ \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}}\left\|P_{X_i|E,X_{I'}=Z} - Q\right\|_1 \leqslant \sqrt{\frac{2}{pn}\cdot\log_2\left(\frac{1}{\alpha}\right)}.$$

Proof. We have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{i\sim[n]}\mathop{\mathbb{E}}_{I\sim_p[n]\setminus\{i\}}\mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}}\left\|P_{X_i|E,X_{I'}=Z}-Q\right\|_1\\
&\quad=\sum_{i\in[n]}\sum_{I\subseteq[n],I\not\ni i}\frac{1}{n}\cdot p^{|I|}(1-p)^{n-1-|I|}\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I\cup\{i\}}}|E}}\left\|P_{X_i|E,X_{\overline{I\cup\{i\}}}=Z}-Q\right\|_1\\
&\quad=\sum_{i\in[n]}\sum_{I\subseteq[n],I\ni i}\frac{1}{n}\cdot p^{|I|-1}(1-p)^{n-|I|}\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\left\|P_{X_i|E,X_{\overline{I}}=Z}-Q\right\|_1\\
&\quad=\sum_{I\subseteq[n],I\neq\emptyset}\sum_{i\in I}\frac{1}{pn}\cdot p^{|I|}(1-p)^{n-|I|}\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\left\|P_{X_i|E,X_{\overline{I}}=Z}-Q\right\|_1\\
&\quad=\mathop{\mathbb{E}}_{I\sim_p[n]}\sum_{i\in I}\left[\frac{1}{pn}\cdot\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\left\|P_{X_i|E,X_{\overline{I}}=Z}-Q\right\|_1\right]\\
&\quad=\mathop{\mathbb{E}}_{I\sim_p[n]}\left[\frac{1}{pn}\cdot\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\sum_{i\in I}\left\|P_{X_i|E,X_{\overline{I}}=Z}-Q\right\|_1\right].
\end{aligned}$$

Now, for any fixed I, Z, we can apply Lemma 3.1, and get that the above at most

$$\begin{aligned}
&\mathop{\mathbb{E}}_{I\sim_p[n]}\left[\frac{1}{pn}\cdot\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\sqrt{2\left|I\right|\cdot\log_2\left(\frac{1}{\Pr\left[E|X_{\overline{I}}=Z\right]}\right)}\right]\\
&\qquad=\mathop{\mathbb{E}}_{I\sim_p[n]}\left[\frac{\sqrt{2\left|I\right|}}{pn}\cdot\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\sqrt{\log_2\left(\frac{1}{\Pr\left[E|X_{\overline{I}}=Z\right]}\right)}\right]
\end{aligned}$$

Note that the function $\sqrt{\log_2(\cdot)}$ is concave over $[1,\infty)$; hence, by Jensen's inequality, the above is at most

$$\begin{aligned}
\mathop{\mathbb{E}}_{I\sim_p[n]}\left[\frac{\sqrt{2\left|I\right|}}{pn}\cdot\sqrt{\log_2\left(\mathop{\mathbb{E}}_{Z\sim P_{X_{\overline{I}}|E}}\frac{1}{\Pr\left[E|X_{\overline{I}}=Z\right]}\right)}\right]&\leqslant\mathop{\mathbb{E}}_{I\sim_p[n]}\left[\frac{\sqrt{2\left|I\right|}}{pn}\cdot\sqrt{\log_2\left(\frac{1}{\Pr[E]}\right)}\right]\\
&\leqslant\frac{1}{pn}\sqrt{2\cdot\log_2\left(\frac{1}{\Pr[E]}\right)}\cdot\mathop{\mathbb{E}}_{I\sim_p[n]}\left[\sqrt{|I|}\right].
\end{aligned}$$

Finally, by Cauchy-Schwarz inequality, we get $\mathbb{E}_{I\sim_p[n]}\left[\sqrt{|I|}\right]\leqslant\sqrt{\mathbb{E}_{I\sim_p[n]}|I|}=\sqrt{pn}$, and plugging this above completes the proof. □

With the above, the lemma we wish to prove follows:

Proof of Lemma 7.4. In the randomized strategy for $\mathcal{G}$, the players choose $p\sim\left\{1,\delta,\ldots,\delta^{T-1}\right\}$, and hence $p\geqslant\delta^T$ almost surely. Plugging this into Lemma 7.6 gives the desired result. □

7.3 Analysis of the Second Term

For the remainder of this section, fix an index $i\in[n]$. We want to analyze the following quantity:

$$\mathop{\mathbb{E}}_{p,\ I\sim_p[n]\setminus\{i\}}\mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}}\mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}}\left[\Pr[L\,|\,\tilde{X},i,p,I,Z]\right].$$

Observe that when $Z \sim P_{X_{I'}|E}$ and $\tilde{X} \sim P_{X_i|E,X_{I'}=Z}$, the consistency condition (in the definition of the embedding strategy) holds almost surely for each player $j \in [k]$, and so they answer $\tilde{A}^j \sim P_{A_i^j \,|\, X^j \in E^j,\ X_i^j = \tilde{X}^j,\ X_{I'}^j = Z^j}$.

Hence, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}} \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}} \Big[\Pr[L \,|\, \tilde{X}, i, p, I, Z]\Big] \\
&= \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}} \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{X}^j,\ X_{I'}^j = Z^j\right]\right] \\
&= \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{X}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{X}^j,\ X_{I'}^j = Z^j\right]\right].
\end{aligned}$$

With the above expression in mind, we prove the following lemma:

Lemma 7.7. *For every $\tilde{x} \in \mathcal{X}, \tilde{a} \in \mathcal{A}$ such that $\{x \in \mathcal{X}^{\otimes n} : x_i = \tilde{x}, x \in E\} \neq \emptyset$, it holds that*

$$\begin{aligned}
&\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right]\right] \\
&\qquad \leqslant \Pr\left[A_i = \tilde{a} | X \in E, X_i = \tilde{x}\right] + \frac{4}{\Pr[E|X_i=\tilde{x}]} \cdot \left(\frac{k}{\delta T} + \epsilon\right).
\end{aligned}$$

Before proving this lemma, we show the proof of Lemma 7.5 assuming this.

Proof of Lemma 7.5. Fix any $i \in [n]$; as before, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}} \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}} \Big[\Pr[L \,|\, \tilde{X}, i, p, I, Z]\Big] \\
&= \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E}} \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,X_{I'}=Z}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{X}^j,\ X_{I'}^j = Z^j\right]\right] \\
&= \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{X}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{X}^j,\ X_{I'}^j = Z^j\right]\right] \\
&\leqslant \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \left[\Pr\left[A_i = \tilde{a} | E, X_i = \tilde{x}\right] + \frac{4}{\Pr[E|X_i=\tilde{X}]} \cdot \left(\frac{k}{\delta T} + \epsilon\right)\right] \\
&= \Pr\left[\mathrm{Lose}_i | E\right] + \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \left[\frac{4}{\Pr[E|X_i=\tilde{X}]} \cdot \left(\frac{k}{\delta T} + \epsilon\right)\right] \\
&\leqslant \Pr\left[\mathrm{Lose}_i | E\right] + \sum_{\tilde{a}\in\mathcal{A}} 4 \cdot \left(\frac{k}{\delta T} + \epsilon\right) \cdot \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}} \left[\frac{1}{\Pr[E|X_i=\tilde{X}]}\right] \\
&\leqslant \Pr\left[\mathrm{Lose}_i | E\right] + 4 \cdot \left(\frac{k}{\delta T} + \epsilon\right) \cdot \frac{1}{\Pr[E]} \cdot |\mathcal{A}|.
\end{aligned}$$

□

7.3.1 Approximate Independence Under Random Restriction

In the remaining part of this section, we prove Lemma 7.7. For this, we fix some $\tilde{x} \in \mathcal{X}, \tilde{a} \in \mathcal{A}$ satisfying $\{x \in \mathcal{X}^{\otimes n} : x_i = \tilde{x}, x \in E\} \neq \emptyset$. Recall that we want to analyze:

$$\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right]\right].$$

We now define a set of functions that we want to have small noise stability. We will ensure this by taking random restrictions and using Lemma 4.6. In words, the function F^j is a function on $n-1$ coordinates, specifically $(\mathcal{X}^j)^{\otimes(n-1)}$, and is the indicator function of when the j-th player fills in the i-th coordinate with $\tilde{x}^j$, whether the full vector satisfies the event E^j. Similarly, f^j is the indicator of F^j being true, and player j answering $\tilde{a}^j$ on coordinate i on the full input.

Definition 7.8. *For each $j \in [k]$, define the functions:*

1. *Let $F^j : (\mathcal{X}^j)^{\otimes n-1} \to \{0,1\}$ be the function given by*
$$F^j(x^j_{-i}) = \mathbb{1}[(x^j_{-i}, \tilde{x}^j) \in E^j],$$
where $(x^j_{-i}, \tilde{x}^j)$ is the vector with $\tilde{x}^j$ in the i^{th} coordinate.
2. *Let $f^j : (\mathcal{X}^j)^{\otimes n-1} \to \{0,1\}$, be the function given by:*
$$f^j(x^j_{-i}) = F^j(x^j_{-i}) \cdot \mathbb{1}[\text{Player } j \text{ has } i^{th} \text{ answer } \tilde{a}^j \text{ on input } (x^j_{-i}, \tilde{x}^j) \in \mathcal{X}^{\otimes n} \text{ in } \mathcal{G}^{\otimes n}].$$
Note that this definition relies on the previously fixed strategy for $\mathcal{G}^{\otimes n}$.

Definition 7.9. *(Good random restriction) Let $\Lambda(I, Z)$ be the event that for each $j \in [k]$:*

$$\mathrm{Stab}^{Q^j}_{1-\delta}[(F^j)_{I'\to Z^j} - Q^j((F^j)_{I'\to Z^j})] < \delta,$$
$$\mathrm{Stab}^{Q^j}_{1-\delta}[(f^j)_{I'\to Z^j} - Q^j((f^j)_{I'\to Z^j})] < \delta,$$

where the functions $(F^j)_{I'\to Z^j}$ (resp. $(f^j)_{I'\to Z^j}$) are the restrictions of the functions F^j (resp. f^j) with Z^j plugged into coordinates $I' = ([n]\setminus\{i\})\setminus I$. We used Q^j to denote the marginal of the query distribution Q on the j^{th} player.

Next, we prove some useful lemmas. First, we show that the event $\Lambda(I, Z)$ occurs with high probability.

Lemma 7.10.
$$\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} [\mathbb{1}[\neg\Lambda(I, Z)]] \leqslant \frac{1}{\Pr[E|X_i = \tilde{x}]} \cdot \frac{4k}{\delta T}.$$

Proof. Using Lemma 4.6 (with $\eta = \delta/2$), and by a union bound (over $2k$ functions), we get

$$\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim Q^{\otimes I'}} [\mathbb{1}[\neg\Lambda(I, Z)]] \leqslant \frac{4k}{\delta T}.$$

Now, for every z, it holds that

$$\Pr[X_{I'} = z \,|\, E, X_i = \tilde{x}] \leqslant \frac{\Pr[X_{I'} = z, X_i = \tilde{x}]}{\Pr[E, X_i = \tilde{x}]} = \frac{\Pr[X_{I'} = z]}{\Pr[E|X_i = \tilde{x}]} = \frac{Q^{\otimes I'}[z]}{\Pr[E|X_i = \tilde{x}]}.$$

Hence we have

$$\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{X}}} [\mathbb{1}[\neg\Lambda(I, Z)]] \leqslant \frac{1}{\Pr[E|X_i = \tilde{x}]} \cdot \frac{4k}{\delta T}.$$ □

We observe that under the event $\Lambda(I,Z)$, our functions satisfy an approximate independence property:

Lemma 7.11. *Let I, Z be such that the event $\Lambda(I,Z)$ holds. Then,*

$$\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j, X^j \in E^j | X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right] \leqslant \Pr\left[A_i = \tilde{a},\ X \in E | X_i = \tilde{x},\ X_{I'} = Z\right] + \epsilon,$$

$$\prod_{j=1}^{k} \Pr\left[X^j \in E^j | X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right] \geqslant \Pr\left[X \in E | X_i = \tilde{x},\ X_{I'} = Z\right] - \epsilon.$$

Proof. Let I, Z be such that the event $\Lambda(I,Z)$ holds. By Corollary 4.2 and the definition of the event $\Lambda(I,Z)$, it holds that

$$\prod_{j=1}^{k} \mathop{\mathbb{E}}_{Y^j \sim (Q^j)^{\otimes I}} \left[(f^j)_{I' \to Z^j}(Y^j)\right] \leqslant \mathop{\mathbb{E}}_{Y \sim Q^{\otimes I}} \left[\prod_{j=1}^{k} (f^j)_{I' \to Z^j}(Y^j)\right] + \epsilon,$$

$$\prod_{j=1}^{k} \mathop{\mathbb{E}}_{Y^j \sim (Q^j)^{\otimes I}} \left[(F^j)_{I' \to Z^j}(Y^j)\right] \geqslant \mathop{\mathbb{E}}_{Y \sim Q^{\otimes I}} \left[\prod_{j=1}^{k} (F^j)_{I' \to Z^j}(Y^j)\right] - \epsilon.$$

Now, by the definitions of the functions $(f^j)_{j\in[k]},\ (F^j)_{j\in[k]}$, we have

$$\mathop{\mathbb{E}}_{Y \sim Q^{\otimes I}} \left[\prod_{j=1}^{k} (f^j)_{I' \to Z^j}(Y^j)\right] = \Pr\left[A_i = \tilde{a},\ X \in E \,|\, X_i = \tilde{x},\ X_{I'} = Z\right],$$

$$\mathop{\mathbb{E}}_{Y \sim Q^{\otimes I}} \left[\prod_{j=1}^{k} (F^j)_{I' \to Z^j}(Y^j)\right] = \Pr\left[X \in E \,|\, X_i = \tilde{x},\ X_{I'} = Z\right],$$

and for every $j \in [k]$,

$$\mathop{\mathbb{E}}_{Y^j \sim (Q^j)^{\otimes I}} \left[(f^j)_{I' \to Z^j}(Y^j)\right] = \Pr\left[A_i^j = \tilde{a}^j, X^j \in E^j | X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right],$$

$$\mathop{\mathbb{E}}_{Y^j \sim (Q^j)^{\otimes I}} \left[(F^j)_{I' \to Z^j}(Y^j)\right] = \Pr\left[X^j \in E^j | X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right].$$

Plugging these into the above inequalities, we obtain the desired result. □

Now, we are ready to complete the proof:

Proof of Lemma 7.7. We have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right]\right] \\
&\leqslant \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right] \cdot \mathbb{1}\left[\Lambda(I,Z)\right]\right] \\
&\qquad + \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\mathbb{1}\left[\neg\Lambda(I,Z)\right]\right].
\end{aligned}$$

By Lemma 7.10, the second term above is at most $\frac{1}{\Pr[E|X_i=\tilde{x}]} \cdot \frac{4k}{\delta T}$. By Lemma 7.11 and Lemma A.3, we can bound the first term as:

$$\begin{aligned}
&\mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j | X^j \in E^j,\ X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right] \cdot \mathbb{1}\left[\Lambda(I,Z)\right]\right] \\
&= \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\frac{\prod_{j=1}^{k} \Pr\left[A_i^j = \tilde{a}^j,\ X^j \in E^j \mid X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right]}{\prod_{j=1}^{k} \Pr\left[X^j \in E^j \mid X_i^j = \tilde{x}^j,\ X_{I'}^j = Z^j\right]} \cdot \mathbb{1}\left[\Lambda(I,Z)\right]\right] \\
&\leqslant \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\frac{\Pr\left[A_i = \tilde{a},\ X \in E \mid X_i = \tilde{x}, X_{I'} = Z\right] + 4\epsilon}{\Pr\left[X \in E \mid X_i = \tilde{x}, X_{I'} = Z\right]}\right] \\
&= \mathop{\mathbb{E}}_{p,I\sim_p[n]\setminus\{i\}} \mathop{\mathbb{E}}_{Z\sim P_{X_{I'}|E,X_i=\tilde{x}}} \left[\Pr\left[A_i = \tilde{a} \mid X \in E, X_i = \tilde{x}, X_{I'} = Z\right] + \frac{4\epsilon}{\Pr\left[X \in E \mid X_i = \tilde{x}, X_{I'} = Z\right]}\right] \\
&= \Pr\left[A_i = \tilde{a} \mid X \in E,\ X_i = \tilde{x}\right] + \frac{4\epsilon}{\Pr\left[X \in E \mid X_i = \tilde{x}\right]}.
\end{aligned}$$

Combining the two terms completes the proof. □

7.4 Some Remarks

We remark that the same proof leads to even better bounds on parallel repetition, in the cases we know better CSP inverse theorems for games with no-Abelian-embeddings. Formally, the same choice of the parameters $\delta = \log(n)^{-1/3}, T = \lceil 1/\delta^2 \rceil, \epsilon = \epsilon(\delta), \alpha = \sqrt{\epsilon}$ works; here $\epsilon = \epsilon(\delta) \geqslant \delta$ is chosen so as to satisfy the CSP inverse theorem in Corollary 4.2. This leads to the bound

$$\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant \epsilon \left(\frac{1}{\sqrt[3]{\log n}}\right)^{\Omega(1)}.$$

This implies the following bounds:

1. For a connected game: $\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant (\log n)^{-\Omega(1)}$, via [Mos10]. In particular, we obtain this bound for all 2-player games, since they are connected without loss of generality.
2. For a 3-player game with no-Abelian-embeddings: $\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant (\log\log n)^{-\Omega(1)}$, via [BKM23b].
3. As before, for a k-player game with no-Abelian-embeddings: $\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant (\log \cdots \log n)^{-\Omega(1)}$, where number of logarithms is at most $k^{O(k)}$, via [BKLM24b].

8 Pairwise Connected Games with No Marginal Abelian Embeddings

The main result of this section is a parallel repetition theorem for pairwise-connected distributions with no-marginal-Abelian embeddings, i.e., Theorem 1.4. Throughout the remainder of this section, we fix a k-player game $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ with $\mathrm{val}(\mathcal{G}) < 1$, and such that the distribution Q is pairwise-connected with no-marginal-Abelian-embeddings.

The proof done in a series of steps, as follows:

1. In Section 8.1, given any sufficiently large $n \in \mathbb{N}$, a strategy for the game $\mathcal{G}^{\otimes n}$, and a product event $E \subseteq \mathcal{X}^{\otimes n}$, we define generalized random restrictions $\mathcal{R}_i$ on $\mathcal{X}^{\otimes n}$, one for each coordinate $i \in [n]$, that make some relevant functions (corresponding to the set E and the answer functions for this coordinate) product-pseudorandom.

2. In Section 8.2, we show that for any sufficiently large $n \in \mathbb{N}$, a strategy for the game $\mathcal{G}^{\otimes n}$, and any product event E of large measure, it is hard for the players to win a coordinate $i \in [n]$, conditioned on the inputs being drawn from E, under a certain pseudorandomness assumption. Namely, we want that conditioned on the event E, a random restriction $\rho \sim \mathcal{R}_i$ does not give too much information on the inputs to the players in coordinate i. We note that the proof of this part is similar in spirit to the proof of Theorem 7.1.

3. In Section 8.3, we show how to achieve the pseudorandom assumption above via an iterative process. More formally, for any sufficiently large $n \in \mathbb{N}$, a strategy for the game $\mathcal{G}^{\otimes n}$, and any product event E of large measure, we show that there exists a generalized random restriction $\mathcal{R}$ on $\mathcal{X}^{\otimes n}$, such that the assumption above is satisfied with high probability when the inputs to the game are drawn conditioned on the restriction $\rho \sim \mathcal{R}$.

4. In Section 8.4, we combine the results in the two sections above. More formally, for any sufficiently large $n \in \mathbb{N}$, a strategy for the game $\mathcal{G}^{\otimes n}$, and any product event E of large measure, we show that there exists a generalized random restriction $\mathcal{R}$ on $\mathcal{X}^{\otimes n}$, such that the game has many hard coordinates when the inputs are drawn condition a restriction $\rho \sim \mathcal{R}$.

5. Finally, in Section 8.5, we use the result of the above section along with an inductive argument to complete the proof of Theorem 1.4. We note that the proof of this part is similar in spirit to Lemma B.1.

8.1 Pseudorandom Partitions For Each Coordinate

For any sufficiently large $n \in \mathbb{N}$, consider the repeated game $\mathcal{G}^{\otimes n}$, and let $(h_i^j : (\mathcal{X}^j)^{\otimes n} \to \mathcal{A}^j)_{i\in[n],j\in[k]}$ be any fixed strategies for the k players, for each of the n coordinates. Let $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event. We define the following set of functions, which are the same functions in Definition 7.8, except for all possible values of $\tilde{x} \in \mathcal{X}$ and $\tilde{a} \in \mathcal{A}$ (in Definition 7.8 we fixed some $\tilde{x} \in \mathcal{X}$ and $\tilde{a} \in \mathcal{A}$ beforehand).

Definition 8.1. *Consider any $i \in [n]$.*
For each $\tilde{x} \in \mathcal{X}, \tilde{a} \in \mathcal{A}, j \in [k]$, define the following functions:

1. *The function $F^j_{i,\tilde{x}^j} : (\mathcal{X}^j)^{\otimes n-1} \to \{0,1\}$ is given by*

$$F^j_{i,\tilde{x}^j}(x^j_{-i}) = \mathbb{1}[(x^j_{-i}, \tilde{x}^j) \in E^j],$$

where $(x^j_{-i}, \tilde{x}^j)$ is the vector with $\tilde{x}^j$ in the i^{th} coordinate.

2. *The function $f^j_{i,\tilde{x}^j,\tilde{a}^j} : (\mathcal{X}^j)^{\otimes n-1} \to \{0,1\}$ is given by:*

$$f^j_{i,\tilde{x}^j,\tilde{a}^j}(x^j_{-i}) = F^j_{i,\tilde{x}^j}(x^j_{-i}) \cdot \mathbb{1}[h^j_i(x^j_{-i}, \tilde{x}^j) = \tilde{a}^j].$$

Observe that for every $i \in [n]$, the total number of functions in Definition 8.1 is $2k\,|\mathcal{X}|\,|\mathcal{A}|$, which is a constant (depending on the base game $\mathcal{G}$). Thus using Corollary 5.10 we can find a generalized random restriction $\mathcal{R}_i$ depending on $i \in [n]$ to make all the functions defined in Definition 8.1 product pseudorandom.

Lemma 8.2. *Let $0 < \gamma \leqslant 1$ be such that $\frac{1}{\gamma} \leqslant o(\log\log n)^{1/4}$, and let $i \in [n]$.*

Then, there exists a $(\frac{1}{\eta}, \eta)$-generalized random restriction $\mathcal{R}_i$ on $\mathcal{X}^{\otimes n-1}$ (with respect to coordinates $[n] \setminus \{i\}$),[12] for $\eta = n^{-\exp(-1/\gamma^4)}$ such that with probability $1-\gamma$ over $\rho \sim \mathcal{R}_i$, each function in Definition 8.1 (with respect to i) is $(\sqrt{m(\rho)}, \gamma)$-product pseudorandom under the restriction ρ.

The constant in the exp *depends only on parameters of the base game $\mathcal{G}$.*

Proof. This follows by applying Corollary 5.10 to the relevant functions.

Formally, we apply the corollary with the probability space $(\mathcal{X}, Q)$, after extending each function in Definition 8.1 to a function $\mathcal{X}^{\otimes n-1} \to \{0,1\}$; for example, a function corresponding to player j will only depend on the inputs from $(\mathcal{X}^j)^{\otimes n-1}$, and ignore the inputs corresonding to the other players. This ensures that the pseudorandomness condition for this function finally holds with respect to the correct marginal Q^j. □

In later sections, we shall also be interested in knowing how the parallel repetition of a multiplayer game behaves under generalized restrictions. Generalized random restrictions effectively turn a multiplayer game into the same game on less coordinates. Formally, we define the following:

Definition 8.3. *(Multiplayer game under generalized restriction) Let $n \in \mathbb{N}$; consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$, and fix any strategy for the k players in this game. Let $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event.*

Let $\rho = (T_1, T_2, \ldots, T_m, I, z)$ be a generalized restriction $\mathcal{X}^{\otimes n}$ with $m = m(\rho) \leqslant n$ free coordinates. For any $x' \in \mathcal{X}^{\otimes m}$, let $x'^{(\rho)} \in \mathcal{X}^{\otimes n}$ be its relevant extension to $\mathcal{X}^{\otimes n}$; formally, we have $x_i'^{(\rho)} = \begin{cases} x'_j, & i \in T_j, j \in [m] \\ z_i, & i \in I \end{cases}$. Then, we have:

1. *Consider the game $\mathcal{G}^{\otimes n}$, with the inputs to the k players drawn conditioned on E_ρ; this input distribution is the same as $Q^{\otimes m}$, the input distribution of the game $\mathcal{G}^{\otimes m}$.*

2. *Under this identification, we define a restricted event $E' \subseteq \mathcal{X}^{\otimes m}$ for the game $\mathcal{G}^{\otimes m}$ by*

$$E' := \left\{ x' \in \mathcal{X}^{\otimes m} : x'^{(\rho)} \in E \right\}.$$

   This is a product event $E' = E'^1 \times \cdots \times E'^k$ with respect to the k players.

3. *Given the strategy for $\mathcal{G}^{\otimes n}$, we can define a restricted strategy for $\mathcal{G}^{\otimes m}$ as follows: Fix indices $i_1 \in T_1, i_2 \in T_2, \ldots, i_m \in T_m$. Now, on input $x' \in \mathcal{X}^{\otimes m}$, the players extend it to an input $x'^{(\rho)} \in \mathcal{X}^{\otimes n}$ of $\mathcal{G}^{\otimes n}$, and output the answers to coordinates $i_1, i_2, \ldots, i_m$ respectively.*

   Note that the players win the game $\mathcal{G}^{\otimes m}$ on input $x' \in \mathcal{X}^{\otimes m}$ with the above strategy if they win the game $\mathcal{G}^{\otimes n}$ on input $x'^{(\rho)} \in \mathcal{X}^{\otimes n}$.

8.2 Embedding for a Single Copy of the Game

In this subsection we show that under a certain pseudorandomness condition (that the generalized random restrictions $\mathcal{R}_i$ do not change the distribution of X_i too much when conditioning on E, see (4)), an embedding argument shows hardness for coordinates of the game $\mathcal{G}^{\otimes n}$ (while conditioning on E).

[12] We shall also think of $\mathcal{R}_i$ as a generalized random restriction (with the same parameters) on $\mathcal{X}^{\otimes n}$ which always leaves coordinate i untouched.

For some sufficiently large n, consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, P = Q^{\otimes n}, V^{\otimes n})$, and fix any strategy for the k players in this game. Let $X = (X^1, \dots, X^k)$ be the random variable denoting the questions to the k players in the game $\mathcal{G}^{\otimes n}$, and let $A = (A^1, \dots, A^k)$ be the random variable denoting the answers of the players using these strategies. For each $i \in [n]$, let Win_i (resp. Lose_i) be the event that $V(X_i, A_i) = 1$ (resp. $V(X_i, A_i) = 0$); that is, the players win (resp. lose) the i^{th} coordinate of the game.

We introduce some parameters:

1. $\beta \in [0, 1]$ is any parameter satisfying $\beta \leqslant o(1)$.
2. $\gamma \in [0, 1]$ is any real number such that $(\log\log\log n)^{-1} \leqslant \gamma \leqslant (\log\log\log\log\log n)^{-1}$.
3. $\eta = \eta(n, \gamma) = n^{-\exp(-1/\gamma^4)}$ is as in Lemma 8.2 with respect to the parameter γ.
4. $\delta := (\log\log\log\log\log n)^{-1}$.
5. Let $1 \leqslant C \leqslant k^{O(k)}$ be a constant so that Corollary 5.3 with respect to the probability space $(\mathcal{X}, Q)$ holds with C logarithms. Let
$$\epsilon := \frac{1}{\log\log\cdots\log 1/\delta} = \frac{1}{\log\log\cdots\log n},$$
where the number of logarithms is C in the first expression and $C+5$ in the second expression; this is chosen so that (ϵ, δ) satisfy Corollary 5.3.
6. $\alpha := \sqrt{\epsilon} = \frac{1}{(\log\log\cdots\log n)^{1/2}}$, where the number of logarithms is $C + 5$.

These satisfy the following inequalities, which we shall use later:

$$\frac{1}{\gamma} \leqslant o(\log\log n)^{1/4}, \qquad \eta \leqslant 2^{-\sqrt{\log n}}, \qquad \delta \geqslant \gamma, \delta \geqslant \sqrt{\eta}, \qquad \eta, \gamma, \epsilon \leqslant o(\alpha).$$

The main result of this subsection is the following:

Proposition 8.4. *Let the parameters $\alpha, \beta, \gamma, \eta$ be as above. Let $E = E^1 \times \cdots \times E^k \subseteq (\mathcal{X}^k)^{\otimes n} = \mathcal{X}^{\otimes n}$ be a product event with $\Pr_{Q^{\otimes n}}[E] \geqslant \alpha$.*

Consider any $i \in [n]$. Let $\mathcal{R}_i$ be the $(\frac{1}{\eta}, \eta)$-generalized random restriction as in Lemma 8.2 (with respect to E, the fixed player strategies, and the parameter $\gamma \geqslant \omega(\log\log n)^{-1/4}$; and $\eta = n^{-\exp(-1/\gamma^4)}$), and suppose that it satisfies

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \left\| P_{X_i | E, E_\rho} - Q \right\|_1 \leqslant \beta \leqslant o(1).^{13} \tag{4}$$

Then, it holds that

$$\Pr\left[\mathrm{Win}_i \,|\, E\right] \leqslant \mathrm{val}(\mathcal{G}) + o_n(1).$$

For the remainder of the section, we fix some $i \in [n]$ and prove the above lemma. We construct the following randomized strategy for the game $\mathcal{G}$:

1. The verifier samples $\tilde{X} \sim Q$, and for each $j \in [k]$, gives player j the input $\tilde{X}^j$.
2. Using shared randomness, the players sample $\rho \sim \mathcal{R}_i | E$.

[13] The conditional random restriction $\mathcal{R}_i | E$ is as in Definition 5.11; this is well-defined since $\eta < \alpha$.

3. For each $j \in [k]$, player j does the following:

   We say that $\tilde{X}^j, \rho$ are consistent with E^j if

   $$\left\{ x^j \in (\mathcal{X}^j)^{\otimes n} : x^j \in E^j \cap E^j_\rho,\ x^j_i = \tilde{X}^j \right\} \neq \emptyset,$$

   where $E^j_\rho \subseteq (\mathcal{X}^j)^{\otimes n}$ is the projection of the set E_ρ on player j.[14]

   (a) If the above consistency condition does not hold, output an arbitrary answer from $\mathcal{A}^j$; for example, we may assume the output is the first element of $\mathcal{A}^j$ under some ordering.

   (b) Else, using private randomness, output $\tilde{A}^j \sim P_{A^j_i \,|\, X^j \in E^j \cap E^j_\rho,\ X^j_i = \tilde{X}^j}$.

4. Let $\tilde{A} = (\tilde{A}^1, \tilde{A}^2, \ldots, \tilde{A}^k)$; the players win if and only if $V(\tilde{X}, \tilde{A}) = 1$.

Let L be the event that $V(\tilde{X}, \tilde{A}) = 0$; that is, the players lose the game $\mathcal{G}$ when using the above strategy. We prove that:

Lemma 8.5. *Under the hypotheses of Proposition 8.4, it holds that*

$$\Pr[L] \leqslant \Pr[\mathrm{Lose}_i \,|\, E] + o(1).$$

Assuming this, we immediately have:

Proof of Proposition 8.4. This follows by Lemma 8.5 and the fact that $\Pr[L] \geqslant 1 - \mathrm{val}(\mathcal{G})$. □

Now, we focus on proving Lemma 8.5. First, we show that using our assumption on $\mathcal{R}_i$, it suffices to bound the losing probability assuming the input $\tilde{X}$ came from a different distribution, which is the one conditioned on E, E_ρ. Formally, we show:

Lemma 8.6. *Under the hypotheses of Proposition 8.4, it holds that*

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \mathop{\mathbb{E}}_{\tilde{X} \sim P_{X_i | E, E_\rho}} \left[\Pr[L \,|\, \tilde{X}, \rho] \right] \leqslant \Pr\left[\mathrm{Lose}_i | E\right] + o(1).$$

Assuming these, the lemma follows easily:

Proof of Lemma 8.5. We can write

$$\begin{aligned}
\Pr[L] &= \mathop{\mathbb{E}}_{\tilde{X} \sim Q} \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \left[\Pr[L \,|\, \tilde{X}, \rho] \right] \\
&= \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \mathop{\mathbb{E}}_{\tilde{X} \sim Q} \left[\Pr[L \,|\, \tilde{X}, \rho] \right] \\
&\leqslant \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \left\| P_{X_i | E, E_\rho} - Q \right\|_1 + \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \mathop{\mathbb{E}}_{\tilde{X} \sim P_{X_i | E, E_\rho}} \left[\Pr[L \,|\, \tilde{X}, \rho] \right] \\
&\leqslant \beta + \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \mathop{\mathbb{E}}_{\tilde{X} \sim P_{X_i | E, E_\rho}} \left[\Pr[L \,|\, \tilde{X}, \rho] \right] \\
&\leqslant o(1) + \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \mathop{\mathbb{E}}_{\tilde{X} \sim P_{X_i | E, E_\rho}} \left[\Pr[L \,|\, \tilde{X}, \rho] \right].
\end{aligned}$$

Now the result follows by Lemma 8.6. □

[14] Recall that this simply ensures that certain coordinates (in $[n]$) are equal, and certain coordinates have some fixed value. Note it also holds that $E_\rho = \prod_{j=1}^k E^j_\rho$.

Next, we complete the proof of Lemma 8.6. We want to upper bound the following quantity:

$$\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E}\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,E_\rho}}\left[\Pr[L\,|\,\tilde{X},\rho]\right].$$

To analyze the above, we prove two lemmas.

First, we show that we may assume (upto a small error) that the chosen random restriction $\rho\sim\mathcal{R}_i|E$ is good, in the sense it makes the relevant functions pseudorandom. This is formally defined as:

Definition 8.7. *For a random restriction $\rho\in\mathrm{supp}(\mathcal{R}_i)$, we denote by $\Lambda(\rho)$ the event that all the $2k\,|\mathcal{X}|\,|\mathcal{A}|$ functions in Definition 8.1 with respect to coordinate i are $(\sqrt{m(\rho)},\gamma)$-product pseudorandom.*

Lemma 8.8.

$$\Pr_{\rho\sim\mathcal{R}_i|E}[\neg\Lambda(\rho)]\leqslant\frac{\gamma}{\alpha}+\frac{\eta}{\alpha}\leqslant o(1).$$

Proof. We know by Lemma 8.2 that $\Pr_{\rho\sim\mathcal{R}_i}[\neg\Lambda(\rho)]\leqslant\gamma$. Hence, by Lemma 5.12 we have

$$\begin{aligned}\Pr_{\rho\sim\mathcal{R}_i|E}[\neg\Lambda(\rho)]&\leqslant\sum_\rho\frac{\mathcal{R}[\rho]\cdot\Pr[E|E_\rho]}{\Pr[E]}\cdot\mathbb{1}[\neg\Lambda(\rho)]+\frac{\eta}{\Pr[E]}\\&\leqslant\frac{1}{\Pr[E]}\cdot\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}}[\mathbb{1}[\neg\Lambda(\rho)]]+\frac{\eta}{\alpha}\leqslant\frac{\gamma}{\alpha}+\frac{\eta}{\alpha}\leqslant o(1).\end{aligned}$$

□

Second, we show that the distribution obtained by sampling $\rho\sim\mathcal{R}_i|E$ and $\tilde{X}\sim P_{X_i|E,E_\rho}$ is essentially the same as the distribution obtained by sampling $\tilde{X}\sim P_{X_i|E}$ and $\rho\sim\mathcal{R}_i|(E,X_i=\tilde{X})$. Formally, we have:

Lemma 8.9. *For every $\tilde{x}\in\mathrm{supp}(Q)$, it holds that $\Pr[E,X_i=\tilde{x}]\geqslant\Omega(\alpha)>2\eta$; in particular, this implies the distribution $\mathcal{R}_i|(E,X_i=\tilde{x})$ is well-defined.*

Moreover, we have

$$\sum_{\substack{\rho\in\mathrm{supp}(\mathcal{R}_i),\\ \tilde{x}\in\mathrm{supp}(Q)}}\Big|\,(\mathcal{R}_i|E)[\rho]\cdot\Pr[X_i=\tilde{x}\,|\,E,E_\rho]-\Pr[X_i=\tilde{x}\,|\,E]\cdot(\mathcal{R}_i|E,X_i=\tilde{x})[\rho]\,\Big|\leqslant\frac{2\,|\mathcal{X}|\cdot\eta}{\alpha}\leqslant o(1).$$

Proof. Observe that by Lemma 5.12, we have

$$\left\|\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E}P_{X_i|E,E_\rho}-P_{X_i|E}\right\|_1\leqslant\frac{2\eta}{\alpha}.$$

This implies

$$\left\|P_{X_i|E}-Q\right\|_1\leqslant\frac{2\eta}{\alpha}+\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E}\left\|P_{X_i|E,E_\rho}-Q\right\|\leqslant\frac{2\eta}{\alpha}+\beta\leqslant o(1).$$

In particular, this implies that for every $\tilde{x}\in\mathrm{supp}(Q)$,

$$\Pr[X_i=\tilde{x},E]\geqslant\alpha\cdot\left(\Pr_Q[\tilde{x}]-o(1)\right)=\Omega(\alpha)>2\eta.$$

Hence, the distribution $\mathcal{R}_i|(E,X_i=\tilde{X})$ is well-defined for each $\tilde{x}\in\mathrm{supp}(Q)$.

Now, by Lemma 5.12, we have

$$\begin{aligned}
&\sum_{\substack{\rho\in\mathrm{supp}(\mathcal{R}_i),\\ \tilde{x}\in\mathrm{supp}(Q)}} \Big|\ (\mathcal{R}_i|E)[\rho]\cdot\Pr[X_i=\tilde{x}\,|\,E,E_\rho]-\Pr[X_i=\tilde{x}\,|\,E]\cdot(\mathcal{R}_i|E,X_i=\tilde{x})[\rho]\ \Big|\\
&\leqslant \sum_{\rho,\tilde{x}}\left|\frac{\mathcal{R}_i[\rho]\cdot\Pr[E|E_\rho]}{\Pr[E]}\cdot\Pr[X_i=\tilde{x}\,|\,E,E_\rho]-\Pr[X_i=\tilde{x}\,|\,E]\cdot\frac{\mathcal{R}_i[\rho]\cdot\Pr[E,X_i=\tilde{x}|E_\rho]}{\Pr[E,X_i=\tilde{x}]}\right|\\
&\quad+\sum_{\rho,\tilde{x}}(\mathcal{R}_i|E)[\rho]\cdot\frac{\eta}{\Pr[E]}\cdot\Pr[X_i=\tilde{x}\,|\,E,E_\rho]\\
&\quad+\sum_{\rho,\tilde{x}}(\mathcal{R}_i|E,X_i=\tilde{x})[\rho]\cdot\frac{\eta}{\Pr[E,X_i=\tilde{x}]}\cdot\Pr[X_i=\tilde{x}\,|\,E]\\
&\leqslant 0+\frac{\eta}{\Pr[E]}+\frac{|\mathcal{X}|\cdot\eta}{\Pr[E]}\leqslant\frac{2\,|\mathcal{X}|\cdot\eta}{\alpha}.
\end{aligned}$$

□

With the above two lemmas, we are ready to complete the proof of Lemma 8.6, assuming the following *approximate-independence lemma*, which we shall prove later:[15]

Lemma 8.10. *For every $\tilde{x}\in\mathcal{X}$, $\tilde{a}\in\mathcal{A}$ such that $\{x\in\mathcal{X}^{\otimes n}: x_i=\tilde{x}, x\in E\}\neq\emptyset$, it holds that*

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E,X_i=\tilde{x}}\left[\prod_{j=1}^{k}\Pr\left[A_i^j=\tilde{a}^j|X^j\in E^j\cap E_\rho^j,\ X_i^j=\tilde{x}^j\right]\cdot\mathbb{1}\left[\Lambda(\rho)\right]\right]\\
&\qquad\leqslant\ \Pr\left[A_i=\tilde{a}\,|\,E,\ X_i=\tilde{x}\right]+\frac{2\eta+8\epsilon}{\Pr\left[E,X_i=\tilde{x}\right]}.
\end{aligned}$$

Proof of Lemma 8.6. Observe that when $\rho\sim\mathcal{R}_i|E$ and $\tilde{X}\sim P_{X_i|E,E_\rho}$, the consistency condition (in the definition of the embedding strategy) holds almost surely for each player $j\in[k]$, and so they answer $\tilde{A}^j\sim P_{A_i^j\,|\,X^j\in E^j\cap E_\rho^j,\ X_i^j=\tilde{X}^j}$. This, along with Lemma 8.8, gives that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E}\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,E_\rho}}\Big[\Pr[L\,|\,\tilde{X},\rho]\Big]\\
&=\ \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E}\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,E_\rho}}\ \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0}\left[\prod_{j=1}^{k}\Pr\left[A_i^j=\tilde{a}^j|X^j\in E^j\cap E_\rho^j,\ X_i^j=\tilde{X}^j\right]\right]\\
&\leqslant\ \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E}\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E,E_\rho}}\ \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0}\left[\prod_{j=1}^{k}\Pr\left[A_i^j=\tilde{a}^j|X^j\in E^j\cap E_\rho^j,\ X_i^j=\tilde{X}^j\right]\cdot\mathbb{1}\left[\Lambda(\rho)\right]\right]+o(1).
\end{aligned}$$

Now, by Lemma 8.9, the above is at most

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}}\ \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E,X_i=\tilde{X}}\ \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0}\left[\prod_{j=1}^{k}\Pr\left[A_i^j=\tilde{a}^j|X^j\in E^j\cap E_\rho^j,\ X_i^j=\tilde{X}^j\right]\cdot\mathbb{1}\left[\Lambda(\rho)\right]\right]+o(1)\\
&\leqslant\ \mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}}\ \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0}\ \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i|E,X_i=\tilde{X}}\left[\prod_{j=1}^{k}\Pr\left[A_i^j=\tilde{a}^j|X^j\in E^j\cap E_\rho^j,\ X_i^j=\tilde{X}^j\right]\cdot\mathbb{1}\left[\Lambda(\rho)\right]\right]+o(1)
\end{aligned}$$

[15] Note that the conditional generalized random restriction in this lemma is well-defined by Lemma 8.9.

Using Lemma 8.10, this is at most

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\tilde{X}\sim P_{X_i|E}} \sum_{\tilde{a}\in\mathcal{A}:V(\tilde{X},\tilde{a})=0} \left[\Pr\left[A_i=\tilde{a}\,|\,E,\ X_i=\tilde{x}\right]+\frac{2\eta+8\epsilon}{\Pr[E,X_i=\tilde{X}]}\right]+o(1)\\
&=\Pr\left[\mathrm{Lose}_i|E\right]+\sum_{\tilde{x}\in\mathcal{X}}\sum_{\tilde{a}\in\mathcal{A}:V(\tilde{x},\tilde{a})=0}\Pr[X_i=\tilde{x}\,|\,E]\cdot\frac{2\eta+8\epsilon}{\Pr\left[E,X_i=\tilde{x}\right]}+o(1)\\
&\leqslant\Pr\left[\mathrm{Lose}_i|E\right]+\frac{(2\eta+8\epsilon)\cdot|\mathcal{X}|\cdot|\mathcal{A}|}{\Pr\left[E\right]}+o(1).\\
&\leqslant\Pr\left[\mathrm{Lose}_i|E\right]+o(1).
\end{aligned}$$

In the last inequality, we used $\Pr[E]\geqslant\alpha$, and $\eta,\epsilon\leqslant o(\alpha)$. □

8.2.1 Approximate Independence Under Random Restriction

Finally, in the remainder of this subsection, we prove Lemma 8.10. For this, we fix some $\tilde{x}\in\mathcal{X},\ \tilde{a}\in\mathcal{A}$ satisfying $\{x\in\mathcal{X}^{\otimes n}:x_i=\tilde{x},x\in E\}\neq\emptyset$.

We start by observing that under the event $\Lambda(\rho)$, our functions satisfy an approximate independence property, by the inverse theorem.

Lemma 8.11. *Let $\rho\in\mathrm{supp}(\mathcal{R}_i)$ be such that the event $\Lambda(\rho)$ holds. Then,*

$$\prod_{j=1}^{k}\Pr\left[A_i^j=\tilde{a}^j,X^j\in E^j|X_i^j=\tilde{x}^j,\ X^j\in E_\rho^j\right]\leqslant\Pr\left[A_i=\tilde{a},\ X\in E|X_i=\tilde{x},\ X\in E_\rho\right]+\epsilon,$$

$$\prod_{j=1}^{k}\Pr\left[X^j\in E^j|X_i^j=\tilde{x}^j,\ X^j\in E_\rho^j\right]\geqslant\Pr\left[X\in E|X_i=\tilde{x},\ X\in E_\rho\right]-\epsilon.$$

Proof. Let ρ be such that the event $\Lambda(\rho)$ holds, and let $m=m(\rho)\geqslant n^{\exp(-1/\gamma^4)}$ be the number of free coordinates in ρ. By Corollary 5.3 and the definition of the event $\Lambda(\rho)$, it holds that

$$\prod_{j=1}^{k}\mathop{\mathbb{E}}_{Y^j\sim(Q^j)^{\otimes m}}\left[(f^j_{i,\tilde{x}^j,\tilde{a}^j})_\rho(Y^j)\right]\leqslant\mathop{\mathbb{E}}_{Y\sim Q^{\otimes m}}\left[\prod_{j=1}^{k}(f^j_{i,\tilde{x}^j,\tilde{a}^j})_\rho(Y^j)\right]+\epsilon,$$

$$\prod_{j=1}^{k}\mathop{\mathbb{E}}_{Y^j\sim(Q^j)^{\otimes m}}\left[(F^j_{i,\tilde{x}^j})_\rho(Y^j)\right]\geqslant\mathop{\mathbb{E}}_{Y\sim Q^{\otimes m}}\left[\prod_{j=1}^{k}(F^j_{i,\tilde{x}^j})_\rho(Y^j)\right]-\epsilon.$$

We used that each of the functions is $(\sqrt{m},\gamma)$-product pseudorandom, and hence also $(\delta m,\delta)$-product pseudorandom, since $\sqrt{m}\leqslant\delta m$ and $\gamma\leqslant\delta$.

By the definitions of the functions $(f^j_{i,\tilde{x}^j,\tilde{a}^j})_{j\in[k]},\ (F^j_{i,\tilde{x}^j})_{j\in[k]}$, we have

$$\mathop{\mathbb{E}}_{Y\sim Q^{\otimes m}}\left[\prod_{j=1}^{k}(f^j_{i,\tilde{x}^j,\tilde{a}^j})_\rho(Y^j)\right]=\Pr\left[A_i=\tilde{a},X\in E|X_i=\tilde{x},\ X\in E_\rho\right],$$

$$\mathop{\mathbb{E}}_{Y\sim Q^{\otimes m}}\left[\prod_{j=1}^{k}(F^j_{i,\tilde{x}^j})_\rho(Y^j)\right]=\Pr\left[X\in E|X_i=\tilde{x},\ X\in E_\rho\right],$$

and for every $j \in [k]$,

$$\mathop{\mathbb{E}}_{Y^j \sim (Q^j)^{\otimes m}} \left[(f^j_{i,\tilde{x}^j,\tilde{a}^j})_\rho(Y^j) \right] = \Pr\left[A^j_i = \tilde{a}^j, X^j \in E^j | X^j_i = \tilde{x}^j,\ X^j \in E^j_\rho \right],$$

$$\mathop{\mathbb{E}}_{Y^j \sim (Q^j)^{\otimes m}} \left[(F^j_{i,\tilde{x}^j})_\rho(Y^j) \right] = \Pr\left[X^j \in E^j | X^j_i = \tilde{x}^j,\ X^j \in E^j_\rho \right].$$

Plugging these into the above inequalities, we obtain the desired result. □

Next, we complete the proof:

Proof of Lemma 8.10. We have

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E, X_i = \tilde{x}} \left[\prod_{j=1}^k \Pr\left[A^j_i = \tilde{a}^j | X^j \in E^j \cap E^j_\rho,\ X^j_i = \tilde{x}^j \right] \cdot \mathbb{1}\left[\Lambda(\rho)\right] \right]$$
$$= \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E, X_i = \tilde{x}} \left[\frac{\prod_{j=1}^k \Pr\left[A^j_i = \tilde{a}^j, X^j \in E^j | X^j \in E^j_\rho,\ X^j_i = \tilde{x}^j \right]}{\prod_{j=1}^k \Pr\left[X^j \in E^j | X^j \in E^j_\rho,\ X^j_i = \tilde{x}^j \right]} \cdot \mathbb{1}\left[\Lambda(\rho)\right] \right]$$

By Lemma 8.11, Lemma A.3, and Lemma 5.12 the above is at most

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E, X_i = \tilde{x}} \left[\frac{\Pr\left[A_i = \tilde{a},\ X \in E \mid X_i = \tilde{x}, X \in E_\rho\right] + 4\epsilon}{\Pr\left[X \in E \mid X_i = \tilde{x}, X \in E_\rho\right]} \right]$$
$$= \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E, X_i = \tilde{x}} \left[\Pr\left[A_i = \tilde{a} \mid X_i = \tilde{x}, E, E_\rho\right] + \frac{4\epsilon}{\Pr\left[E \mid X_i = \tilde{x}, E_\rho\right]} \right]$$
$$\leqslant \Pr\left[A_i = \tilde{a} \mid E, X_i = \tilde{x}\right] + \frac{2\eta}{\Pr\left[E, X_i = \tilde{x}\right]} + \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E, X_i = \tilde{x}} \left[\frac{4\epsilon}{\Pr\left[E \mid X_i = \tilde{x}, E_\rho\right]} \right].$$

The last term can now be bounded as

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E, X_i = \tilde{x}} \left[\frac{4\epsilon}{\Pr\left[E \mid X_i = \tilde{x}, E_\rho\right]} \right]$$
$$= \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i} \left[\Pr[E, X_i = \tilde{x} \,|\, E_\rho] \cdot \frac{4\epsilon}{\Pr\left[E \mid X_i = \tilde{x}, E_\rho\right]} \right] \cdot \frac{1}{\mathbb{E}_{\rho' \sim \mathcal{R}_i}\left[\Pr[E, X_i = \tilde{x} \,|\, E_{\rho'}]\right]}$$
$$= 4\epsilon \cdot \frac{\mathbb{E}_{\rho \sim \mathcal{R}_i}\left[\Pr[X_i = \tilde{x} \,|\, E_\rho]\right]}{\mathbb{E}_{\rho' \sim \mathcal{R}_i}\left[\Pr[E, X_i = \tilde{x} \,|\, E_{\rho'}]\right]}$$
$$\leqslant 4\epsilon \cdot \frac{1}{\Pr[E, X_i = \tilde{x}] - \eta}.$$
$$\leqslant 8\epsilon \cdot \frac{1}{\Pr[E, X_i = \tilde{x}]}.$$

We used Lemma 8.9 to say that $\Pr[E, X_i = \tilde{x}] \geqslant \Omega(\alpha) > 2\eta$. □

8.3 Ensuring That Random Restrictions Don't Give Too Much Information

In this subsection, we show how to obtain the pseudorandomness assumption (see (4)) in the above section, via a generalized random restriction. The following lemma says that if some $\mathcal{R}_i$ changes the distribution of X_i conditioned on E by a lot, then in fact conditioning on $\rho \sim \mathcal{R}_i,\ X_i \sim Q$ increases the ℓ_2 energy of E.

Lemma 8.12. *For some sufficiently large n, consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, P = Q^{\otimes n}, V^{\otimes n})$, and let $X = (X^1, \ldots, X^k)$ be the random variable denoting the questions to the k players in the game $\mathcal{G}^{\otimes n}$. Let $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event with $\alpha := \Pr_{Q^{\otimes n}}[E]$.*

Let $i \in [n]$, and let $\mathcal{R}_i$ be any (m, ϵ)-generalized random restriction on $\mathcal{X}^{\otimes n-1}$ (on coordinates $[n] \setminus \{i\}$),[16]*, with $\epsilon < \alpha$, and such that*

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \left\| P_{X_i | E, E_\rho} - Q \right\|_1 \geqslant \beta,$$[17]

for some $0 < \beta \leqslant 1$.

Let $\mathcal{R}$ be the generalized random restriction on $\mathcal{X}^{\otimes n}$ defined as follows: choose $\rho \sim \mathcal{R}_i$, $\tilde{x} \sim Q$; output the generalized restriction ρ' which performs ρ on coordinates $[n] \setminus \{i\}$, and fixes the input value in coordinate i to $\tilde{x}$. Then, it holds:

1. *$\mathcal{R}$ is a (m, ϵ)-generalized random restriction on $\mathcal{X}^{\otimes n}$.*
2. *The conditional mass of E under $E_{\rho'}$ has non-trivially increased variance, i.e.,*

$$\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}} \left[\Pr[E | E_{\rho'}]^2 \right] \geqslant \alpha^2 \left(1 + \beta^2 - \frac{6\epsilon}{\alpha} \right).$$

Before we give the formal proof, we explain why this should hold. Since $\rho \sim \mathcal{R}_i$ does not act on coordinate i, we know that just conditioning on ρ (ignoring E) should not affect the marginal distribution of X_i. However, adding in E does change it. This means that restricting by $\rho \sim \mathcal{R}_i$, $X_i \sim Q$ should split[18] the mass of E in an uneven manner, which then increases the ℓ_2 energy.

Proof. It follows by definitions that $\mathcal{R}$ is a (m, ϵ)-generalized random restriction on $\mathcal{X}^{\otimes n}$.

By Lemma 5.12, we have

$$\begin{aligned}
\beta &\leqslant \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i | E} \left\| P_{X_i | E, E_\rho} - Q \right\|_1 \\
&\leqslant \sum_{\rho} \mathcal{R}_i[\rho] \cdot \frac{\Pr[E | E_\rho]}{\Pr[E]} \cdot \left\| P_{X_i | E, E_\rho} - Q \right\|_1 + \frac{2\epsilon}{\Pr[E]} \\
&= \sum_{\tilde{x}, \rho} \mathcal{R}_i[\rho] \cdot \frac{\Pr[E | E_\rho]}{\alpha} \cdot \left| \Pr[X_i = \tilde{x} \,|\, E, E_\rho] - Q[\tilde{x}] \right| + \frac{2\epsilon}{\alpha} \\
&= \frac{1}{\alpha} \cdot \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i} \mathop{\mathbb{E}}_{\tilde{x} \sim Q} \left| \frac{\Pr[E, X_i = \tilde{x} | E_\rho]}{Q[\tilde{x}]} - \Pr[E | E_\rho] \right| + \frac{2\epsilon}{\alpha}.
\end{aligned}$$

Since ρ only acts on coordinates $[n] \setminus \{i\}$, it holds that $\Pr[E, X_i = \tilde{x} | E_\rho] = \Pr[E | E_\rho, X_i = \tilde{x}] \cdot Q[\tilde{x}]$, and hence the above gives

$$\beta\alpha - 2\epsilon \leqslant \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_i} \mathop{\mathbb{E}}_{\tilde{x} \sim Q} \left| \Pr[E | E_\rho, X_i = \tilde{x}] - \Pr[E | E_\rho] \right|.$$

[16] We shall also regard $\mathcal{R}_i$ as a random restriction on $\mathcal{X}^{\otimes n}$ which does nothing to coordinate i.

[17] Note that the conditional random restriction $\mathcal{R}_i | E$ is well-defined when $\epsilon < \alpha$.

[18] The term *split* is justified by Property 2 of Definition 5.7.

Using Cauchy-Schwarz, we get

$$\begin{aligned}
(\beta\alpha - 2\epsilon)^2 &\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i}\mathop{\mathbb{E}}_{\tilde{x}\sim Q}\left|\Pr[E|E_\rho, X_i=\tilde{x}] - \Pr[E|E_\rho]\right|^2 \\
&= \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i}\left[\mathop{\mathbb{E}}_{\tilde{x}\sim Q}\Pr[E|E_\rho, X_i=\tilde{x}]^2 + \Pr[E|E_\rho]^2 - 2\mathop{\mathbb{E}}_{\tilde{x}\sim Q}\Pr[E|E_\rho, X_i=\tilde{x}]\cdot\Pr[E|E_\rho]\right] \\
&= \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i}\left[\mathop{\mathbb{E}}_{\tilde{x}\sim Q}\Pr[E|E_\rho, X_i=\tilde{x}]^2 - \Pr[E|E_\rho]^2\right] \\
&\leqslant \mathop{\mathbb{E}}_{\rho'\sim\mathcal{R}}\left[\Pr[E|E_{\rho'}]^2\right] - \left(\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}_i}\Pr[E|E_\rho]\right)^2 \\
&\leqslant \mathop{\mathbb{E}}_{\rho'\sim\mathcal{R}}\left[\Pr[E|E_{\rho'}]^2\right] - (\alpha-\epsilon)^2\,.
\end{aligned}$$

Rearranging, we get

$$\mathop{\mathbb{E}}_{\rho'\sim\mathcal{R}}\left[\Pr[E|E_{\rho'}]^2\right] \geqslant (\beta\alpha-2\epsilon)^2 + (\alpha-\epsilon)^2 \geqslant \alpha^2(1+\beta^2) - 6\epsilon\alpha. \qquad \square$$

With the above, we state and prove the main result of this subsection. This follows by repeatedly iterating Lemma 8.12 while some $\mathcal{R}_i$ violates (4).

Lemma 8.13. *Let $n\in\mathbb{N}$ be sufficiently large. Let $\gamma\in(0,1)$ be such that $\gamma \geqslant \frac{1}{\log\log\log\log n}$; let $\alpha,\beta,\kappa\in(0,1)$ be parameters, such that $\kappa,\alpha,\beta\geqslant\gamma$.*

Consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$, and fix any strategy for the k players in this game. Let $E = E^1\times\cdots\times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event with $\Pr_{Q^{\otimes n}}[E]\geqslant\alpha$. Then, there exists a $(n^{\exp(-1/\gamma^{10})}, n^{-\exp(-1/\gamma^{10})})$-generalized random restriction $\mathcal{R}$ on $\mathcal{X}^{\otimes n}$, such that with probability at least $1-\kappa$ over $\rho\sim\mathcal{R}|E$,[19] it holds that:

Let $m = m(\rho)$ be the number of free coordinates in ρ. Consider the game $\mathcal{G}^{\otimes n}$ when the inputs to the k-players are conditioned to be in the set E_ρ; this is the same as $\mathcal{G}^{\otimes m} = (\mathcal{X}^{\otimes m}, \mathcal{A}^{\otimes m}, P = Q^{\otimes m}, V^{\otimes m})$. Let $E'\subseteq\mathcal{X}^{\otimes m}$ be the corresponding restricted event, and also consider restricted strategies for the game $\mathcal{G}^{\otimes m}$, as in Definition 8.3. Let X be the random variable denoting the questions to the k players in this game $\mathcal{G}^{\otimes m}$. For every coordinate $i\in[m]$ in this game, let $\mathcal{R}_i$ be the $(\frac{1}{\eta},\eta)$-generalized random restriction as in Lemma 8.2 (with respect to E', the above player strategies, and the parameter γ;[20] and $\eta = \eta(m,\gamma) = m^{-\exp(-1/\gamma^4)}$). Then, it holds that

1. $\Pr_{Q^{\otimes m}}[E'] = \Pr_{Q^{\otimes n}}[E|E_\rho] \geqslant \frac{\kappa}{4}\cdot\Pr_{Q^{\otimes n}}[E]$.
2. *For every $i\in[m]$ it holds:*

$$\mathop{\mathbb{E}}_{\rho'\sim\mathcal{R}_i|E'}\left\|P_{X_i|E',E_{\rho'}} - Q\right\|_1 \leqslant \beta.^{21} \tag{5}$$

Proof. The proof proceeds via an iterative argument; we start with the generalized random restriction $\mathcal{R}^{(0)}$ on $\mathcal{X}^{\otimes n}$ that does nothing, and in each step refine it. We shall use a progress measure defined as follows: for any generalized random restriction $\mathcal{R}$ on $\mathcal{X}^{\otimes n}$, define

$$\mathcal{Z}(\mathcal{R}) := \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}}\left[\Pr\left[E\,|\,E_\rho\right]^2\right].$$

[19] This conditional random restriction is well defined as $n^{-\exp(-1/\gamma^{10})} < 2^{-\sqrt{\log n}} < \gamma \leqslant \alpha \leqslant \Pr[E]$.

[20] Note that this is well-defined since $m\geqslant n^{-\exp(-1/\gamma^{10})}\geqslant 2^{\sqrt{\log n}}$, and hence $1/\gamma\leqslant o(\log\log m)^{1/4}$.

[21] This conditional random restriction is well-defined as $\eta\leqslant 2^{-\sqrt{\log n}}$ and $\Pr[E']\geqslant\frac{\kappa}{4}\cdot\Pr[E]\geqslant\frac{\gamma^2}{4} > \eta$.

This satisfies $\mathcal{Z}\left(\mathcal{R}^{(0)}\right) = \Pr[E]^2 \geqslant \alpha^2$.

Let $T = \left\lceil \frac{8}{\gamma^6} \right\rceil$; for $t = 1, 2, \ldots, T$, we define the random restriction $\mathcal{R}^{(t)}$ in the following manner:

1. Choose $\rho \sim \mathcal{R}^{(t-1)}$. Let $m = m(\rho)$ be the number of free coordinates in ρ, let $\eta = \eta(m(\rho), \gamma) = m^{-\exp(-1/\gamma^4)}$ be as in Lemma 8.2, let $E' \subseteq \mathcal{X}^{\otimes m}$ be the restriction of E corresponding to ρ, and also consider restricted strategies for the game $\mathcal{G}^{\otimes m}$ as in Definition 8.3.

2. We say that ρ is bad if $\Pr[E'] = \Pr[E|E_\rho] \geqslant \frac{\kappa}{4} \cdot \Pr[E]$, and Equation 5 fails for some coordinate $i \in [m]$; else, we say it is good.

   (a) If ρ is good, we do nothing and output ρ.

   (b) Otherwise, $\Pr[E'] \geqslant \frac{\kappa}{4} \cdot \Pr[E]$ and there exists $i \in [m]$ such that Equation 5 fails. Now, we apply Lemma 8.12 with respect to $\mathcal{R}_i$ to find a relevant $(\frac{1}{\eta}, \eta)$-generalized random restriction $\mathcal{R}_\rho$ on $\mathcal{X}^{\otimes m}$. Choose $\rho' \sim \mathcal{R}_\rho$ and output $\rho' \circ \rho$.

By induction, it is verified that for each $t = 0, 1, \ldots, T$, the random restriction $\mathcal{R}^{(t)}$ is a $(m^{(t)}, \epsilon^{(t)})$-generalized random restriction on Σ^n, with

$$m^{(t)} = \frac{1}{\eta(m^{(t-1)}, \gamma)} = \left(m^{(t-1)}\right)^{\exp(-1/\gamma^4)} = n^{\exp(-t/\gamma^4)} \geqslant n^{\exp(-1/\gamma^{10})} \geqslant 2^{\sqrt{\log n}},$$

and

$$\epsilon^{(t)} = \epsilon^{(t-1)} + \left(m^{(t-1)}\right)^{-\exp(-1/\gamma^4)} \leqslant t \cdot n^{-\exp(-t/\gamma^4)} \leqslant n^{-\exp(-1/\gamma^{10})} \leqslant 2^{-\sqrt{\log n}} := \epsilon.$$

Note that this satisfies $\frac{1}{\gamma^4} \leqslant o(\log\log m^{(T)})^{1/4} \leqslant o(\log\log m(\rho))^{1/4}$ whenever Lemma 8.2 is applied, as needed in the assumption of the lemma. Also, $\epsilon < \alpha$, so the distribution $\mathcal{R}^{(t)}|E$ is well-defined for every $t = 0, 1, \ldots, T$.

Now, suppose that at some point we found some ρ that is bad, and applied Step 2b above to get $\mathcal{R}_\rho$; then, for $\eta = \eta(m(\rho), \gamma)$, by Lemma 8.12, we have

$$\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_\rho} \left[\Pr[E'|E_{\rho'}]^2\right] \geqslant \Pr[E']^2 \left(1 + \beta^2 - \frac{6\eta}{\Pr[E']}\right).$$

Since $\Pr[E'] \geqslant \frac{\kappa}{4} \cdot \Pr[E] \geqslant \frac{\gamma^2}{4}$ (as ρ is bad), and $\eta \leqslant \epsilon = 2^{-\sqrt{\log n}}$, we have $\frac{6\eta}{\Pr[E']} \leqslant \frac{24\epsilon}{\gamma^2} \leqslant \frac{\gamma^2}{2} \leqslant \frac{\beta^2}{2}$, and

$$\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_\rho} \left[\Pr[E'|E_{\rho'}]^2\right] \geqslant \Pr[E']^2 \left(1 + \frac{\beta^2}{2}\right).$$

This implies that for any $t = 1, \ldots, T$, we have

$$\begin{aligned}
\mathcal{Z}(\mathcal{R}^{(t)}) &= \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}^{(t)}} \left[\Pr[E|E_\rho]^2\right] \\
&\geqslant \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}^{(t-1)}} \left[\Pr[E|E_\rho]^2 \cdot \left(1 + \frac{\beta^2}{2} \cdot \mathbb{1}\left[\rho \text{ is bad}\right]\right)\right] \\
&= \mathcal{Z}(\mathcal{R}^{(t-1)}) + \frac{\beta^2}{2} \cdot \mathop{\mathbb{E}}_{\rho \sim \mathcal{R}^{(t-1)}} \left[\Pr[E|E_\rho]^2 \cdot \mathbb{1}\left[\rho \text{ is bad}\right]\right]
\end{aligned}$$

Suppose, for the sake of contradiction, that $\mathcal{R}^{(t)}$ does not satisfy the statement of the lemma, for any $t = 0, 1, \ldots, T$. By Lemma 5.12, we get

$$\begin{aligned}
\kappa &\leqslant \Pr_{\rho\sim\mathcal{R}^{(t)}|E}\left[\Pr[E|E_\rho] < \frac{\kappa}{4}\cdot\Pr[E] \text{ or } \rho \text{ is bad}\right] \\
&\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}^{(t)}}\left[\frac{\Pr[E|E_\rho]}{\Pr[E]}\cdot\left(\mathbb{1}\left[\Pr[E|E_\rho] < \frac{\kappa}{4}\cdot\Pr[E]\right] + \mathbb{1}\left[\rho \text{ is bad}\right]\right)\right] + \frac{\epsilon}{\Pr[E]} \\
&\leqslant \frac{\frac{\kappa}{4}\cdot\Pr[E]}{\Pr[E]} + \frac{1}{\Pr[E]}\cdot\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}^{(t)}}\left[\Pr[E|E_\rho]\cdot\mathbb{1}\left[\rho \text{ is bad}\right]\right] + \frac{\epsilon}{\Pr[E]} \\
&\leqslant \frac{1}{\alpha}\cdot\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}^{(t)}}\left[\Pr[E|E_\rho]\cdot\mathbb{1}\left[\rho \text{ is bad}\right]\right] + \frac{\kappa}{4} + \frac{\epsilon}{\alpha} \\
&\leqslant \frac{1}{\alpha}\cdot\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}^{(t)}}\left[\Pr[E|E_\rho]\cdot\mathbb{1}\left[\rho \text{ is bad}\right]\right] + \frac{\kappa}{2}
\end{aligned}$$

In the last inequality, we used $\kappa, \alpha \geqslant \gamma$ and $\epsilon \leqslant 2^{-\sqrt{\log n}}$. By Cauchy-Schwarz, this implies

$$\mathop{\mathbb{E}}_{\rho\sim\mathcal{R}^{(t)}}\left[\Pr[E|E_\rho]^2\cdot\mathbb{1}[\rho \text{ is bad}]\right] \geqslant \frac{\alpha^2\kappa^2}{4}.$$

Plugging into the above, we get that for any $t = 1, \ldots, T$ it holds that

$$\mathcal{Z}(\mathcal{R}^{(t)}) \geqslant \mathcal{Z}(\mathcal{R}^{(t-1)}) + \frac{\alpha^2\beta^2\kappa^2}{8} \geqslant \mathcal{Z}(\mathcal{R}^{(0)}) + \frac{\alpha^2\beta^2\kappa^2 t}{8} \geqslant \alpha^2 + \frac{\alpha^2\beta^2\kappa^2 t}{8} \geqslant \alpha^2 + \frac{\gamma^6 t}{8}.$$

This is a contradiction for $t = T$. Hence, for some t, the generalized random restriction $\mathcal{R}^{(t)}$ satisfies the statement of the lemma. □

8.4 Combining Together: Hard Coordinates under Random Restrictions

We combine the results in the above sections, and prove the following lemma. It shows that in the game $\mathcal{G}^{\otimes n}$, conditioned on a large product event E, we can find hard coordinates after a generalized random restriction.

Lemma 8.14. *Let $1 \leqslant C \leqslant k^{O(k)}$ be a constant so that Corollary 5.3 with respect to the probability space $(\mathcal{X}, Q)$ holds with C logarithms. Let $n \in \mathbb{N}$ be sufficiently large; let $\alpha, \gamma \in (0,1)$ be such that $(\log\log\log\log n)^{-1} \leqslant \gamma \leqslant (\log\log\log\log\log n)^{-1}$, and $\alpha \geqslant \frac{1}{\log\log\cdots\log n}$ where the number of logarithms is $C+6$. Let $\kappa \in (0,1)$ be such that $\kappa \geqslant \alpha$.*

Consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$, and fix any strategy for the k players in this game. Let $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event with $\Pr_{Q^{\otimes n}}[E] \geqslant \alpha$. Then, there exists a $\left(n', \frac{1}{n'}\right)$-generalized random restriction $\mathcal{R}$ on $\mathcal{X}^{\otimes n}$, with $n' \geqslant n^{\exp(-1/\gamma^{10})}$, such that with probability at least $1-\kappa$ over $\rho \sim \mathcal{R}|E$,[22] it holds that:

Let $m = m(\rho)$ be the number of free coordinates in ρ. Consider the game $\mathcal{G}^{\otimes n}$ when the inputs to the k-players are conditioned to be in the set E_ρ; this is the same as $\mathcal{G}^{\otimes m} = (\mathcal{X}^{\otimes m}, \mathcal{A}^{\otimes m}, P = Q^{\otimes m}, V^{\otimes m})$. Let $E' \subseteq \mathcal{X}^{\otimes m}$ be the corresponding restricted event, and also consider restricted strategies for the game $\mathcal{G}^{\otimes m}$, as in Definition 8.3. Let X be the random variable denoting the questions to the k players in this game $\mathcal{G}^{\otimes m}$, and let A be the random variable denoting their answers with respect to the above strategies. For each $i \in [m]$, let Win_i be the event $V(X_i, A_i) = 1$, denoting that the players win coordinate i of the game. Then, we have:

[22] This conditional random restriction is well-defined as $\frac{1}{n'} < \alpha$.

1. $\Pr_{Q^{\otimes m}}[E'] = \Pr_{Q^{\otimes n}}[E|E_\rho] \geqslant \frac{\kappa}{4} \cdot \Pr_{Q^{\otimes n}}[E]$.
2. *For each* $i \in [m]$,
$$\Pr_{Q^{\otimes m}}\left[\mathrm{Win}_i|E'\right] \leqslant \mathrm{val}(\mathcal{G}) + o(1) \leqslant 1 - \Omega(1).$$

Proof. Let $\mathcal{R}$ be the $(n', \frac{1}{n'})$-generalized random restriction, for $n' = n^{\exp(-1/\gamma^{10})}$, as in Lemma 8.13, with the choice $\beta = \gamma$;[23] note that the assumption $\alpha, \kappa \geqslant \gamma$ is satisfied. This satisfies that with probability $1 - \kappa$ over $\rho \sim \mathcal{R}|E$:

Let $m := m(\rho)$, and let E' and player strategies for $\mathcal{G}^{\otimes m}$ be defined as in the lemma statement. Then, we know

1. $\Pr[E'] = \Pr[E|E_\rho] \geqslant \frac{\kappa}{4} \cdot \Pr[E]$.
2. For each $i \in [m]$, let $\mathcal{R}_i$ be the $(\frac{1}{\eta}, \eta)$-generalized random restriction as in Lemma 8.2 (with respect to E', the restricted player strategies, and the parameter γ; and $\eta = \eta(m, \gamma) = m^{-\exp(-1/\gamma^4)}$). Then,
$$\mathop{\mathbb{E}}_{\rho' \sim \mathcal{R}_i|E'} \left\| P_{X_i|E', E_{\rho'}} - Q \right\|_1 \leqslant \beta.$$

Now, the result will follow by applying Proposition 8.4 on the game $\mathcal{G}^{\otimes m}$, with respect to the event E'. We verify that the parameter assumptions in Section 8.2 hold, as follows:

1. C is the constant for Corollary 5.3 with respect to the space $(\mathcal{X}, Q)$.
2. $\beta \leqslant o(1)$ holds as $\beta = \gamma$, and $m \to \infty$ as $n \to \infty$.
3. We have $(\log\log\log\log n)^{-1} \leqslant \gamma \leqslant (\log\log\log\log\log n)^{-1}$, and $n \geqslant m \geqslant n^{\exp(-1/\gamma^{10})} \geqslant 2^{\sqrt{\log n}}$. Hence, it holds that $(\log\log\log m)^{-1} \leqslant \gamma \leqslant (\log\log\log\log\log m)^{-1}$.
4. $\eta = \eta(m, \gamma) = m^{-\exp(-1/\gamma^4)}$ is as in Lemma 8.2, as required.
5. We know $\alpha \geqslant \frac{1}{\log\log\cdots\log n}$ where the number of logarithms is $C + 6$, and $\Pr[E] \geqslant \alpha$. Hence,
$$\Pr[E'] \geqslant \frac{\kappa}{4} \cdot \Pr[E] \geqslant \frac{\alpha^2}{4} \geqslant \frac{1}{4 \cdot (\underbrace{\log\log\cdots\log}_{C+6} n)^2} \geqslant \frac{1}{(\underbrace{\log\log\cdots\log}_{C+5} m)^{1/2}}.$$ □

8.5 Final Induction

In this section, we finally prove the main theorem via an inductive argument. The goal is to prove that for sufficiently large $N \in \mathbb{N}$, it holds that $\mathrm{val}(\mathcal{G}^{\otimes N}) \leqslant \frac{1}{\underbrace{\log\log\cdots\log}_{O(1)} N}$. For this purpose, we shall fix some large enough $N \in \mathbb{N}$ for the remainder of this section. Also, we fix the following parameters:

$$\alpha := \frac{1}{\underbrace{\log\log\cdots\log}_{C+7} N}, \quad \gamma := \frac{1}{\log\log\log\log\log N}, \quad c = \frac{1 - \mathrm{val}(\mathcal{G})}{2} \geqslant \Omega(1),$$

where $1 \leqslant C \leqslant k^{O(k)}$ is a constant so that Corollary 5.3 with respect to the space $(\mathcal{X}, Q)$ holds with C logarithms. We make the following definition:

[23] It is fine to choose β to be any anything that satisfies $\gamma \leqslant \beta \leqslant o(1)$.

Definition 8.15. *For integer $1 \leqslant n \leqslant N$, and $\alpha \in [0,1]$, we define the quantity* $\mathrm{val}(n,\alpha)$ *as*

$$\mathrm{val}(n,\alpha) := \max_{n \leqslant n' \leqslant N} \max_{E} \mathrm{val}\left(\mathcal{G}^{\otimes n'} | E\right),$$

where the second maximum is over product events $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n'}$ satisfying $\Pr_{Q^{\otimes n'}}[E] \geqslant \alpha$, and $\mathcal{G}^{\otimes n'}|E := (\mathcal{X}^{\otimes n'}, \mathcal{A}^{\otimes n'}, Q^{\otimes n'}|E, V^{\otimes n'})$ is the game $\mathcal{G}^{\otimes n'}$ with the questions to the k players drawn conditioned on E.

Note that by definition it holds that $\mathrm{val}(\mathcal{G}^{\otimes N}) = \mathrm{val}(N,1)$.

We start by proving a simple lemma, which was (implicitly) used in Lemma B.1.

Lemma 8.16. *For any integer $n \leqslant N$, consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$, and fix any strategy for the k players in this game. For each $i \in [n]$, let Win_i be the event that this strategy wins the i^{th} coordinate of the game. Let $\mathrm{Win} = \wedge_{i=1}^{n} \mathrm{Win}_i$.*

Let $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event, and let $i \in [n]$. Then, for any $\theta \in [0,1]$, we have

$$\Pr[\mathrm{Win} \,|\, E] \leqslant \Pr[\mathrm{Win}_i \,|\, E] \cdot \mathrm{val}(n, \theta \Pr[E]) + |\mathcal{X}|\,|\mathcal{A}|\,\theta.$$

Proof. Let X be the random variable denoting the questions to the players in $\mathcal{G}^{\otimes n}$, and let A be the random variable denoting their answers with respect to the fixed strategy.

Fix some $i \in [n]$, and let $Z = (X_i, A_i)$ be the random variable denoting the tuple of questions and answers in coordinate i. Let $\mathcal{T} = \{z \in \mathcal{X} \times \mathcal{A} : V(z) = 1\}$ be the set of winning question and answer pairs in the base game $\mathcal{G}$, and let $\mathcal{T}' \subseteq \mathcal{T}$ consist of z such that $\Pr[Z = z \,|\, E] \geqslant \theta$. Then, it holds that

$$\begin{aligned}
\Pr[\mathrm{Win} \,|\, E] &= \Pr[\mathrm{Win} \wedge \mathrm{Win}_i \,|\, E] \\
&= \sum_{z \in \mathcal{T}} \Pr[\mathrm{Win} \wedge (Z = z) \,|\, E] \\
&\leqslant \sum_{z \in \mathcal{T}'} \Pr[\mathrm{Win} \wedge (Z = z) \,|\, E] + |\mathcal{X}|\,|\mathcal{A}|\,\theta \\
&= \sum_{z \in \mathcal{T}'} \Pr[Z = z \,|\, E] \cdot \Pr[\mathrm{Win} \,|\, E, Z = z] + |\mathcal{X}|\,|\mathcal{A}|\,\theta \\
&\leqslant \sum_{z \in \mathcal{T}'} \Pr[Z = z \,|\, E] \cdot \mathrm{val}(n,\ \theta \Pr[E]) + |\mathcal{X}|\,|\mathcal{A}|\,\theta \\
&\leqslant \Pr[\mathrm{Win}_i \,|\, E] \cdot \mathrm{val}(n, \theta \Pr[E]) + |\mathcal{X}|\,|\mathcal{A}|\,\theta.
\end{aligned}$$

□

We used that for every $z \in \mathcal{T}'$, it holds that $E, Z = z$ is a product event with measure $\Pr[E, Z = z] = \Pr[E] \cdot \Pr[Z = z|E] \geqslant \Pr[E] \cdot \theta$.

Now, using Lemma 8.14, we prove the following inductive lemma:

Lemma 8.17. *Let $\theta \in (0,1)$ be a paramter satisfying $4\theta^3 \geqslant \alpha$.*

Let $n \in \mathbb{N}$, $\mu \in [0,1]$ be such that $2^{\sqrt{\log N}} \leqslant n \leqslant N$ and $\mu \geqslant \alpha$. Then, it holds that

$$\mathrm{val}(n,\mu) \leqslant (1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})} \right\rceil, \theta^6 \mu\right) + 6\,|\mathcal{X}|\,|\mathcal{A}| \cdot \theta^3.$$

Proof. Let θ be as in the lemma statement. Let $2^{\sqrt{\log N}} \leqslant n \leqslant N$, and $\mu \geqslant \alpha$. Consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$, and let $E = E^1 \times \cdots \times E^k \subseteq \mathcal{X}^{\otimes n}$ be a product event such that

$\Pr_{Q^{\otimes n}}[E] \geqslant \mu$. Fix any strategy for the k players for the game $\mathcal{G}^{\otimes n}$, and let W be the event that the players win this game. It suffices to show[24] that

$$\Pr[W|E] \leqslant (1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})}\right\rceil, \theta^6\mu\right) + 6\,|\mathcal{X}|\,|\mathcal{A}| \cdot \theta^3.$$

First, we apply Lemma 8.14 (with $\kappa = 4\theta^3$), to find a $\left(\frac{1}{\eta}, \eta\right)$-generalized random restriction $\mathcal{R}$ on $\mathcal{X}^{\otimes n}$, with $\eta \leqslant n^{-\exp(-1/\gamma^{10})} \leqslant 2^{-(\log N)^{1/4}}$, such that with probability at least $1-4\theta^3$ over $\rho \sim \mathcal{R}|E$: Let $m = m(\rho)$ be the number of free coordinates in ρ. Consider the game $\mathcal{G}^{\otimes n}$ when the inputs to the k-players are conditioned to be in the set E_ρ; this is the same as $\mathcal{G}^{\otimes m} = (\mathcal{X}^{\otimes m}, \mathcal{A}^{\otimes m}, P = Q^{\otimes m}, V^{\otimes m})$. Let $E' \subseteq \mathcal{X}^{\otimes m}$ be the corresponding restricted event, and also consider restricted strategies for the game $\mathcal{G}^{\otimes m}$, as in Definition 8.3. Let X be the random variable denoting the questions to the k players in this game $\mathcal{G}^{\otimes m}$, and let A be the random variable denoting their answers with respect to the above strategies. For each $i \in [m]$, let Win_i be the event $V(X_i, A_i) = 1$ denoting that the players win coordinate i of the game. Then, we have:

1. $\Pr[E'] = \Pr[E|E_\rho] \geqslant \theta^3 \cdot \Pr[E] \geqslant \theta^3\mu$.
2. For each $i \in [m]$, $\Pr_{Q^{\otimes m}}[\mathrm{Win}_i|E'] \leqslant \mathrm{val}(\mathcal{G}) + o(1) \leqslant 1 - c$.

Note that the lemma is applicable by the choice of parameters as:

1. $\kappa = 4\theta^3 \geqslant \alpha$ by the lemma hypothesis.
2. Note that $2^{\sqrt{\log N}} \leqslant n \leqslant N$, and so $\frac{1}{2}\log\log N \leqslant \log\log n \leqslant \log\log N$.
   (a) $\alpha := \frac{1}{\log\log\cdots\log N}$ where the number of logarithms is $C+7$, and hence $\Pr[E] \geqslant \mu \geqslant \alpha \geqslant \frac{1}{\log\log\cdots\log n}$ where the number of logarithms is $C+6$.
   (b) $\gamma = \frac{1}{\log\log\log\log\log N}$ and hence $(\log\log\log\log n)^{-1} \leqslant \gamma \leqslant (\log\log\log\log\log n)^{-1}$.

Now, let Γ be the set of all ρ for which the above properties hold. Consider any $\rho \in \Gamma$ and let $m = m(\rho)$. Denoting $\mathrm{Win} = \wedge_{i=1}^m \mathrm{Win}_i$, we have by Lemma 8.16 (with parameters m, θ^3), that

$$\begin{aligned}
\Pr_{Q^{\otimes m}}[\mathrm{Win}\,|\,E'] &\leqslant \Pr_{Q^{\otimes m}}\left[\mathrm{Win}_i|E'\right] \cdot \mathrm{val}(m,\ \theta^3 \Pr[E']) + |\mathcal{X}|\,|\mathcal{A}|\,\theta^3 \\
&\leqslant (1-c) \cdot \mathrm{val}(m,\ \theta^6\mu) + |\mathcal{X}|\,|\mathcal{A}|\,\theta^3 \\
&\leqslant (1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})}\right\rceil,\ \theta^6\mu\right) + |\mathcal{X}|\,|\mathcal{A}|\,\theta^3
\end{aligned}$$

Finally, by Lemma 5.12, in the game $\mathcal{G}^{\otimes n}$ we have that

$$\begin{aligned}
\Pr_{Q^{\otimes n}}[W|E] &\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}|E}[\Pr[W\,|\,E, E_\rho]] + \frac{2 \cdot 2^{-(\log N)^{1/4}}}{\Pr[E]} \\
&\leqslant \mathop{\mathbb{E}}_{\rho\sim\mathcal{R}|E}[\Pr[W\,|\,E, E_\rho]] + \theta^3 \\
&\leqslant (1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})}\right\rceil,\ \theta^6\mu\right) + |\mathcal{X}|\,|\mathcal{A}|\,\theta^3 + \Pr_{\rho\sim\mathcal{R}|E}[\rho \notin \Gamma] + \theta^3 \\
&\leqslant (1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})}\right\rceil,\ \theta^6\mu\right) + |\mathcal{X}|\,|\mathcal{A}|\,\theta^3 + 4\theta^3 + \theta^3, \\
&\leqslant (1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})}\right\rceil,\ \theta^6\mu\right) + 6\,|\mathcal{X}|\,|\mathcal{A}| \cdot \theta^3.
\end{aligned}$$

[24] Formally, to actually prove the lemma one can apply this claim to an arbitrary n' satisfying $n \leqslant n' \leqslant N$, and $E \subseteq \mathcal{X}^{\otimes n'}$, and then use that $\left\lceil n'^{\exp(-1/\gamma^{10})}\right\rceil \geqslant \left\lceil n^{\exp(-1/\gamma^{10})}\right\rceil$.

We used that for each $\rho \in \Gamma$, the quantity $\Pr[W \mid E, E_\rho]$ is bounded by the quantity $\Pr_{Q^{\otimes m}}[\mathrm{Win}|E']$ analyzed before. □

Next, we complete the proof of the main theorem:

Proof of Theorem 1.4. The proof shall follow by applying Lemma 8.17 iteratively. We define the parameters

$$\theta := \frac{1}{\log\left(\frac{1}{\alpha}\right)} = \frac{1}{\underbrace{\log\log\cdots\log}_{C+8} N}, \quad T = \left\lceil \frac{2}{c} \log\log\left(\frac{1}{\alpha}\right) \right\rceil.$$

Note that this satisfies $4\theta^3 \geqslant \alpha$.

We show by induction that for every $t = 0, 1, 2, \ldots, T$, it holds that

$$\mathrm{val}(\mathcal{G}^{\otimes N}) \leqslant (1-c)^t \cdot \mathrm{val}\left(\left\lceil N^{\exp(-t/\gamma^{10})} \right\rceil,\ \theta^{6t}\right) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3 t.$$

The base case $t = 0$ follows from the observation that $\mathrm{val}(\mathcal{G}^{\otimes N}) = \mathrm{val}(N, 1)$.

For the inductive step, consider any $t = 0, \ldots, T-1$, and assume that the statement holds for t. Let $n = \left\lceil N^{\exp(-t/\gamma^{10})} \right\rceil$, $\mu = \theta^{6t}$; observe that by the choice of parameters, these satisfy $2^{\sqrt{\log N}} \leqslant n \leqslant N$ and $\mu \geqslant \alpha$. By the inductive hypothesis, and Lemma 8.17, we get

$$\begin{aligned}
\mathrm{val}(\mathcal{G}^{\otimes N}) &\leqslant (1-c)^t \cdot \mathrm{val}\left(\left\lceil N^{\exp(-t/\gamma^{10})} \right\rceil,\ \theta^{6t}\right) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3 t \\
&= (1-c)^t \cdot \mathrm{val}(n, \mu) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3 t \\
&\leqslant (1-c)^t \cdot \left((1-c) \cdot \mathrm{val}\left(\left\lceil n^{\exp(-1/\gamma^{10})} \right\rceil,\ \theta^6 \mu\right) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3\right) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3 t \\
&\leqslant (1-c)^{t+1} \cdot \mathrm{val}\left(\mathcal{G}, \left\lceil N^{\exp(-(t+1)/\gamma^{10})} \right\rceil,\ \theta^{6(t+1)}\right) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3 \cdot (t+1),
\end{aligned}$$

as desired.

Finally, applying the above claim for $t = T$, we get

$$\begin{aligned}
\mathrm{val}(\mathcal{G}^{\otimes N}) &\leqslant (1-c)^T \cdot \mathrm{val}\left(\left\lceil N^{\exp(-T/\gamma^{10})} \right\rceil,\ \theta^{6T}\right) + 6\,|\mathcal{X}|\,|\mathcal{A}|\,\theta^3 T \\
&\leqslant (1-c)^T + 6\,|\mathcal{X}|\,|\mathcal{A}| \cdot \theta^3 \cdot T \\
&\leqslant (\log(1/\alpha))^{-2} + 6\,|\mathcal{X}|\,|\mathcal{A}| \cdot (\log(1/\alpha))^{-3} \cdot \frac{4}{c} \log\log\left(\frac{1}{\alpha}\right) \\
&\leqslant (\log(1/\alpha))^{-1} = \frac{1}{\underbrace{\log\log\cdots\log}_{C+8} N}.
\end{aligned}$$

Hence, the theorem holds with the constant $C + 8 \leqslant k^{O(k)}$. □

A Some Useful Lemmas

The following simple lemma appears as [BKLM24a, Lemma 8.4].

Lemma A.1. *Let (Σ, μ) be a finite probability space, and let $k \leqslant n \in \mathbb{N}$. For each $T \subseteq [n]$, let $E_T \subseteq \Sigma^n$ be the event defined as:*

$$E_T = \{x \in \Sigma^n : x_i = x_j \text{ for each } i, j \in T\}.$$

Now, let $S \subseteq [n]$ be any set, and let $1 \leqslant k \leqslant \sqrt{|S|}$ be an integer. Let ν be the distribution on Σ^n obtained sampling $T \subseteq S, |T| = k$ uniformly at random, and then sampling $x \sim \mu^{\otimes n}|E_T$. Then, it holds that

$$\left\|\nu - \mu^{\otimes n}\right\|_1 \leqslant \frac{Ck}{\sqrt{|S|}},$$

where C is a constant depending on μ.

Proof. For a fixed $a \in \Sigma$, let ν' be the distribution obtained sampling $T \subseteq S, |T| = k$, and then sampling x from $\mu^{\otimes n}$ conditioned on $x_i = a$ for each $i \in T$. It suffices to show $\|\nu' - \mu^{\otimes n}\|_1 \leqslant \frac{Ck}{\sqrt{|S|}}$. This is proven by induction on k as follows:

1. For $k = 1$, this follows by direct calculation.
2. For $k > 1$, note that sampling $T \subseteq S$ can be done by sampling $T' \subseteq S$, $|T'| = k - 1$, and then sampling $t \in S \setminus T'$ and setting $T = T' \cup \{t\}$. The result follows by induction. □

Lemma A.2. *Let $x, y \in \mathbb{C}$ be such that $|y| \leqslant 1$ and $|x - y| \leqslant \epsilon$. Then, $|x|^2 \geqslant |y|^2 - 2\epsilon$.*

Proof. We have

$$|x|^2 = |y + (x - y)|^2 = |y|^2 + |x - y|^2 - (y \cdot \overline{(x - y)} - \overline{y} \cdot (x - y)) \geqslant |y|^2 + \epsilon^2 - 2\,|y|\,\epsilon \geqslant |y|^2 - 2\epsilon.$$ □

Lemma A.3. *Let $x, y, a, b, \epsilon \in [0, 1]$ be real numbers such that*

$$b > 0,\ y > 0, \quad a \leqslant b,\ x \leqslant y, \quad a \leqslant x + \epsilon,\ b \geqslant y - \epsilon.$$

Then,

$$\frac{a}{b} \leqslant \frac{x + 4\epsilon}{y}.$$

Proof. If $y \leqslant 2\epsilon$, the statement holds since $\frac{a}{b} \leqslant 1 \leqslant 2 \leqslant \frac{4\epsilon}{y} \leqslant \frac{x+4\epsilon}{y}$. Otherwise, if $y > 2\epsilon$, we have $y - \epsilon \geqslant \frac{y}{2} > 0$, and hence

$$\frac{a}{b} - \frac{x}{y} \leqslant \frac{x + \epsilon}{y - \epsilon} - \frac{x}{y} = \frac{(x + y)\epsilon}{(y - \epsilon)y} \leqslant \frac{2y\epsilon}{\frac{y}{2} \cdot y} = \frac{4\epsilon}{y}.$$ □

Fact A.4. *(Chernoff Bounds, see [MU05] for reference) Let $X_1, \ldots, X_n \in \{0, 1\}$ be independent random variables each with mean μ, and let $X = \sum_{i=1}^n X_i$. Then, for all $\delta \in (0, 1)$, it holds that*

$$\Pr\left[X \leqslant (1 - \delta)\mu n\right] \leqslant e^{-\frac{\delta^2 \mu n}{2}},$$

$$\Pr\left[X \geqslant (1 + \delta)\mu n\right] \leqslant e^{-\frac{\delta^2 \mu n}{3}}.$$

B An Inductive Parallel Repetition Criterion

The following lemma is an inductive parallel repetition criterion, similar to one in [Raz98]. The proof is identical to the proof of Lemma 3.18 in [GHM+22], and is included here for the sake of completeness.

Lemma B.1. *Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, Q, V)$ be a k-player game, and consider its n-folds repetition $\mathcal{G}^{\otimes n} = (\mathcal{X}^{\otimes n}, \mathcal{A}^{\otimes n}, Q^{\otimes n}, V^{\otimes n})$ for some sufficiently large $n \in \mathbb{N}$. Fix an optimal strategy for the k players in this game, and each for $i \in [n]$, let Win_i be the event that this strategy wins the i^{th} coordinate of the game.*

Suppose that $\epsilon, \alpha \in (0, 1]$, $\alpha \geqslant 2^{-n}$ are such that the following condition holds: For every product event $E = E^1 \times \cdots \times E^k \subseteq (\mathcal{X}^1)^{\otimes n} \times \cdots \times (\mathcal{X}^k)^{\otimes n} = \mathcal{X}^{\otimes n}$ with $\Pr_{Q^{\otimes n}}[E] \geqslant \alpha$, there exists a coordinate $i \in [n]$ such that $\Pr[\mathrm{Win}_i \,|\, E] \leqslant 1 - \epsilon$. Then, it holds that

$$\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant \left(1 - \frac{\epsilon}{2}\right)^{\frac{\log_2(1/\alpha)}{2 \cdot \log_2(4|\mathcal{X}||\mathcal{A}|)}}.$$

In particular, if $\epsilon > 0$ is a constant, we get $\mathrm{val}(\mathcal{G}^{\otimes n}) \leqslant \alpha^c$, for some constant $c = c(\mathcal{G}) > 0$.

Proof. Let $X \in \mathcal{X}^{\otimes n}$ be the random variable denoting the questions to the players in $\mathcal{G}^{\otimes n}$, and let $A \in \mathcal{A}^{\otimes n}$ be the random variable denoting their answers with respect to a fixed optimal strategy. Define a sequence of random variables $J_1, \ldots, J_n \in [n]$, and $Z_1, \ldots, Z_n \in \mathcal{X} \times \mathcal{A}$ as follows: for each $i = 1, 2, \ldots, n$, let $J_i \in [n] \setminus \{J_1, \ldots, J_{i-1}\}$ be a coordinate with the lowest winning probability, conditioned on the value of $(Z_1, \ldots, Z_{i-1})$; let $Z_i = (X_{J_i}, A_{J_i})$. For each $i \in [n]$, let W_i denote the event Win_{J_i} of the players winning the coordinate J_i.

Let $s = |\mathcal{X}|\,|\mathcal{A}|$, and $m = \lfloor \frac{\log_2(1/\alpha)}{\log_2(4s)} \rfloor < n$; assume $m \geqslant 2$ or else the theorem holds trivially from the lemma hypothesis by taking $E = \mathcal{X}^{\otimes n}$. We claim that for every integer $k \in [m]$, it holds that $\Pr[W_{\leqslant k}] \leqslant (1 - \epsilon/2)^k$, where $W_{\leqslant k} = W_1 \wedge W_2 \wedge \cdots \wedge W_k$. Then, substituting $k = m$ gives the desired result.

We prove above claim by induction on k. The base case $k = 1$ follows from the lemma hypothesis by taking $E = \mathcal{X}^{\otimes n}$. For the inductive step, consider any $1 \leqslant k \leqslant m - 1$, and suppose that $\Pr[W_{\leqslant k}] \leqslant (1 - \epsilon/2)^k$. Further, we may assume that $\Pr[W_{\leqslant k}] \geqslant 1/2^{k+1}$, or else we already have $\Pr[W_{\leqslant k+1}] \leqslant \Pr[W_{\leqslant k}] \leqslant 2^{-(k+1)} \leqslant (1 - \epsilon/2)^{k+1}$. Now, observe that $W_{\leqslant k}$ depends deterministically on the random variables $Z_{\leqslant k} = (Z_1, \ldots, Z_k)$. Let $\mathcal{T}$ denote the set of all such tuples $z_{\leqslant k}$ satisfying $W_{\leqslant k}$, and let $\mathcal{T}' \subseteq \mathcal{T}$ consist of $z_{\leqslant k} \in \mathcal{T}$ such that $\Pr[Z_{\leqslant k} = z_{\leqslant k}] \geqslant \alpha$; note that by the lemma hypothesis, for each $z_{\leqslant k} \in \mathcal{T}'$, it holds that $\Pr[W_{k+1} \,|\, Z_{\leqslant k} = z_{\leqslant k}] \leqslant 1 - \epsilon$, since $Z_{\leqslant k} = z_{\leqslant k}$ is a product event of measure at least α. Hence,

$$\begin{aligned}
\Pr[W_{k+1} \,|\, W_{\leqslant k}] &= \sum_{z_{\leqslant k} \in \mathcal{T}} \Pr[W_{k+1} \,|\, Z_{\leqslant k} = z_{\leqslant k}] \cdot \frac{\Pr[Z_{\leqslant k} = z_{\leqslant k}]}{\Pr[W_{\leqslant k}]} \\
&\leqslant \sum_{z_{\leqslant k} \in \mathcal{T}'} (1 - \epsilon) \cdot \frac{\Pr[Z_{\leqslant k} = z_{\leqslant k}]}{\Pr[W_{\leqslant k}]} + \sum_{z_{\leqslant k} \in \mathcal{T} \setminus \mathcal{T}'} 1 \cdot \frac{\Pr[Z_{\leqslant k} = z_{\leqslant k}]}{\Pr[W_{\leqslant k}]} \\
&= (1 - \epsilon) + \epsilon \cdot \sum_{z_{\leqslant k} \in \mathcal{T} \setminus \mathcal{T}'} \frac{\Pr[Z_{\leqslant k} = z_{\leqslant k}]}{\Pr[W_{\leqslant k}]} \\
&\leqslant (1 - \epsilon) + \epsilon \cdot \frac{\alpha}{2^{-(k+1)}} \cdot s^k \\
&\leqslant (1 - \epsilon) + \epsilon \cdot \alpha \cdot (2s)^m \leqslant (1 - \epsilon) + \frac{\epsilon}{2} \cdot \alpha \cdot (4s)^m \leqslant 1 - \frac{\epsilon}{2}.
\end{aligned}$$

This implies $\Pr[W_{\leqslant k+1}] = \Pr[W_{k+1} \,|\, W_{\leqslant k}] \cdot \Pr[W_{\leqslant k}] \leqslant (1 - \epsilon/2) \cdot (1 - \epsilon/2)^k = (1 - \epsilon/2)^{k+1}$. □

References

[ABSS97] Sanjeev Arora, László Babai, Jacques Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *J. Comput. Syst. Sci.*, 54(2):317–331, 1997. 3

[AK09] Noga Alon and Bo'az Klartag. Economical toric spines via Cheeger's inequality. *J. Topol. Anal.*, 1(2):101–111, 2009. 3

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. 3

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. 3

[BBCR13] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM J. Comput.*, 42(3):1327–1363, 2013. 3

[BBK+25] Amey Bhangale, Mark Braverman, Subhash Khot, Yang P. Liu, and Dor Minzer. Parallel repetition for 3-player XOR games. In Michal Koucký and Nikhil Bansal, editors, *Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC 2025, Prague, Czechia, June 23-27, 2025*, pages 104–110. ACM, 2025. 1, 4, 7

[BBLV13] Jop Briët, Harry Buhrman, Troy Lee, and Thomas Vidick. Multipartite entanglement in XOR games. *Quantum Inf. Comput.*, 13(3-4):334–360, 2013. 3

[BG15] Mark Braverman and Ankit Garg. Small value parallel repetition for general games. In Rocco A. Servedio and Ronitt Rubinfeld, editors, *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 335–340. ACM, 2015. 3

[BGKW88] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In Janos Simon, editor, *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 113–131. ACM, 1988. 3

[BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, pcps, and nonapproximability-towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998. 3

[BKLM24a] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. On approximability of satisfiable k-CSPs: VI, 2024. Available at https://arxiv.org/pdf/2411.15133. 4, 10, 17, 18, 49

[BKLM24b] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. On approximability of satisfiable k-CSPs: VII, 2024. Available at https://arxiv.org/pdf/2411.15136. 4, 8, 9, 13, 14, 17, 33

[BKLM24c] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. Reasonable bounds for combinatorial lines of length three, 2024. Available at https://arxiv.org/pdf/2411.15137. 4, 10, 18, 20

[BKM23a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: III. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 643–655. ACM, 2023. 8

[BKM23b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-CSPs: II. In *STOC'23—Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 632–642. ACM, New York, [2023] ©2023. 8, 33

[BKM23c] Mark Braverman, Subhash Khot, and Dor Minzer. Parallel repetition for the GHZ game: Exponential decay. In *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2023, Santa Cruz, CA, USA, November 6-9, 2023*, pages 1337–1341. IEEE, 2023. 1, 4, 7

[BKM24] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: IV. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1423–1434. ACM, 2024. 8

[BKM25a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: I. *Comput. Complex.*, 34(2):8, 2025. 8

[BKM25b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: V. In Michal Koucký and Nikhil Bansal, editors, *Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC 2025, Prague, Czechia, June 23-27, 2025*, pages 62–71. ACM, 2025. 8

[BM21] Mark Braverman and Dor Minzer. Optimal tiling of the euclidean space using permutation-symmetric bodies. In Valentine Kabanets, editor, *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPIcs*, pages 5:1–5:48. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 3

[BRR+09] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. Strong parallel repetition theorem for free projection games. In *APPROX-RANDOM*, pages 352–365, 2009. 3

[BRWY13] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 746–755. IEEE Computer Society, 2013. 3

[CHTW04] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 236–249. IEEE Computer Society, 2004. 3

[DGKR05] Irit Dinur, Venkatesan Guruswami, Subhash Khot, and Oded Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. *SIAM J. Comput.*, 34(5):1129–1146, 2005. 3

[DHVY17] Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expanding games. In *ITCS*, pages Art. No. 37, 16, 2017. 1, 3, 4, 6, 16

[DRS05] Irit Dinur, Oded Regev, and Clifford D. Smyth. The hardness of 3-uniform hypergraph coloring. *Comb.*, 25(5):519–535, 2005. 3

[DS14] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 624–633. ACM, 2014. 1, 3

[Fei91] Uriel Feige. On the success probability of the two provers in one-round proof systems. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference, Chicago, Illinois, USA, June 30 - July 3, 1991*, pages 116–123. IEEE Computer Society, 1991. 3

[Fei98] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45(4):634–652, 1998. 3

[FGL+96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996. 3

[FK91] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem. *J. Anal. Math.*, 57:64–119, 1991. 3

[FKO07] Uriel Feige, Guy Kindler, and Ryan O'Donnell. Understanding parallel repetition requires understanding foams. In *22nd Annual IEEE Conference on Computational Complexity (CCC 2007), 13-16 June 2007, San Diego, California, USA*, pages 179–192. IEEE Computer Society, 2007. 3

[For89] Lance Fortnow. *Complexity theoretic aspects of interactive proof systems*. PhD thesis, MIT, 1989. 3

[FRS94] Lance Fortnow, John Rompel, and Michael Sipser. On the power of multi-prover interactive protocols. *Theor. Comput. Sci.*, 134(2):545–557, 1994. 3

[FV02] Uriel Feige and Oleg Verbitsky. Error reduction by parallel repetition—a negative result. *Combinatorica*, 22(4):461–478, 2002. 3, 4

[GHM+21] Uma Girish, Justin Holmgren, Kunal Mittal, Ran Raz, and Wei Zhan. Parallel repetition for the GHZ game: A simpler proof. In *APPROX-RANDOM*, pages 62:1–62:19, 2021. 1, 3, 4

[GHM+22] Uma Girish, Justin Holmgren, Kunal Mittal, Ran Raz, and Wei Zhan. Parallel repetition for all 3-player games over binary alphabet. In *STOC*, pages 998–1009, 2022. 1, 3, 4, 6, 7, 16, 50

[GHS02] Venkatesan Guruswami, Johan Hastad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM J. Comput.*, 31(6):1663–1686, 2002. 3

[GMRZ22] Uma Girish, Kunal Mittal, Ran Raz, and Wei Zhan. Polynomial bounds on parallel repetition for all 3-player games with binary inputs. In *APPROX-RANDOM*, pages 6:1–6:17, 2022. 1, 3, 4

[Has01] Johan Hastad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. 3

[HHR16] Jan Hązła, Thomas Holenstein, and Anup Rao. Forbidden subgraph bounds for parallel repetition and the density hales-jewett theorem. *CoRR*, abs/1604.05757, 2016. Available at http://arxiv.org/abs/1604.05757. 4

[Hol09] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. *Theory Comput.*, 5:141–172, 2009. (also in STOC 2007). 1, 3, 8, 10, 12

[HR20] Justin Holmgren and Ran Raz. A parallel repetition theorem for the ghz game. *arXiv preprint arXiv:2008.05059*, 2020. Available at https://arxiv.org/pdf/2008.05059.pdf. 1, 3, 4

[Kho02a] Subhash Khot. Hardness results for approximate hypergraph coloring. In John H. Reif, editor, *Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada*, pages 351–359. ACM, 2002. 3

[Kho02b] Subhash Khot. Hardness results for coloring 3 -colorable 3 -uniform hypergraphs. In *43rd Symposium on Foundations of Computer Science (FOCS 2002), 16-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 23–32. IEEE Computer Society, 2002. 3

[KORW08] Guy Kindler, Ryan O'Donnell, Anup Rao, and Avi Wigderson. Spherical cubes and rounding in high dimensions. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 189–198. IEEE Computer Society, 2008. 3

[Mit25] Kunal Mittal. Multiplayer parallel repetition is the same as high-dimensional extremal combinatorics, 2025. Available at https://arxiv.org/abs/2510.24910. 4

[Mos10] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geom. Funct. Anal.*, 19(6):1713–1756, 2010. 33

[MR21] Kunal Mittal and Ran Raz. Block rigidity: strong multiplayer parallel repetition implies super-linear lower bounds for Turing machines. In *ITCS*, pages Art. No. 71, 15, 2021. 4

[MU05] Michael Mitzenmacher and Eli Upfal. *Probability and computing*. Cambridge University Press, Cambridge, 2005. Randomized algorithms and probabilistic analysis. 50

[Pal64] Ilona Palásti. On the connectedness of bichromatic random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 8:431–441 (1964), 1964. 6

[Pol12] D. H. J. Polymath. A new proof of the density Hales-Jewett theorem. *Ann. of Math. (2)*, 175(3):1283–1327, 2012. 3

[PRW97] Itzhak Parnafes, Ran Raz, and Avi Wigderson. Direct product results and the GCD problem, in old and new communication models. In Frank Thomson Leighton and Peter W. Shor, editors, *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, El Paso, Texas, USA, May 4-6, 1997*, pages 363–372. ACM, 1997. 3

[Rao11] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011. 3

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. (also in STOC 1995). 1, 3, 8, 10, 12, 50

[Raz10] Ran Raz. Parallel repetition of two prover games. In *CCC*, pages 3–6, 2010. 3

[Raz11] Ran Raz. A counterexample to strong parallel repetition. *SIAM J. Comput.*, 40(3):771–777, 2011. (also in FOCS 2008). 3

[RR12] Ran Raz and Ricky Rosen. A strong parallel repetition theorem for projection games on expanders. In *CCC*, pages 247–257, 2012. 3

[Ver96] Oleg Verbitsky. Towards the parallel repetition conjecture. *Theoret. Comput. Sci.*, 157(2):277–282, 1996. 1, 3, 4