Parallel Repetition for the GHZ Game: A Simpler Proof

Uma Girish
Department of Computer Science, Princeton University, NJ, USA

Justin Holmgren
NTT Research, Sunnyvale, CA, USA

Kunal Mittal
Department of Computer Science, Princeton University, NJ, USA

Ran Raz
Department of Computer Science, Princeton University, NJ, USA

Wei Zhan
Department of Computer Science, Princeton University, NJ, USA

Abstract

We give a new proof of the fact that the parallel repetition of the (3-player) GHZ game reduces the value of the game to zero polynomially quickly. That is, we show that the value of the n-fold GHZ game is at most $n^{-\Omega(1)}$. This was first established by Holmgren and Raz [18]. We present a new proof of this theorem that we believe to be simpler and more direct. Unlike most previous works on parallel repetition, our proof makes no use of information theory, and relies on the use of Fourier analysis.

The GHZ game [15] has played a foundational role in the understanding of quantum information theory, due in part to the fact that quantum strategies can win the GHZ game with probability 1. It is possible that improved parallel repetition bounds may find applications in this setting.

Recently, Dinur, Harsha, Venkat, and Yuen [7] highlighted the GHZ game as a simple three-player game, which is in some sense maximally far from the class of multi-player games whose behavior under parallel repetition is well understood. Dinur et al. conjectured that parallel repetition decreases the value of the GHZ game exponentially quickly, and speculated that progress on proving this would shed light on parallel repetition for general multi-player (multi-prover) games.

2012 ACM Subject Classification Theory of computation → Interactive proof systems

Keywords and phrases Parallel Repetition, GHZ, Polynomial, Multi-player

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2021.62

Category RANDOM

Related Version *Full Version*: https://eccc.weizmann.ac.il/report/2021/101/

Funding *Uma Girish*: Simons Collaboration on Algorithms and Geometry, by a Simons Investigator Award and by the National Science Foundation grants No. CCF-1714779, CCF-2007462.
Kunal Mittal: Simons Collaboration on Algorithms and Geometry, by a Simons Investigator Award and by the National Science Foundation grants No. CCF-1714779, CCF-2007462.
Ran Raz: Simons Collaboration on Algorithms and Geometry, by a Simons Investigator Award and by the National Science Foundation grants No. CCF-1714779, CCF-2007462.
Wei Zhan: Simons Collaboration on Algorithms and Geometry, by a Simons Investigator Award and by the National Science Foundation grants No. CCF-1714779, CCF-2007462.

1 Introduction

The focus of this paper is multi-player games, and in particular their asymptotic behavior under parallel repetition.

Multi-player games consist of a one-round interaction between a referee and k players. In this interaction, the referee first samples a "query" $(q_1, \ldots, q_k)$ from some joint query distribution $\mathcal{Q}$, and for each i sends q_i to the i^{th} player. The players are required to respectively produce "answers" $a_1, \ldots, a_k$ without communicating with one another (that is, each a_i is a function only of q_i) and they are said to *win* the game if $(q_1, \ldots, q_k, a_1, \ldots, a_k)$ satisfy some predicate W that is fixed and associated with the game.

Suppose that a game G has the property that the maximum probability with which players can win is $1 - \epsilon$, no matter what strategy they use. This quantity is called the value of G. The parallel repetition question [13] asks

How well can the players concurrently play in n independent copies of G?

More precisely, consider the following k-player game, which we call the n-wise parallel repetition of G and denote by G^n:

1. The referee samples, for each $i \in [n]$ independently, query tuples $(q_1^i, \ldots, q_k^i) \sim \mathcal{Q}$. We refer to the index i as a coordinate of the parallel repeated game.
2. The j^{th} player is given $(q_j^1, \ldots, q_j^n)$ and is required to produce a tuple $(a_j^1, \ldots, a_j^n)$.
3. The players are said to win in coordinate i if $(q_1^i, \ldots, q_k^i, a_1^i, \ldots, a_k^i)$ satisfies W. They are said to win (without qualification) if they win in every coordinate $i \in [n]$.

One might initially conjecture that the value of G^n is $(1 - \epsilon)^n$. However, this turns out not to be true [14, 9, 12, 25], as players may benefit from correlating their answers across different coordinates. Still, Raz showed that if G is a two-player game, then the value of G^n is $2^{-\Omega(n)}$, where the Ω hides a game-dependent constant [23, 17]. Tighter results, based on the value of the initial game are also known [8, 5]. For many applications, such bounds are qualitatively as good as the initial flawed conjecture.

Games involving three or more players have proven more difficult to analyze, and the best known general bound on their parallel repeated value is due to Verbitsky [26]. This bound states that the value of G^n approaches 0, but the bound is *extremely* weak (it shows that the value is at most $\frac{1}{\alpha(n)}$, where α denotes an inverse Ackermann function). The weakness of this bound is generally conjectured to reflect limitations of current proof techniques rather than a fundamental difference in the behavior of many-player games. In the technically incomparable but related *no-signaling setting* however, Holmgren and Yang showed that three-player games genuinely behave differently than two-player games [19]. Specifically, they showed that there exists a three-player game with "no-signaling value" bounded away from 1 such that no amount of parallel repetition reduces the no-signaling value at all.

Parallel repetition is a mathematically natural operation that we find worthy of study in its own right. At the same time, parallel repetition bounds have found several applications in theoretical computer science (see this survey by [24]). For example, parallel repetition of 2 player games shares intimate connections with multi-player interactive proofs [4], probabilistically checkable proofs and hardness of approximation [3, 10, 16], geometry of foams [11, 20, 1], quantum information [6], and communication complexity [22, 2]. Recent work also shows that strong parallel repetition for a particular class of multiprover games implies new time lower bounds on Turing machines that can take advice [21].

Dinur et al. [7] describe a restricted class of multi-player games for which Raz's approach generalizes (giving exponential parallel bounds). Specifically, they consider games whose query distribution satisfies a certain connectivity property. For games outside this class,

Verbitsky's bound was the best known. Dinur et al. highlighted one simple three-player game, called the GHZ game [15], that in some sense is maximally far from the aforementioned tractable class of multi-player games. In the GHZ game, the players' queries are (q_1, q_2, q_3) chosen uniformly at random from $\{0,1\}^3$ such that $q_1 \oplus q_2 \oplus q_3 = 0$, and the players' goal is to produce (a_1, a_2, a_3) such that $a_1 \oplus a_2 \oplus a_3 = q_1 \vee q_2 \vee q_3$. Dinur et al. conjectured that parallel repetition decreases the value of the GHZ game exponentially quickly, and speculated that progress on proving this would shed light on parallel repetition for general games. The GHZ game has also played a foundational role in the understanding of quantum information theory, due in part to the fact that quantum strategies can win the GHZ game with probability 1. It is possible that improved parallel repetition bounds will find applications in this setting as well.

In a recent work, Holmgren and Raz [18] proved the following polynomial upper bound on the parallel repetition of the GHZ game:

▶ **Theorem 1.** *The value of the n-wise repeated GHZ game is at most $n^{-\Omega(1)}$.*

Our main contribution is a different proof of this theorem that, in our view, is significantly simpler and more direct than the proof of [18]. Like [18], we actually do not rely on any properties of the GHZ game other than its query distribution, and in particular we do not rely on specifics of the win condition. Furthermore, unlike most previous works on parallel repetition, our proof makes no use of information theory, and instead relies on the use of Fourier analysis.

1.1 Technical Overview

Let $\mathcal{P}$ denote the distribution of queries in the n-wise parallel repeated GHZ game. Let $\alpha = \Theta(1/n^{\varepsilon})$ for a small constant $\varepsilon > 0$ and $E = E_1 \times E_2 \times E_3$ be any product event with significant probability under $\mathcal{P}$, i.e., $\mathcal{P}(E) \geq \alpha$. The core of our proof is establishing that for a random coordinate $i \in [n]$, the query distribution $\mathcal{P}|E$ ($\mathcal{P}$ conditioned on E) is mildly hard in the i^{th} coordinate. That is, given queries sampled from $\mathcal{P}|E$, the players' maximum winning probability in the i^{th} coordinate is bounded away from 1. Using standard arguments from the parallel repetition literature, this will imply an inverse polynomial bound for the value of the n-fold GHZ game. The difficulty, as usual, is that the n different queries in $\mathcal{P}|E$ may not be independent.

Our approach at a high level is to:

1. Identify a class $\mathcal{D}$ of simple distributions (over queries for the n-wise repeated GHZ game) such that it is easy to analyze (in step 3 below) which coordinates are hard for any given $D \in \mathcal{D}$. By hard, we mean that the players' maximum winning probability in the i^{th} coordinate is $\frac{3}{4}$.
2. Approximate $\mathcal{P}|E$ by a convex combination of distributions from $\mathcal{D}$. That is, we write
$$\mathcal{P}|E \approx \sum_j p_j D_j,$$
where $\{D_j\}$ are distributions in $\mathcal{D}$, p_j are non-negative reals summing to 1, and $\approx$ denotes closeness in total variational distance.
3. Show that in the above convex combination, "most" of the D_i have many hard coordinates. More precisely, if we sample j with probability p_j, then the expected fraction of coordinates in which D_j is hard is at least a constant (say 1/3).

Completing this approach implies that if $i \in [n]$ is uniformly random, then the i^{th} coordinate of $\mathcal{P}|E$ can be won with probability at most $1 - \Omega(1)$. We elaborate on each of these steps below.

Bow Tie Distributions

For our class of "simple" distributions $\mathcal{D}$, we introduce the notion of a "bow tie" distribution. We then define $\mathcal{D}$ to be the set of all bow tie distributions. A bow tie is a set B of the form

$$\left\{ \begin{array}{l} (x_0, y_0, z_0), \\ (x_0, y_1, z_1), \\ (x_1, y_0, z_1), \\ (x_1, y_1, z_0) \end{array} \right\} \subseteq (\mathbb{F}_2^n)^3$$

such that for each (x, y, z) in B, we have $x + y + z = 0$. In particular this requires that $x_0 + x_1 = y_0 + y_1 = z_0 + z_1$. A bow tie distribution is the uniform distribution on a bow tie. Our name of "bow tie" is based on the fact that bow ties are thus determined by $\{(x_0, y_0), (x_0, y_1), (x_1, y_0), (x_1, y_1)\}$, which we sometimes view as a set of edges in a graph. In this case, bow ties are special kinds of $K_{2,2}$ subgraphs, where $K_{2,2}$ denotes the complete bipartite graph.

The main property of a bow tie distribution D is that for every coordinate i for which $(x_0)_i \neq (x_1)_i$ (equivalently $(y_0)_i \neq (y_1)_i$, or $(z_0)_i \neq (z_1)_i$), the i^{th} coordinate of D is as hard as the GHZ game (i.e. players cannot produce winning answers for the i^{th} coordinate with probability more than $\frac{3}{4}$). This follows by "locally embedding" the (unrepeated) GHZ query distribution into the i^{th} coordinate of D as follows. We first swap $x_0 \leftrightarrow x_1$, $y_0 \leftrightarrow y_1$, $z_0 \leftrightarrow z_1$ as necessary to ensure that

$$(x_0)_i = (y_0)_i = (z_0)_i = 0. \tag{1}$$

An even number of swaps are required to do this by the assumption that $x_0 + y_0 + z_0 = 0$, and bow ties are invariant under an even number of such swaps. Thus Equation (1) is without loss of generality. Suppose $\bar{f}_1, \bar{f}_2, \bar{f}_3 : \mathbb{F}_2^n \to \mathbb{F}_2$ comprise a strategy for the i^{th} coordinate of D. Then a strategy $f_1, f_2, f_3 : \mathbb{F}_2 \to \mathbb{F}_2$ for the basic (unrepeated) GHZ game can be constructed as

$$\begin{aligned} f_1(b) &= \bar{f}_1(x_b) \\ f_2(b) &= \bar{f}_2(y_b) \\ f_3(b) &= \bar{f}_3(z_b). \end{aligned}$$

The winning probability of this strategy is the same as the winning probability of $\bar{f}_1, \bar{f}_2, \bar{f}_3$ in the i^{th} coordinate because $\big((x_{b_1})_i, (y_{b_2})_i, (z_{b_3})_i\big) = (b_1, b_2, b_3)$. Hence both probabilities are at most $3/4$.

Approximating $\mathcal{P}|E$ by Bow Ties

We now sketch how to approximate $\mathcal{P}|E$ by a convex combination of bow tie distributions, where E is a product event $E_1 \times E_2 \times E_3$. We assume for now that the non-zero Fourier coefficients of each E_j are small. We will return to this assumption at the end of the overview – it turns out to be nearly without loss of generality.

We show that $\mathcal{P}|E$ is close in total variational distance to the distribution obtained by sampling a *uniformly random* bow tie $B \subseteq E$, and then outputting a random element of B. The latter distribution is equivalent to sampling (x, y, z) with probability proportional to the number of bow ties $B \subseteq E$ that contain (x, y, z). This number is

$$\begin{cases} \left(\sum_{z' \in \mathbb{F}_2^n} E_1(y + z') E_2(x + z') E_3(z')\right) - 1 & \text{if } (x, y, z) \in \mathsf{supp}(\mathcal{P}|E) \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where we identify E_1, E_2, and E_3 with their indicator functions. Note that we are subtracting 1 to cancel the term corresponding to $z' = z$.

Intuitively, the fact that all E_j have small Fourier coefficients means that they look random with respect to linear functions. Thus, one might guess that the above sum is close to $2^n \cdot \mu(E_1)\mu(E_2)\mu(E_3)$ for most $(x, y, z) \in \mathsf{supp}(\mathcal{P}|E)$, where $\mu(S) = |S|/2^n$ denotes the measure of S under the uniform distribution on $\mathbb{F}_2^n$. If "close to" and "most" have the right meanings, then this would imply that our distribution is close in total variational distance to $\mathcal{P}|E$ as desired.

Our full proof indeed establishes this. More precisely, we view Equation (2) as a vector indexed by (x, y, z) and establish bounds on that vector's ℓ_1 and ℓ_2 norms as a criterion for near-uniformity. In the process our proof repeatedly uses the following claims (see Lemma 16). For all sets $S, T \subseteq \mathbb{F}_2^n$ that are sufficiently large, we have

$$\mathop{\mathbb{E}}_{\substack{z\sim\mathbb{F}_2^n \\ x\sim\mathbb{F}_2^n}} [S(x) \cdot T(x+z) \cdot E_3(z)] \approx \mu(S) \cdot \mu(T) \cdot \mu(E_3)$$

and

$$\mathop{\mathbb{E}}_{z\sim\mathbb{F}_2^n} \left[\left(\mathop{\mathbb{E}}_{x\sim\mathbb{F}_2^n} [S(x) \cdot E_2(x+z)] \right)^2 \cdot E_3(z) \right] \approx \mu(S)^2 \cdot \mu(E_2)^2 \cdot \mu(E_3).$$

Most Bow Ties are Hard in Many Coordinates

For the final step of our proof, we need to show that the distribution of bow ties analyzed in the previous step produces (with high probability) bow ties that differ in many coordinates.

We begin by parameterizing a bow tie by $(x_0, y_0, x_0 \oplus x_1)$ and noting that in the previous step, we essentially showed that E contains $2^{3n-O(\log n)}$ different bow ties. The $O(\log n)$ term in the exponent arises from the fact that the events $\{E_j\}$ have density in $\mathbb{F}_2^n$ that is inverse polynomial in n. A simple counting argument then shows that for a random bow tie, the min-entropy of $x_0 \oplus x_1$ is close to n. This means that $x_0 \oplus x_1$ is close to the uniform distribution in the sense that any event occurring with probability p under the uniform distribution occurs with probability $p \cdot n^{O(1)}$ under the distribution of $x_0 \oplus x_1$. Thus we can finally apply a Chernoff bound to deduce that with all but $2^{-\Omega(n)}$ probability, $x_0 \oplus x_1$ has Hamming weight at least $n/3$.

In other words, a bow tie sampled uniformly at random differs in at least a $\frac{1}{3}$ fraction of coordinates. By the main property of bow ties, this implies that the corresponding bow tie distribution is hard on a $\frac{1}{3}$ fraction of coordinates (indeed, the same set of coordinates).

Handling General Events

For general (product) events $E = E_1 \times E_2 \times E_3$ (where the sets $\{E_i\}$ need not have small Fourier coefficients), we can partition the universe $\mathbb{F}_2^n \times \mathbb{F}_2^n \times \mathbb{F}_2^n$ into parts π such that for most of the parts π, the event E restricted to π has the structure that we already analyzed. For this to make sense, we ensure several properties of the partition. First, π should be a product set ($\pi = \pi_1 \times \pi_2 \times \pi_3$) so that $E \cap \pi$ is a product set as well, i.e. $E \cap \pi$ has the form $\tilde{E}_1 \times \tilde{E}_2 \times \tilde{E}_3$. Second, each π_i should be an affine subspace of $\mathbb{F}_2^n$ so that we can do Fourier analysis with respect to this subspace. Finally π_1, π_2, and π_3 should all be affine shifts of the *same* linear subspace so that the set $\{(x, y, z) \in \pi : x + y + z = 0\}$ has the same Fourier-analytic structure as the parallel repeated GHZ query set $\{(x, y, z) \in (\mathbb{F}_2^{n'})^3 : x + y + z = 0\}$ for some $n' < n$.

We prove the existence of such a partition with n' not too small ($n' = n - o(n)$) by a simple iterative approach, which is similar to [18].

1.2 Comparison to [18]

Our proof has some similarity to [18] – in particular, both proofs partition $(\mathbb{F}_2^n)^3$ into subspaces according to Fourier-analytic criteria and analyze these subspaces separately – but the resemblance ends there. In fact, there are fundamental high-level differences between the two proofs.

The biggest qualitative difference is that our high-level approach decomposes any conditional distribution $\mathcal{P}|E$ into components (bow tie distributions) for which many coordinates are hard. [18] takes an analogous approach, but it establishes a weaker result that differs in the order of quantifiers: it first fixes a strategy f, and then decomposes $\mathcal{P}|E$ into components such that f performs poorly on many coordinates of many components. This difference is due to the fact that [18] uses uniform distributions on high-dimensional affine spaces as their basic "hard" distributions. It is not in general possible to express $\mathcal{P}|E$ as a convex combination of such distributions (for example if each E_j is a uniformly random subset of $\mathbb{F}_2^n$). Instead, [18] expresses $\mathcal{P}|E$ as a convex combination of "pseudo-affine" distributions. This significantly complicates their proof, and we avoid this complication entirely by our use of bow tie distributions, which are novel to this work.

The remainder of our proof (the analysis of hardness within each part of the partition) is entirely different.

2 Notation & Preliminaries

A significant portion of these preliminaries is taken verbatim from [18].

We write $\exp(t)$ to denote e^t for $t \in \mathbb{R}$.

Let $n \in \mathbb{N}$. For a vector $v \in \mathbb{R}^n$ and $i \in [n]$, we write $v(i)$ or v^i to denote the i-th coordinate of v. For $p \in \mathbb{N}$, we write $\|v\|_p \stackrel{\mathsf{def}}{=} \left(\sum_{i\in[n]} |v(i)|^p\right)^{1/p}$ to denote the ℓ_p norm of v. For $z \in \{0,1\}^*$, $\mathsf{hwt}(z) \stackrel{\mathsf{def}}{=} \|z\|_1$ denotes the Hamming weight of z.

We crucially rely on the Cauchy-Schwarz inequality.

▶ **Fact 2** (Cauchy-Schwarz). *Let $k \in \mathbb{N}$ and $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{R}$. Then, $\sum_{i=1}^k |a_i \cdot b_i| \leq \sqrt{\sum_{i=1}^k a_i^2} \cdot \sqrt{\sum_{i=1}^k b_i^2}$.*

2.1 Set Theory

Let Ω be a universe. By a partition of Ω, we mean a collection of pairwise disjoint subsets of Ω, whose union equals Ω. If Π is a partition of Ω and ω is an element of Ω, we will write $\Pi(\omega)$ to denote the (unique) element of Π that contains ω. Thus, we can view Π as a function $\Pi : \Omega \to 2^\Omega$.

For a set $S \subseteq \Omega$, we identify S with its indicator function $S : \Omega \to \{0,1\}$ defined at $\omega \in \Omega$ by

$$S(\omega) = \begin{cases} 1 & \text{if } \omega \in S \\ 0 & \text{otherwise.} \end{cases}$$

For sets $S, T \subseteq \Omega$ such that $T \neq \emptyset$, we use $S|_T \subseteq T$ to denote the set $S \cap T$ when viewed as a subset of T. In particular, $S|_T$ is an indicator function from T to $\{0,1\}$.

2.2 Probability Theory

Probability Distributions

Let P be a distribution over a universe Ω. We sometimes think of P as a vector in $\mathbb{R}^{|\Omega|}$ whose value in coordinate $\omega \in \Omega$ is $P(\omega)$. In particular, we use $\|P-Q\|_1$ to denote the ℓ_1 norm of the vector $P - Q \in \mathbb{R}^{|\Omega|}$, where P and Q are probability distributions. We use $\omega \sim P$ to denote a random element ω distributed according to P. We use $\mathsf{supp}(P) = \{\omega \in \Omega : P(\omega) > 0\}$ to denote the support of the distribution P.

Random Variables

Let Σ be any alphabet. We say that $X : \Omega \to \Sigma$ is a Σ-valued random variable. If $\Sigma = \mathbb{R}$, we say that the random variable is real-valued. If X is a real-valued random variable, the expectation of X under P is denoted $\mathbb{E}_{\omega\sim P}[X(\omega)]$. Often, the underlying distribution P is implicit, in which case we simply use $\mathbb{E}[X]$. If X is a Σ-valued random variable and P is a probability distribution, we write P_X or $X(P)$ to denote the induced probability distribution of X under P, i.e., $P_X(\sigma) = (X(P))(\sigma) \stackrel{\mathsf{def}}{=} P(X = \sigma)$ for all $\sigma \in \Sigma$. In particular, we say that X is distributed according to P_X and we use $\sigma \sim X(P)$ to denote a random variable σ distributed according to P_X. The distribution P is often implicit, and we identify X with the underlying distribution P_X.

Events

We refer to subsets of Ω as events. We use standard shorthand for denoting events. For instance, if X is a Σ-valued random variable and $x \in \Sigma$, we write $X = x$ to denote the event $\{\omega \in \Omega : X(\omega) = x\}$. Similarly, for a subset $F \subseteq \Sigma$, we write $X \in F$ to denote the event $\{\omega \in \Omega : X(\omega) \in F\}$. We use $P(E)$ to denote the probability of E under P. When P is implicit, we use the notation $\Pr(E)$ to denote $P(E)$.

Conditional Probabilities

Let $E \subseteq \Omega$ be an event with $P(E) > 0$. Then the conditional distribution of P given E is denoted $(P|E) : \Omega \to \mathbb{R}$ and is defined to be

$$(P|E)(\omega) = \begin{cases} P(\omega)/P(E) & \text{if } \omega \in E \\ 0 & \text{otherwise.} \end{cases}$$

If E is an event, we write $P_{X|E}$ as shorthand for $(P|E)_X$.

Measure under Uniform Distribution

For any set $S \subseteq \Omega$, we sometimes identify S with the uniform distribution over S. In particular, we use $x \sim S$ to denote x sampled according to the uniform distribution on S. For $S, \pi \subseteq \Omega$ such that $\pi \neq \emptyset$, we use $\mu_\pi(S) = \frac{|S\cap\pi|}{|\pi|}$ to denote the measure of S under the uniform distribution over π. When $\pi = \Omega$, we omit the subscript and simply use $\mu(S)$.

2.3 Fourier Analysis

Fourier Analysis over Subspaces

For any (finite) vector space $\mathcal{V}$ over $\mathbb{F}_2$, the character group of $\mathcal{V}$, denoted $\widehat{\mathcal{V}}$, is the set of group homomorphisms mapping $\mathcal{V}$ (viewed as an additive group) to $\{-1, 1\}$ (viewed as a

multiplicative group). Each such homomorphism is called a character of $\mathcal{V}$. For functions mapping $\mathcal{V} \to \mathbb{R}$, we define the inner product

$$\langle f, g \rangle \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x \sim \mathcal{V}} [f(x)g(x)] .$$

The character group of $\mathcal{V}$ forms an orthonormal basis under this inner product. We refer to the all-ones functions $\chi : \mathcal{V} \to \{-1, 1\}, \chi \equiv 1$ as the trivial character or the zero character and denote this by $\chi = \emptyset$.

For all characters $\chi \neq \emptyset$, since $\langle \chi, \emptyset \rangle = 0$, we have $\mathbb{E}_{x \sim \mathcal{V}} [\chi(x)] = 0$, in particular, $\chi(\mathcal{V})$ is a uniform $\{\pm 1\}$-random variable. Let $\emptyset \neq S \subseteq \mathcal{V}$ be a set. Then $\mu_{\mathcal{V}}(S) \triangleq \frac{|S \cap \mathcal{V}|}{|\mathcal{V}|} = \widehat{S}(\emptyset)$, where we identify S with its indicator function $S : \mathcal{V} \to \{0, 1\}$ as mentioned before. For $\chi \in \widehat{\mathcal{V}}$, we have $\mathbb{E}_{x \sim S} [\chi(x)] = \frac{\widehat{S}(\chi)}{\widehat{S}(\emptyset)}$.

▶ **Fact 3.** *Given a choice of basis for $\mathcal{V}$, there is a canonical isomorphism between $\mathcal{V}$ and $\widehat{\mathcal{V}}$. Specifically, if $\mathcal{V} = \mathbb{F}_2^n$, then the characters of $\mathcal{V}$ are the functions of the form*

$$\chi_\gamma(v) = (-1)^{\gamma \cdot v}$$

for $\gamma \in \mathbb{F}_2^n$.

▶ **Definition 4.** *For any function $f : \mathcal{V} \to \mathbb{R}$, its* Fourier transform *is the function $\widehat{f} : \widehat{\mathcal{V}} \to \mathbb{R}$ defined by*

$$\widehat{f}(\chi) \stackrel{\text{def}}{=} \langle f, \chi \rangle = \mathop{\mathbb{E}}_{x \sim \mathcal{V}} [f(x)\chi(x)] .$$

Since the characters of $\mathcal{V}$ are orthonormal and $\mathcal{V}$ is finite, we can deduce that f is equal to $\sum_{\chi \in \widehat{\mathcal{V}}} \widehat{f}(\chi) \cdot \chi$.

▶ **Theorem 5** (Plancherel)**.** *For any $f, g : \mathcal{V} \to \mathbb{R}$,*

$$\langle f, g \rangle = \sum_{\chi \in \widehat{\mathcal{V}}} \widehat{f}(\chi) \cdot \widehat{g}(\chi).$$

An important special case of Plancherel's theorem is Parseval's theorem:

▶ **Theorem 6** (Parseval)**.** *For any $f : \mathcal{V} \to \mathbb{R}$,*

$$\mathop{\mathbb{E}}_{x \sim \mathcal{V}} \left[f(x)^2 \right] = \sum_{\chi \in \widehat{\mathcal{V}}} \widehat{f}(\chi)^2.$$

Fourier Analysis over Affine Subspaces

Fix any subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$ and a vector $a \in \mathbb{F}_2^n$. Let $\mathcal{U} = a + \mathcal{V}$ denote the affine subspace obtained by shifting $\mathcal{V}$ by a. For every function $f : \mathcal{V} \to \mathbb{R}$, we associate it with a function $f_a : \mathcal{U} \to \mathbb{R}$ defined by $f_a(x) = f(x + a)$ for all $x \in \mathcal{U}$. This is a bijective correspondence between the set of functions from $\mathcal{U}$ to $\mathbb{R}$ and the set of functions from $\mathcal{V}$ to $\mathbb{R}$. Under this association, we can identify $\chi \in \widehat{\mathcal{V}}$ with $\chi_a : \mathcal{U} \to \{-1, 1\}$ where $\chi_a(x) = \chi(x + a)$ for all $x \in \mathcal{U}$. This defines an orthonormal basis $\widehat{\mathcal{U}_a} := \{\chi_a : \mathcal{U} \to \{-1, 1\} \mid \chi \in \widehat{\mathcal{V}}\}$ for the vector space of functions from $\mathcal{U}$ to $\mathbb{R}$. We call this the Fourier basis for $\mathcal{U}$ with respect to a. This basis depends on the choice of the shift $a \in \mathcal{U}$. However, for all possible shifts $b \in \mathcal{U}$ and character functions $\chi \in \widehat{\mathcal{V}}$, the functions χ_a and χ_b only differ by a sign. To see this, observe that

$$\chi_a(x) = \chi(a + x) = \chi(b + x) \cdot \chi(a + b) = \chi_b(x) \cdot \chi(a + b)$$

We will sometimes ignore the subscript and simply use $\chi \in \widehat{\mathcal{V}}$ to index functions in the Fourier basis of $\mathcal{U}$. This is particularly the case when the properties we are dealing are independent of choice of basis (for example, the absolute values of Fourier coefficients of a function).

2.4 Multi-Player Games

In parallel repetition we often work with Cartesian product sets of the form $(\mathcal{X}_1 \times \cdots \times \mathcal{X}_k)^n$. For these sets, we will use subscripts to index the inner product and superscripts to index the outer product. That is, for $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_k$ we view elements x of $\mathcal{X}^n$ as tuples $(x_1, \ldots, x_k)$, where $x_i \in \mathcal{X}_i^n$. We use x_i^j or $x_i(j)$ to refer to the j^{th} coordinate of x_i. We use x^j to denote the vector $(x_1^j, \ldots, x_k^j)$.

If $\{E_i \subseteq \mathcal{X}_i\}_{i\in[k]}$ is a collection of subsets, we write $E_1 \times \cdots \times E_k$ to denote the set $\{x \in \mathcal{X} : \forall i \in [k], x_i \in E_i\}$. We say that $f : (\mathcal{X}_1 \times \cdots \times \mathcal{X}_k)^n \to (\mathcal{Y}_1 \times \cdots \times \mathcal{Y}_k)^n$ is a product function if $f = f_1 \times \cdots \times f_k$ for some functions $f_i : \mathcal{X}_i^n \to \mathcal{Y}_i^n$.

▶ **Definition 7** (Multi-player Games). *A* k-player game *is a tuple* $(\mathcal{X}, \mathcal{Y}, Q, W)$*, where* $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ *and* $\mathcal{Y} = \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_k$ *are finite sets,* Q *is a probability measure on* $\mathcal{X}$*, and* $W : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ *is a "winning" predicate. We refer to* Q *as the* query distribution *or the* input distribution *of the game.*

▶ **Definition 8** (Deterministic Strategies). *A* deterministic strategy *for a* k*-player game* $\mathcal{G} = (\mathcal{X}, \mathcal{Y}, Q, W)$ *is a function* $f = f_1 \times \cdots \times f_k$ *where each* $f_i : \mathcal{X}_i \to \mathcal{Y}_i$*. The* success probability *of* f *in* $\mathcal{G}$ *is denoted and defined as*

$$\mathsf{val}(\mathcal{G}, f) \stackrel{\mathsf{def}}{=} \Pr_{x\sim Q}\Big[W\big(x, f(x)\big) = 1\Big].$$

The most important quantity associated with a game is the maximum probability with which the game can be "won".

▶ **Definition 9.** *The* value *of a* k*-player game* $\mathcal{G} = (\mathcal{X}, \mathcal{Y}, Q, W)$*, denoted* $\mathsf{val}(\mathcal{G})$*, is the maximum, over all deterministic strategies* f*, of* $\mathsf{val}(\mathcal{G}, f)$*.*

It is often easier to construct *probabilistic* strategies for a game, i.e. strategies in which players may use shared and/or individual randomness in computing their answers.

▶ **Definition 10** (Probabilistic Strategies). *Let* $\mathcal{G} = (\mathcal{X}, \mathcal{Y}, Q, W)$ *be a* k*-player game. A* probablistic strategy *for* $\mathcal{G}$ *is a distribution* $\mathcal{F}$ *of deterministic strategies for* $\mathcal{G}$*. The* success probability *of* $\mathcal{F}$ *in* $\mathcal{G}$ *is denoted and defined as*

$$\mathsf{val}(\mathcal{G}, \mathcal{F}) \stackrel{\mathsf{def}}{=} \Pr_{\substack{x\sim Q\\ f\sim\mathcal{F}}}\Big[W\big(x, f(x)\big) = 1\Big].$$

A standard averaging argument implies that for every game, probabilistic strategies cannot achieve better success probability than deterministic strategies:

▶ **Fact 11.** *Replacing "deterministic strategies" by "probabilistic strategies" in Definition 9 yields an equivalent definition.*

The main operation on multi-player games that we consider in this paper is parallel repetition:

▶ **Definition 12** (Parallel Repetition). *Given a k-player game $\mathcal{G} = (\mathcal{X}, \mathcal{Y}, Q, W)$, its n-fold* parallel repetition, *denoted $\mathcal{G}^n$, is defined as the k-player game $(\mathcal{X}^n, \mathcal{Y}^n, Q^n, W^n)$, where $W^n(x, y) \stackrel{\text{def}}{=} \bigwedge_{j=1}^n W(x^j, y^j)$. For $x \in \mathcal{X}^n$, we refer to $x_i \in \mathcal{X}_i^n$ as the input to the i-th player.*

To bound the value of parallel repeated games, it is helpful to analyze the probability of winning in a particular instance of the game under various modified query distributions.

▶ **Definition 13** (Value in j^{th} coordinate). *If $\mathcal{G} = (\mathcal{X}, \mathcal{Y}, Q, W^n)$ is a game (with a product winning predicate), the* value of $\mathcal{G}$ in the j^{th} coordinate *for $j \in [n]$, denoted $\mathsf{val}^{(j)}(\mathcal{G})$, is the value of the game $(\mathcal{X}, \mathcal{Y}, Q, W')$, where $W'(x, y) = W(x^j, y^j)$.*

▶ **Definition 14** (Game with Modified Query Distribution). *Let $\mathcal{G} = (\mathcal{X}, \mathcal{Y}, Q, W)$ be a game. For a probability measure P on $\mathcal{X}$, we write $\mathcal{G}|P$ to denote the game $(\mathcal{X}, \mathcal{Y}, P, W)$. For an event E on $\mathcal{X}$, we write $\mathcal{G}|E$ to denote the game $(\mathcal{X}, \mathcal{Y}, Q_E, W)$.*

2.5 GHZ Distribution

Let $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$ and $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3$ where $\mathcal{X}_i = \mathcal{Y}_i = \mathbb{F}_2$. Let $\mathcal{Q}$ denote the uniform distribution over $\{(0,0,0), (0,1,1), (1,0,1), (1,1,0)\}$. Define $W : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ at $x \in \mathcal{X}, y \in \mathcal{Y}$ by $W(x, y) = 1$ if and only if $x_1 \vee x_2 \vee x_3 = y_1 + y_2 + y_3 \pmod 2$. The GHZ game refers to the 3-player game $(\mathcal{X}, \mathcal{Y}, \mathcal{Q}, W)$, which has value $3/4$. The n-fold repeated GHZ game refers to the n-fold parallel repetition of $(\mathcal{X}, \mathcal{Y}, \mathcal{Q}, W)$. Our parallel repetition results easily generalize with any other (constant-sized) answer alphabet $\mathcal{Y}'$ and any predicate W', as long as the game $(\mathcal{X}, \mathcal{Y}', \mathcal{Q}, W')$ has value less than 1.

We typically use $X = (X_1, X_2, X_3) \in \mathcal{X}^n$ to denote a random variable distributed according to $\mathcal{Q}^n$ where $X_i \in \mathcal{X}_i^n$ denotes the input to the i-th player.

3 Partitioning into Pseudorandom Subspaces

We make use of the notion of affine partition similar to the one defined in [18]. We say that Π is an affine partition of $(\mathbb{F}_2^n)^3$ of codimension at most d if Π is a partition on $(\mathbb{F}_2^n)^3$ and:

- Each part $\pi \in \Pi$ has the form $a_\pi + \mathcal{V}_\pi^3$ where $\mathcal{V}_\pi$ is a subspace of $\mathbb{F}_2^n$ and $a_\pi \in (\mathbb{F}_2^n)^3$, and
- Each $\mathcal{V}_\pi$ has codimension at most d.

The main take-away from this section is Proposition 15, which states the following: Given the query distribution to the n-fold GHZ game, and a product event $E \subseteq (\mathbb{F}_2^n)^3$ with large enough probability mass, we can find an affine partition Π of $(\mathbb{F}_2^n)^3$ such that on a typical part $\pi \in \Pi$, the non-zero Fourier coefficients of the indicator functions $E_1|_{\pi_1}, E_2|_{\pi_2}, E_3|_{\pi_3}$ are small. Recall that $E_i|_{\pi_i} : \pi_i \to \{0,1\}$ is the indicator function of the set $E_i \cap \pi_i \subseteq \pi_i$.

Formally, the proposition is as follows:

▶ **Proposition 15.** *Let $\mathcal{P} = \mathcal{Q}^n$. Let $E = E_1 \times E_2 \times E_3 \subseteq (\mathbb{F}_2^n)^3$ be such that $\mathcal{P}(E) = \alpha$. For all $\delta > 0$, there exists an affine partition Π of $(\mathbb{F}_2^n)^3$ of codimension at most $\frac{3}{\delta^3}$ such that the following holds. With probability at least $1 - \frac{\delta}{\alpha}$ over $\pi \sim \Pi(\mathcal{P}|E)$, for all $i \in [3]$ and non-zero $\chi \in \widehat{\mathcal{V}}$, we have $\left|\widehat{E_i|_{\pi_i}}(\chi)\right| \leq \delta$, where π is of the form $\pi_1 \times \pi_2 \times \pi_3$ for affine shifts π_1, π_2, π_3 of some subspace $\mathcal{V}$ of $\mathbb{F}_2^n$.*

Recall that $\Pi(\mathcal{P}|E)$ is the distribution induced by sampling $x \sim \mathcal{P}|E$ and outputting the part of Π to which x belongs. Note that in the statement of the proposition, we don't specify a choice of Fourier basis for π_i. This is because for any set $S \subseteq \pi_i$, the quantity $\left|\widehat{S}(\chi_{a_i})\right|$ is

independent of choice of $a_i \in \pi_i$ so we simply write $\left|\widehat{S}(\chi)\right|$. The proof of Proposition 15 is similar in nature to the proof of Lemma 6.2 in [18], but is much simpler and is presented in the full version of the paper.

4 Key Fourier Analytic Lemmas

We crucially make use of the following lemma, the proof of which can be found in the full version of the paper.

▶ **Lemma 16.** *Let $\mathcal{V} \subseteq \mathbb{F}_2^n$ be a subspace and $a_1, a_2, a_3 \in \mathbb{F}_2^n$ be such that $a_1 + a_2 + a_3 = 0$. Let $\pi = \pi_1 \times \pi_2 \times \pi_3$ where $\pi_i = a_i + \mathcal{V}$. Let $A \subseteq \pi_1, B \subseteq \pi_2, C \subseteq \pi_3$ be sets such that for all non-zero $\chi \in \widehat{\mathcal{V}}$, we have $\left|\widehat{C}(\chi)\right| \leq \delta_1$. Then,*

$$\left| \mathop{\mathbb{E}}_{\substack{z \sim \pi_3 \\ x \sim \pi_1}} [A(x) \cdot B(x+z) \cdot C(z)] - \mu_{\pi_1}(A) \cdot \mu_{\pi_2}(B) \cdot \mu_{\pi_3}(C) \right| \leq \delta_1.$$

If furthermore for all non-zero $\chi \in \widehat{\mathcal{V}}$, we have $\left|\widehat{B}(\chi)\right| \leq \delta_2$, then

$$\left| \mathop{\mathbb{E}}_{z \sim \pi_3} \left[\left(\mathop{\mathbb{E}}_{x \sim \pi_1} [A(x) \cdot B(x+z)] \right)^2 \cdot C(z) \right] - \mu_{\pi_1}(A)^2 \cdot \mu_{\pi_2}(B)^2 \cdot \mu_{\pi_3}(C) \right| \leq \delta_2^2 + \delta_1.$$

Recall from Section 2.2 that $\mu_{\pi_i}(S) \triangleq \frac{|S \cap \pi_i|}{|\pi_i|}$. In the statement of this lemma, we don't specify a choice of Fourier basis for π_2 and π_3. Since the properties $\left|\widehat{C}(\chi_{a_3})\right| \leq \delta_1$ and $\left|\widehat{B}(\chi_{a_2})\right| \leq \delta_2$ are independent of the choice of a_2 and a_3, we simply write $\left|\widehat{C}(\chi)\right| \leq \delta_1$ and $\left|\widehat{B}(\chi)\right| \leq \delta_2$.

5 Main Proof

We use the following Parallel Repetition Criterion which is similar to, but weaker than the one from [18] for the GHZ game and has a slightly simpler proof.

Let $\mathcal{G}$ refer to the n-fold parallel repetition of the GHZ game. Let $\mathcal{P} = \mathcal{Q}^n$.

▶ **Lemma 17** (Parallel Repetition Criterion)**.** *Let $c \in (0, 1]$ be a constant and $\rho(n) : \mathbb{N} \to \mathbb{R}$ be a function such that $\rho(n) \geq \exp(-n)$. Suppose for all large $n \in \mathbb{N}$ and all subsets $E_1, E_2, E_3 \subseteq \mathbb{F}_2^n$ such that $\mathcal{P}(E) \geq \rho(n)$ where $E = E_1 \times E_2 \times E_3$, we have $\mathbb{E}_{i \sim [n]} \left[\mathsf{val}^{(i)}(\mathcal{G}|E) \right] \leq 1 - c$. Then,*

$$\mathsf{val}(\mathcal{G}) \leq \rho(n)^{\Omega(1)}.$$

This lemma is proved in [18] under the weaker assumption that there is *some* coordinate $i \in [n]$ for which $\mathsf{val}^{(i)}(\mathcal{G}|E) \leq 1 - c$. The proof is slightly simpler under our stronger assumption that $\mathbb{E}_{i \sim [n]} \left[\mathsf{val}^{(i)}(\mathcal{G}|E) \right] \leq 1 - c$. We prove this in Appendix A.1.

Given this criterion, our goal of showing an inverse polynomial bound for $\mathsf{val}(\mathcal{G})$ reduces to showing the following. Let $E = E_1 \times E_2 \times E_3$ be any event such that $\mathcal{P}(E) = \alpha \geq \frac{1}{n^{1/100}}$ and n be large enough. It suffices to show that $\mathbb{E}_{i \sim [n]} \left[\mathsf{val}^{(i)}(\mathcal{G}|E) \right] \leq 0.95$. We do this as follows.

Let $\delta = \frac{\alpha^{20}}{n^{1/40}}$. Proposition 15 implies the existence of a partition Π of $(\mathbb{F}_2^n)^3$ into affine subspaces of codimension at most $O\left(\frac{1}{\delta^3}\right) = o(n)$ such that:

- Every $\pi \in \Pi$ is of the form $a + \mathcal{V}^3$ where $\mathcal{V} \subseteq \mathbb{F}_2^n$ is a subspace and $a \in (\mathbb{F}_2^n)^3$.
- With probability at least $1 - \frac{\delta}{\mathcal{P}(E)} \geq 1 - o(1)$ over $\pi \sim \Pi(\mathcal{P}|E)$, we have $\left|\widehat{E_i|_{\pi_i}}(\chi)\right| \leq \delta$ for all $i \in [3]$ and non-zero $\chi \in \widehat{\mathcal{V}}$, where $\mathcal{V}$ is the subspace of $\mathbb{F}_2^n$ for which π is an affine shift of $\mathcal{V}^3$.

Under the distribution $\Pi(\mathcal{P}|E)$, the probability that π is sampled equals $\frac{(\mathcal{P}|\pi)(E)\cdot\mathcal{P}(\pi)}{\mathcal{P}(E)}$ by Bayes' rule. This implies that the probability that $\pi \sim \Pi(\mathcal{P}|E)$ satisfies $(\mathcal{P}|\pi)(E) \leq \mathcal{P}(E)/10$ is at most $1/10$. We will focus on $\pi = \pi_1 \times \pi_2 \times \pi_3$ that satisfy both these properties, namely, the measure of E under $\mathcal{P}|\pi$ is significant, furthermore, for all $i \in [3]$, all non-zero Fourier coefficients of the sets E_i restricted to π_i are small.

▶ **Definition 18.** *We say that π is* good *if*

$$(\mathcal{P}|\pi)(E) \geq \alpha/10, \text{ and for all non-zero } \chi \in \widehat{\mathcal{V}} \text{ and } i \in [3], \text{ we have } \left|\widehat{E_i|_{\pi_i}}(\chi)\right| \leq \delta. \quad (3)$$

By a union bound, a random $\pi \sim \Pi(\mathcal{P}|E)$ will be good with probability at least $1 - \frac{1}{10} - \frac{\delta}{\alpha}$. Fix any such good $\pi = \pi_1 \times \pi_2 \times \pi_3 \in \Pi$, and let $\mathcal{V}$ be the subspace such that π is an affine shift of $\mathcal{V}^3$.

For all $z \in E_3 \cap \pi_3$, define a (partial) matching M_z between π_1 and π_2 as follows. For $x \in \pi_1 \cap E_1, y \in \pi_2 \cap E_2, z \in \pi_3 \cap E_3$ such that $x + y = z$, put an edge (x, y). Let L_z (resp. R_z) be the left (resp. right) endpoints of M_z. Let $G = \cup_{z \in E_3 \cap \pi_3} M_z$ be the bipartite graph between π_1 and π_2 obtained by combining edges from the matchings for $z \in E_3 \cap \pi_3$. Let $E(G)$ denote the set of edges in G. For every edge $e \in E(G)$, we can identify e with a valid input to the n-fold GHZ game that is contained in $E \cap \pi$. Namely, we associate $(x_0, y_0) \in E(G)$ to the input $(x_0, y_0, x_0 + y_0) \in \mathsf{supp}(\mathcal{P}) \cap E \cap \pi$. This is a bijective correspondence because of the way we defined the graph G. Under this correspondence, the uniform distribution over edges of G corresponds to the distribution $\mathcal{P}|E, \pi$. We now introduce the important notion of a bow tie.

▶ **Definition 19** (Bow Tie). *We say that a subset of edges $b \subseteq E(G)$ is a* bow tie *if $b = \{x_0, x_1\} \times \{y_0, y_1\}$ for some $x_0 \neq x_1 \in \pi_1$, $y_0 \neq y_1 \in \pi_2$ such that $x_0 + y_0 = x_1 + y_1$ (or equivalently $x_0 + y_1 = x_1 + y_0$). Alternatively, for $z_0 = x_0 + y_0$ and $z_1 = x_0 + y_1$, we have $(x_i, y_j, z_k) \in \mathsf{supp}(\mathcal{P})$ for all $(i, j, k) \in \mathsf{supp}(\mathcal{Q})$.*

Let $b = \{x_0, x_1\} \times \{y_0, y_1\}$ be a bow tie. As before, we identify b with the indicator vector $b \in \{0,1\}^{E(G)}$ of the edges of b, that is, $b(e) = 1$ iff $e \in \{(x_i, y_j) : i, j \in \{0, 1\}\}$. We use $\tilde{b}$ to denote the uniform distribution on the edges of the bow tie, when viewed as inputs to the n-fold GHZ game. More precisely, $\tilde{b}$ denotes the uniform distribution on $\{(x_i, y_j, x_i + y_j) \,|\, i, j \in \{0, 1\}\}$.

We say that b differs in the i-th coordinate *for $i \in [n]$ if $x_0(i) \neq x_1(i)$, or equivalently, $y_0(i) \neq y_1(i)$, or equivalently, $z_0(i) \neq z_1(i)$.*

Let b be a bow tie and $I \subseteq [n]$ be the coordinates on which b differs. The following claim shows that $\mathsf{val}^{(i)}(\mathcal{G}|\tilde{b}) \leq 3/4$ for all $i \in I$. The proof is deferred to Appendix A.2

▷ Claim 20. Let $b = \{x_0, x_1\} \times \{y_0, y_1\}$ be a bow tie. Let $I \subseteq [n]$ be the subset of coordinates on which b differs. Then, $\mathsf{val}^{(i)}(\mathcal{G}|\tilde{b}) \leq 3/4$ for all $i \in I$.

Let B denote the set of all bow ties. Consider the distribution on edges defined by first sampling a uniformly random bow tie from B, and then a uniformly random edge from the bow tie. We now provide an alternate description of this distribution. For each $z \in E_3 \cap \pi_3$, define $1_z \in \{0,1\}^{|E(G)|}$ as follows. For each $e = (x, y) \in E(G)$, define $1_z(e) = 1$ if x and y

are both matched in M_z but not to each other, and define $1_z(e) = 0$ otherwise. Alternatively, 1_z is the indicator of the set $((L_z \times R_z) \setminus M_z) \cap E(G)$. Let $v := \mathbb{E}_{z \sim E_3 \cap \pi_3}[1_z]$. Note that v has $|E(G)|$ coordinates, each of which have non-negative values, so v induces a distribution on $E(G)$. Consider this distribution $\tilde{v} = \frac{v}{\|v\|_1}$ on $E(G)$ defined by normalizing v. We show that this distribution is an alternate description of the aforementioned distribution.

▷ Claim 21. $v = |E_3 \cap \pi_3|^{-1} \cdot \left(\sum_{b \in B} b\right)$. In particular, we can think of the distribution $\tilde{v} := \frac{v}{\|v\|_1}$ on $E(G)$ as obtained by sampling a uniformly random bow tie b in G and outputting a uniformly random edge of b.

The proof of this is deferred to Appendix A.3. Our goal now is to show that the distribution $\tilde{v}$ is close to the uniform distribution over edges of G. To do so, we study some properties of G. Observe that $|E(G)| \triangleq |\mathcal{V}|^2 \cdot \mathbb{E}_{\substack{x \sim \pi_1 \\ z \sim \pi_3}}[E_1(x) \cdot E_2(x+z) \cdot E_3(z)]$. We apply Lemma 16 with parameters $A = E_1 \cap \pi_1, B = E_2 \cap \pi_2, C = E_3 \cap \pi_3$. Since $\pi \in \mathsf{supp}(\Pi(\mathcal{P}|E))$, the set $\pi \cap \mathsf{supp}(\mathcal{P})$ is non-empty, therefore, we may choose $a \in \mathsf{supp}(\mathcal{P})$ so that $\pi = a + \mathcal{V}^3$. This, along with Equation (3) implies that the first hypothesis of Lemma 16 is satisfied. Lemma 16 implies that

$$\left||E(G)| - |\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \cdot \mu_{\pi_3}(E_3)\right| \leq |\mathcal{V}|^2 \cdot \delta. \tag{4}$$

We make use of the following bounds on the ℓ_1 and ℓ_2 norms of v. The proofs of these are by Fourier analysis and are deferred to Appendices A.4 and A.5.

▷ Claim 22.

$$\begin{aligned}\|v\|_1 \geq\ & |\mathcal{V}|^2 \cdot \left(\mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \cdot \mu_{\pi_3}(E_3) - 3 \cdot \delta\right) \\ & - |\mathcal{V}| \cdot \left(\mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) + 2 \cdot \delta \cdot \mu_{\pi_3}(E_3)^{-1}\right)\end{aligned} \tag{5}$$

▷ Claim 23.

$$\|v\|_2^2 \leq |\mathcal{V}|^2 \cdot \left(\mu_{\pi_1}(E_1)^3 \cdot \mu_{\pi_2}(E_2)^3 \cdot \mu_{\pi_3}(E_3) + 10 \cdot \sqrt{\delta}\right) \tag{6}$$

We now bound $\|\tilde{v}\|_2 = \frac{\|v\|_2}{\|v\|_1}$ by plugging in appropriate bounds on δ and dividing Equation (6) by Equation (5). Our choice of $\delta = \alpha^{20}/n^{1/40}$, and our assumption that $\alpha/10 \leq (\mathcal{P}|\pi)(E)$ (which in turn is at most $\min_{i \in [3]}(\mu_{\pi_i}(E_i))$) implies that δ is much smaller than any $\mu_{\pi_i}(E_i)$. In particular, we highlight that

$$\begin{aligned}\sqrt{\delta} &= o\left(\mu_{\pi_1}(E_1)^3 \cdot \mu_{\pi_2}(E_2)^3 \cdot \mu_{\pi_3}(E_3)\right) \\ \delta &= o\left(\mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \cdot \mu_{\pi_3}(E_3)\right) \\ \delta &= o\left(\mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \cdot \mu_{\pi_3}(E_3)\right)\end{aligned}$$

Furthermore, since $|\mathcal{V}| = 2^{\Omega(n)}$ and $1 \geq \mu_{\pi_i}(E_i) = \Omega(\alpha) = n^{-O(1)}$, we have

$$|\mathcal{V}| \cdot \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) = o\left(|\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \cdot \mu_{\pi_3}(E_3)\right).$$

Thus the dominant term on the right-hand side of Equation (5) is $|\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \cdot \mu_{\pi_3}(E_3)$, and the dominant term on the right-hand side of Equation (6) is $|\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1)^3 \cdot \mu_{\pi_2}(E_2)^3 \cdot \mu_{\pi_3}(E_3)$. More precisely, we have

$$\|v\|_1 \geq (1 - o(1)) \cdot |\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \cdot \mu_{\pi_3}(E_3) \tag{7}$$

$$\|v\|_2^2 \leq (1 + o(1)) \cdot |\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1)^3 \cdot \mu_{\pi_2}(E_2)^3 \cdot \mu_{\pi_3}(E_3). \tag{8}$$

This implies that

$$\|\tilde{v}\|_2^2 = \frac{\|v\|_2^2}{\|v\|_1^2} \leq \frac{1+o(1)}{|\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \cdot \mu_{\pi_3}(E_3)} \tag{9}$$

In comparison, Equation (4) gave that

$$|E(G)| \in (1 \pm o(1)) \cdot |\mathcal{V}|^2 \cdot \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \cdot \mu_{\pi_3}(E_3).$$

Thus we can rewrite Equation (9) as

$$\|\tilde{v}\|_2 \leq \frac{1+o(1)}{\sqrt{|E(G)|}} \tag{10}$$

This, together with the fact that by construction $\|\tilde{v}\|_1 = 1$, is sufficient to deduce that $\tilde{v}$ is close to the "uniform distribution" vector $\tilde{u} \stackrel{\text{def}}{=} (\frac{1}{|E(G)|}, \ldots, \frac{1}{|E(G)|})$. More formally, we have:

▶ **Fact 24.** *Suppose that $\tilde{v} \in \mathbb{R}^m$ is an m-dimensional vector such that $\|\tilde{v}\|_1 = 1$, and $\|\tilde{v}\|_2 = \frac{1+\beta}{\sqrt{m}}$ for some $\beta \in [0,1]$. Then*

$$\|\tilde{v} - \tilde{u}\|_1 \leq \sqrt{3\beta},$$

where $\tilde{u}$ denotes the vector $(\frac{1}{m}, \ldots, \frac{1}{m})$.

The proof of Fact 24 is deferred to Appendix A.6

Applying Fact 24 to Equation (10) shows that $d_{\mathsf{TV}}(\tilde{v}, \tilde{u}) = o(1)$. In other words, a uniformly random edge of a uniformly random bow tie is distributed close to uniformly on $E(G)$.

We now show that a typical bow tie differs in a considerable fraction of coordinates.

▷ Claim 25. $\Pr_{\substack{i\sim[n]\\ b\sim B}}[b \text{ differs in } i\text{-th coordinate}] \geq 1/3 - o(1)$.

The proof of Claim 25 is deferred to Appendix A.7.

Claim 20, along with Claim 25 implies that $\Pr_{\substack{i\sim[n]\\ b\sim B}}[\mathsf{val}^{(i)}(\mathcal{G}|\tilde{b}) \leq 3/4] \geq 1/3 - o(1) \geq 0.3$. For those $i \in [n]$ and $b \in B$ such that b doesn't differ at the i-th coordinate, we bound $\mathsf{val}^{(i)}(\mathcal{G}|\tilde{b})$ by 1. This, along with Claim 21 implies that $\mathbb{E}_{i\sim[n]}\left[\mathsf{val}^{(i)}(\mathcal{G}|\tilde{v})\right] \leq \mathbb{E}_{\substack{i\sim[n]\\ b\sim B}}[\mathsf{val}^{(i)}(\mathcal{G}|\tilde{b})] \leq 0.75 \times 0.3 + 1 \times 0.7 \leq 0.925$. Since $d_{\mathsf{TV}}(\tilde{u}, \tilde{v}) \leq o(1)$ and $\tilde{u}$ corresponds to $\mathcal{P}|\pi, E$, this implies that $\mathbb{E}_{i\sim[n]}\left[\mathsf{val}^{(i)}(\mathcal{G}|\pi, E)\right] = 0.925 + o(1) \leq 0.93$. Since $\pi \sim \Pi(\mathcal{P}|E)$ is good with probability at least $1 - \delta \cdot \alpha^{-1} - 1/10 \geq 0.9 - o(1) \geq 0.8$, we have $\mathbb{E}_{i\sim[n]}\left[\mathsf{val}^{(i)}(\mathcal{G}|E)\right] \leq \mathbb{E}_{\substack{i\sim[n]\\ \pi\sim\Pi(\mathcal{P}|E)}}\left[\mathsf{val}^{(i)}(\mathcal{G}|E,\pi)\right] \leq 0.8 \times 0.93 + 0.2 \times 1 < 0.95$. This, along with Lemma 17 completes the proof.

References

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

2 Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM J. Comput.*, 42(3):1327–1363, 2013. (also in STOC 2010).

3 Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998. (also in FOCS 1995).

4 Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In *STOC*, pages 113–131, 1988.

5 Mark Braverman and Ankit Garg. Small value parallel repetition for general games. In *STOC*, pages 335–340, 2015.

6 Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *CCC*, pages 236–249, 2004.

7 Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expanding games. In *ITCS*, volume 67 of *LIPIcs*, pages Art. No. 37, 16, 2017.

8 Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *STOC*, pages 624–633, 2014.

9 Uriel Feige. On the success probability of the two provers in one-round proof systems. In *CCC*, pages 116–123. IEEE Computer Society, 1991.

10 Uriel Feige. A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45(4):634–652, 1998. (also in STOC 1996).

11 Uriel Feige, Guy Kindler, and Ryan O'Donnell. Understanding parallel repetition requires understanding foams. In *CCC*, pages 179–192, 2007.

12 Uriel Feige and Oleg Verbitsky. Error reduction by parallel repetition - A negative result. *Comb.*, 22(4):461–478, 2002.

13 Lance Fortnow, John Rompel, and Michael Sipser. On the power of multi-power interactive protocols. In *CCC*, pages 156–161. IEEE Computer Society, 1988.

14 Lance Jeremy Fortnow. *Complexity-theoretic aspects of interactive proof systems*. PhD thesis, MIT, 1989.

15 Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. *Going Beyond Bell's Theorem*, pages 69–72. Springer Netherlands, Dordrecht, 1989.

16 Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. (also in STOC 1997).

17 Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. *Theory Comput.*, 5:141–172, 2009. (also in STOC 2007).

18 Justin Holmgren and Ran Raz. A parallel repetition theorem for the GHZ game. *CoRR*, abs/2008.05059, 2020. URL: `https://arxiv.org/abs/2008.05059`.

19 Justin Holmgren and Lisa Yang. The parallel repetition of non-signaling games: counterexamples and dichotomy. In *STOC*, pages 185–192. ACM, 2019.

20 Guy Kindler, Ryan O'Donnell, Anup Rao, and Avi Wigderson. Spherical cubes and rounding in high dimensions. In *FOCS*, pages 189–198, 2008.

21 Kunal Mittal and Ran Raz. Block rigidity: Strong multiplayer parallel repetition implies superlinear lower bounds for turing machines. In *ITCS*, volume 185 of *LIPIcs*, pages 71:1–71:15, 2021.

22 Itzhak Parnafes, Ran Raz, and Avi Wigderson. Direct product results and the GCD problem, in old and new communication models. In *STOC*, pages 363–372. 1997.

23 Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. (also in STOC 1995).

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

25 Ran Raz. A counterexample to strong parallel repetition. *SIAM J. Comput.*, 40(3):771–777, 2011.

26 Oleg Verbitsky. Towards the parallel repetition conjecture. In *CCC*, pages 304–307. IEEE Computer Society, 1994.

A Appendix

A.1 Proof of Lemma 17

Proof of Lemma 17. Let $\mathcal{P} = \mathcal{Q}^n$. Choose the largest integer $m \geq 0$ such that $32^{-m} \geq \rho(n) \cdot \frac{2}{c}$. Note that $m = \Theta(\log(1/\rho(n)))$. Fix any deterministic product strategy $\bar{f} = (\bar{f}_1, \bar{f}_2, \bar{f}_3)$ for the players where $\bar{f}_i : \mathbb{F}_2^n \to \mathbb{F}_2^n$ denotes the strategy for the i-th player. Let $Y_i = \bar{f}_i(X_i) \in \mathbb{F}_2^n$ denote the output of player i on input X_i. Let $\{j_1, \ldots, j_m\} \subseteq [n]$ be a set of coordinates. Let W_i denote the event of winning the GHZ game in the j_i-th coordinate under the strategy $\bar{f}$ and let $W_{\leq i} := W_1 \wedge \ldots \wedge W_i$. Observe that

$$\mathsf{val}(\mathcal{G}, \bar{f}) \leq \prod_{i=0}^{m-1} \Pr\left[W_{i+1} \mid W_{\leq i}\right].$$

We show how to construct a sequence of coordinates so that every term in the above product is at most $1 - c/2$. This would imply that $\mathsf{val}(\mathcal{G}) \leq (1 - c/2)^{\Theta(\log(1/\rho(n)))} = \rho(n)^{\Omega(1)}$. Fix any $i \in \{0, \ldots, m-1\}$ and assume that we have found $j_1, \ldots, j_i$. Let $X \sim \mathcal{P}$ and $X_{\leq i}$ denote X restricted to the coordinates $\{j_1, \ldots, j_i\}$. Let $Y_{\leq i}$ denote the outputs of the players restricted to the coordinates $\{j_1, \ldots, j_i\}$. Let $Z_{\leq i} = (X_{\leq i}, Y_{\leq i})$. Since $W_{\leq i}$ is a function of $Z_{\leq i}$, we have

$$\begin{aligned}\Pr\left[W_{i+1} \mid W_{\leq i}\right] &= \mathop{\mathbb{E}}_{z_{\leq i} \sim Z_{\leq i} | W_{\leq i}} \left[\Pr\left[W_{i+1} \mid Z_{\leq i} = z_{\leq i}\right]\right] \\ &\leq \mathop{\mathbb{E}}_{z_{\leq i} \sim Z_{\leq i} | W_{\leq i}} \left[\mathsf{val}^{(j_{i+1})}\left(\mathcal{G} | Z_{\leq i} = z_{\leq i}\right)\right]. \end{aligned} \tag{11}$$

Let $F = F(z_{\leq i})$ denote the event that $\mathcal{P}\left[Z_{\leq i} = z_{\leq i} | W_{\leq i}\right] \geq \frac{c}{2} \cdot \frac{1}{N}$ where $N = 32^i \geq \mathsf{supp}(Z_{\leq i})$. We argue that F occurs with probability at least $1 - c/2$. This is because we are sampling $z_{\leq i}$ with probability $\mathcal{P}[Z_{\leq i} = z_{\leq i} | W_{\leq i}]$, hence the measure of $z_{\leq i}$ for which $\mathcal{P}[Z_{\leq i} = z_{\leq i} | W_{\leq i}] \leq \frac{c}{2} \cdot \frac{1}{N}$ is at most $\frac{c}{2}$. Fix any $z_{\leq i}$ such that F holds. Our choice of m implies that $\frac{1}{N} \cdot \frac{c}{2} \geq \rho(n)$. Note that we can express the distribution $\mathcal{P} | Z_{\leq i} = z_{\leq i}$ as $\mathcal{P} | E$ where $E = E_1 \times E_2 \times E_3$ for $E_1, E_2, E_3 \subseteq \mathbb{F}_2^n$ and $\mathcal{P}(E) \geq \rho(n)$. The hypothesis of Lemma 17 implies that $\mathbb{E}_{j \sim [n]} \left[\mathsf{val}^{(j)}\left(\mathcal{G} | Z_{\leq i} = z_{\leq i}\right)\right] \leq 1 - c$. This implies that

$$\begin{aligned}\mathop{\mathbb{E}}_{\substack{z_{\leq i} \sim Z_{\leq i} | W_{\leq i} \\ j \sim [n]}} \left[\mathsf{val}^{(j)}\left(\mathcal{G} | Z_{\leq i} = z_{\leq i}\right)\right] &\leq \Pr_{z_{\leq i} \sim Z_{\leq i} | W_{\leq i}} [\neg F] \\ &\quad + \mathop{\mathbb{E}}_{\substack{z_{\leq i} \sim Z_{\leq i} | W_{\leq i}, F \\ j \sim [n]}} \left[\mathsf{val}^{(j)}\left(\mathcal{G} | Z_{\leq i} = z_{\leq i}\right)\right] \\ &\leq \tfrac{c}{2} + 1 - c = 1 - \tfrac{c}{2}.\end{aligned}$$

By linearity of expectation, we can fix a $j \in [n]$ such that $\mathbb{E}_{z_{\leq i} \sim Z_{\leq i} | W_{\leq i}} \left[\mathsf{val}^{(j)}\left(\mathcal{G} | Z_{\leq i} = z_{\leq i}\right)\right] \leq 1 - \frac{c}{2}$. Note that $j \notin \{j_1, \ldots, j_i\}$ since we already win the game on these coordinates. This, along with Equation (11) completes the proof. ◀

A.2 Proof of Claim 20

Proof of Claim 20. Let $i \in I$. Since the bow tie b differs in the i-th coordinate, we have

$$\{x_0(i), x_1(i)\} = \{y_0(i), y_1(i)\} = \{z_0(i), z_1(i)\} = \{0, 1\}.$$

We may thus assume without loss of generality that $x_0(i) = y_0(i) = 0$. Define embeddings $\phi_1 : \mathbb{F}_2 \to \{x_0, x_1\}$, $\phi_2 : \mathbb{F}_2 \to \{y_0, y_1\}$ and $\phi_3 : \mathbb{F}_2 \to \{z_0, z_1\}$ at $a \in \mathbb{F}_2$ by $\phi_1(a) = x_a$, $\phi_2(a) = y_a$ and $\phi_3(a) = z_a$. It follows for all $a \in \{0,1\}$ and $j \in [3]$, we have $(\phi_j(a))(i) = a$. In particular, for $\phi = \phi_1 \times \phi_2 \times \phi_3$, the distribution $\phi(\mathcal{Q})$ is exactly the distribution $\tilde{b}$. Given any strategies $\bar{f}_1, \bar{f}_2, \bar{f}_3 : \mathbb{F}_2^n \to \mathbb{F}_2^n$ for the players for the n-fold GHZ game restricted to the query distribution $\tilde{b}$, the functions ϕ_1, ϕ_2, ϕ_3 induce a strategy for the GHZ game as follows. Define $f_j : \mathbb{F}_2 \to \mathbb{F}_2$ by $f_j(a) = (\bar{f}_j(\phi_j(a)))(i)$. The success probability of the strategy $f_1 \times f_2 \times f_3$ on the distribution $\mathcal{Q}$ is exactly the success probability in the i-th coordinate of the strategy $\bar{f}_1 \times \bar{f}_2 \times \bar{f}_2$ on the distribution $\tilde{b}$. It follows that $\mathsf{val}^{(i)}(\mathcal{G}|\tilde{b}) \leq 3/4$. ◁

A.3 Proof of Claim 21

Proof of Claim 21. Fix any $e \in E(G), e = (x_0, y_0)$. This implies that $x_0 \in E_1 \cap \pi_1, y_0 \in E_2 \cap \pi_2$ and $z_0 := x_0 + y_0 \in E_3 \cap \pi_3$. Note that $v(e) = \Pr_{z \sim E_3 \cap \pi_3}[(x_0, y_0) \in (L_z \times R_z) \setminus M_z]$. For any $z_1 \in E_3 \cap \pi_3$,

$$\begin{aligned} e \in (L_{z_1} \times R_{z_1}) \setminus M_{z_1} &\iff x_0 + z_1 \in E_2 \cap \pi_2, y_0 + z_1 \in E_1 \cap \pi_1, z_1 \neq z_0 \\ &\iff x_0, x_1 \in E_1 \cap \pi_1, y_0, y_1 \in E_2 \cap \pi_2, z_1 \neq z_0 \in E_3 \cap \pi_3 \\ &\quad \text{where } x_1 := y_0 + z_1, y_1 := x_0 + z_1 \\ &\iff \{x_0, x_1\} \times \{y_0, y_1\} \text{ is a bow tie} \\ &\quad \text{where } x_1 := y_0 + z_1, y_1 := x_0 + z_1. \end{aligned}$$

This implies that for all $e = (x_0, y_0) \in E(G)$ and $z_1 \in E_3 \cap \pi_3$, we have $1_{z_1}(e) = 1$ if and only if $b = \{x_0, x_1\} \times \{y_0, y_1\}$ is a bow tie. Observe that as we vary $z_1 \in E_3 \cap \pi_3$, we obtain all possible bow ties that contain the edge e, i.e. the bow ties b for which $b(e) \neq 0$. This implies that $v \triangleq \mathbb{E}_{z_1 \sim E_3 \cap \pi_3}[1_z] = |E_3 \cap \pi_3|^{-1} \cdot \left(\sum_{b \in B} b\right)$. ◁

A.4 Proof of Claim 22

For ease of notation, we define weight functions as follows.

▶ **Definition 26** (Weight functions). *Let $\mathcal{P} = \mathcal{Q}^n$. For $z \in \pi_3$, let*

$$\mathsf{wt}_\pi(z) := \Pr_{X \sim \mathcal{P}}[(X_1 \in E_1 \text{ and } X_2 \in E_2) | (X \in \pi \text{ and } X_3 = z)] = \mathop{\mathbb{E}}_{x \sim \pi_1}[E_1(x) E_2(x+z)].$$

Proof of Claim 22. Let $z \in E_3 \cap \pi_3$. Note that $\mathsf{wt}_\pi(z) = \mu_{\pi_1}(L_z) = \mu_{\pi_2}(R_z)$. Observe that $\|1_z\|_1 = |E(G) \cap (L_z \times R_z) \setminus M_z|$. We apply Lemma 16 with parameters $A = L_z \cap \pi_1, B = R_z \cap \pi_2, C = E_3 \cap \pi_3$. The first hypothesis of Lemma 16 is satisfied due to Equation (3). Lemma 16 implies that

$$\begin{aligned} |E(G) \cap (L_z \times R_z)| &\triangleq |\mathcal{V}|^2 \cdot \mathop{\mathbb{E}}_{\substack{z' \sim \pi_3 \\ x \sim \pi_1}} [L_z(x) \cdot R_z(x+z') \cdot E_3(z')] \\ &\geq |\mathcal{V}|^2 \cdot (\mu_{\pi_1}(L_z) \cdot \mu_{\pi_2}(R_z) \cdot \mu_{\pi_3}(E_3) - \delta) \\ &\triangleq |\mathcal{V}|^2 \cdot \left(\mathsf{wt}_\pi(z)^2 \cdot \mu_{\pi_3}(E_3) - \delta\right). \end{aligned}$$

Similarly, $|M_z| \triangleq |\mathcal{V}| \cdot \mathbb{E}_{x \sim \pi_1}[E_1(x) \cdot E_2(x+z)] = |\mathcal{V}| \cdot \mathsf{wt}_\pi(z)$. We apply Lemma 16 with parameters $A = E_1, B = E_2, C = E_3$. All the hypothesis are satisfied due to Equation (3). Lemma 16, along with conditioning $z \sim \pi_3$ on $z \in E_3$ implies that

$$\left| \mathop{\mathbb{E}}_{z \sim E_3 \cap \pi_3} \left[\mathsf{wt}_\pi(z)^2\right] - \mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \right| \leq 2 \cdot \delta \cdot \mu_{\pi_3}(E_3)^{-1}. \tag{12}$$

$$\left| \mathop{\mathbb{E}}_{z \sim E_3 \cap \pi_3} [\mathsf{wt}_\pi(z)] - \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \right| \leq 2 \cdot \delta \cdot \mu_{\pi_3}(E_3)^{-1}.$$

Substituting this in the previous inequalities and taking an expectation over $z \sim E_3 \cap \pi_3$,

$$\begin{aligned}
\|v\|_1 = \mathop{\mathbb{E}}_{z \sim E_3 \cap \pi_3} [\|1_z\|_1] &= \mathop{\mathbb{E}}_{z \sim E_3 \cap \pi_3} [|E(G) \cap (L_z \times R_z)| - |M_z|] \\
&\geq |\mathcal{V}|^2 \cdot \left(\mathop{\mathbb{E}}_{z \sim E_3 \cap \pi_3} \left[\mathsf{wt}_\pi(z)^2 \right] \cdot \mu_{\pi_3}(E_3) - \delta \right) \\
&\quad - |\mathcal{V}| \cdot \mathop{\mathbb{E}}_{z \sim E_3 \cap \pi_3} [\mathsf{wt}_\pi(z)] \\
&\geq |\mathcal{V}|^2 \cdot \left(\mu_{\pi_1}(E_1)^2 \cdot \mu_{\pi_2}(E_2)^2 \cdot \mu_{\pi_3}(E_3) - 3 \cdot \delta \right) \\
&\quad - |\mathcal{V}| \cdot \left(\mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) + 2 \cdot \delta \cdot \mu_{\pi_3}(E_3)^{-1} \right).
\end{aligned}$$

◀

A.5 Proof of Claim 23

Proof of Claim 23. Define $\mathsf{wt}_\pi(\cdot)$ as in the proof of Claim 22. Let $z, z' \in E_3 \cap \pi_3$. Observe that $\langle 1_z, 1_{z'} \rangle = |E(G) \cap ((L_z \cap L_{z'}) \times (R_z \cap R_{z'})) \setminus (M_z \cup M_{z'})|$. We apply Lemma 16 with parameters $A = L_z \cap L_{z'} \cap \pi_1$, $B = R_z \cap R_{z'} \cap \pi_2$ and $C = E_3 \cap \pi_3$. The first hypothesis is satisfied due to Equation (3). Lemma 16 implies that

$$\begin{aligned}
\langle 1_z, 1_{z'} \rangle &= \left| E(G) \cap ((L_z \cap L_{z'}) \times (R_z \cap R_{z'})) \setminus (M_z \cup M_{z'}) \right| \\
&\leq |\mathcal{V}|^2 \cdot (\mu_{\pi_1}(L_z \cap L_{z'}) \cdot \mu_{\pi_2}(R_z \cap R_{z'}) \cdot \mu_{\pi_3}(E_3) + \delta).
\end{aligned}$$

Taking an expectation over $z' \sim E_3 \cap \pi_3$ and applying Cauchy-Schwartz yields that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{z' \sim E_3 \cap \pi_3} [\langle 1_z, 1_{z'} \rangle] \\
&\leq |\mathcal{V}|^2 \cdot \mathop{\mathbb{E}}_{z' \sim E_3 \cap \pi_3} [\mu_{\pi_1}(L_z \cap L_{z'}) \cdot \mu_{\pi_2}(R_z \cap R_{z'}) \cdot \mu_{\pi_3}(E_3) + \delta] \\
&\leq |\mathcal{V}|^2 \cdot \left(\sqrt{\mathop{\mathbb{E}}_{z' \sim E_3 \cap \pi_3} [\mu_{\pi_1}(L_z \cap L_{z'})^2]} \cdot \sqrt{\mathop{\mathbb{E}}_{z' \sim E_3 \cap \pi_3} [\mu_{\pi_2}(R_z \cap R_{z'})^2]} \cdot \mu_{\pi_3}(E_3) + \delta \right).
\end{aligned}$$

Observe that $\mu_{\pi_1}(L_z \cap L_{z'}) = \mathbb{E}_{x \sim \pi_1} [L_z(x) E_2(x + z')]$ for all $z' \in E_3 \cap \pi_3$. We now apply Lemma 16 with parameters $A = L_z \cap \pi_1, B = E_2 \cap \pi_2, C = E_3 \cap \pi_3$. All the hypotheses are satisfied due to Equation (3). Lemma 16, along with the aforementioned observation implies that

$$\left| \mathop{\mathbb{E}}_{z' \sim E_3 \cap \pi_3} \left[\mu_{\pi_1}(L_z \cap L_{z'})^2 \right] - \mu_{\pi_1}(L_z)^2 \cdot \mu_{\pi_2}(E_2)^2 \right| \leq 2 \cdot \delta \cdot \mu_{\pi_3}(E_3)^{-1}.$$

An analogous inequality holds for $|R_z \cap R_{z'}|$. Substituting this in the previous inequality and using the fact that $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{z' \sim E_3 \cap \pi_3} [\langle 1_z, 1_{z'} \rangle] \\
&\leq |\mathcal{V}|^2 \cdot \Bigg(\left(\mu_{\pi_1}(L_z) \cdot \mu_{\pi_2}(E_2) + \sqrt{\tfrac{2 \cdot \delta}{\mu_{\pi_3}(E_3)}} \right) \cdot \left(\mu_{\pi_2}(R_z) \cdot \mu_{\pi_1}(E_1) + \sqrt{\tfrac{2 \cdot \delta}{\mu_{\pi_3}(E_3)}} \right) \cdot \mu_{\pi_3}(E_3) \\
&\quad + \delta \Bigg) \\
&\leq |\mathcal{V}|^2 \cdot \left(\mu_{\pi_1}(L_z) \cdot \mu_{\pi_2}(R_z) \cdot \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \cdot \mu_{\pi_3}(E_3) + 8 \cdot \sqrt{\delta} \right) \\
&= |\mathcal{V}|^2 \cdot \left(\mathsf{wt}_\pi(z)^2 \cdot \mu_{\pi_1}(E_1) \cdot \mu_{\pi_2}(E_2) \cdot \mu_{\pi_3}(E_3) + 8 \cdot \sqrt{\delta} \right).
\end{aligned}$$

We now take an expectation over $z \sim E_3 \cap \pi_3$ and use Equation (12) to conclude that

$$\mathop{\mathbb{E}}_{z, z' \sim E_3 \cap \pi_3} [\langle 1_z, 1_{z'} \rangle] \leq |\mathcal{V}|^2 \cdot \left(\mu_{\pi_1}(E_1)^3 \cdot \mu_{\pi_2}(E_2)^3 \cdot \mu_{\pi_3}(E_3) + 10 \cdot \sqrt{\delta} \right).$$

◀

A.6 Proof of Fact 24

Proof of Fact 24.

$$\begin{aligned}\|\tilde{v}-\tilde{u}\|_2^2 &= \langle \tilde{v}-\tilde{u}, \tilde{v}-\tilde{u}\rangle \\ &= \|\tilde{v}\|_2^2 + \|\tilde{u}\|_2^2 - 2\langle \tilde{u}, \tilde{v}\rangle \\ &= \frac{1+2\beta+\beta^2}{m} + \frac{1}{m} - \frac{2}{m} \\ &= \frac{2\beta+\beta^2}{m} \leq \frac{3\beta}{m}.\end{aligned}$$

Finally, we bound the ℓ_1 distance in terms of the ℓ_2 distance:

$$\|\tilde{v}-\tilde{u}\|_1 \leq \|\tilde{v}-\tilde{u}\|_2 \cdot \sqrt{m} \leq \sqrt{3\beta}.$$ ◀

A.7 Proof of Claim 25

Proof of Claim 25. It suffices to show that a random $b \sim B$ differs in less than $n/3$ coordinates with probability at most $2^{-\Omega(n)} = o(1)$.

The Chernoff bound implies that $\Pr_{x_0,x_1\sim\mathbb{F}_2^n}[\mathsf{hwt}(x_0+x_1) < n/3] \leq 2^{-\Omega(n)}$. We condition on $x_0, x_1 \in \pi_1$ to conclude that $\Pr_{x_0,x_1\sim\pi_1}[\mathsf{hwt}(x_0+x_1) < n/3] \leq 2^{-\Omega(n)} \cdot \frac{2^{2n}}{|\mathcal{V}|^2}$.

Let $b = \{x_0, x_1\} \times \{y_0, y_1\}$ be a bow tie. By definition, we have $y_1 = x_0 + x_1 + y_0$. In particular, the bow tie b is uniquely identified by x_0, x_1, y_0. This implies that the probability that a random $b \sim B$ differs in less than $n/3$ coordinates is precisely

$$\begin{aligned}&\frac{|\mathcal{V}|^3}{|B|} \Pr_{\substack{x_0,x_1\sim\pi_1 \\ y_0\sim\pi_2 \\ y_1=x_0+x_1+y_0}} [\{x_0,x_1\}\times\{y_0,y_1\} \in B \text{ and } \mathsf{hwt}(x_0+x_1) < n/3] \\ \leq\ &\frac{|\mathcal{V}|^3}{|B|} \Pr_{x_0,x_1\sim\pi_1}[\mathsf{hwt}(x_0+x_1) < n/3] \\ \leq\ &\frac{|\mathcal{V}|^3}{|B|} \cdot 2^{-\Omega(n)} \cdot \frac{2^{2n}}{|\mathcal{V}|^2}\end{aligned}$$

Recall that $v = \mathbb{E}_{z\sim E_3\cap\pi_3}[1_z] = \frac{1}{\mu_{\pi_3}(E_3)\cdot|\mathcal{V}|}\sum_{z\in E_3\cap\pi_3} 1_z$, where for each e, $\sum_{z\in E_3\cap\pi_3} 1_z(e)$ equals the number of bow ties containing the edge e. Since each bow tie contains 4 edges, we have that $\|v\|_1 = \frac{4}{\mu_{\pi_3}(E_3)\cdot|\mathcal{V}|}\cdot|B|$. Then, equation (7) implies that

$$|B| \geq \frac{1}{8}\cdot|\mathcal{V}|^3\cdot\mu_{\pi_1}(E_1)^2\cdot\mu_{\pi_2}(E_2)^2\cdot\mu_{\pi_3}(E_3)^2 \geq \frac{1}{8}\cdot|\mathcal{V}|^3\cdot\alpha^6.$$

This implies that $\frac{|\mathcal{V}|^3}{|B|} \leq 8/\alpha^6$. Recall that $\alpha \geq n^{-O(1)}$ and the co-dimension of $\mathcal{V}$ is $o(n)$. This implies that $\frac{2^{2n}}{|\mathcal{V}|^2} = 2^{o(n)}$. This along with the above calculation implies that the probability that a uniformly random $b \sim B$ differs in less than $n/3$ coordinates is at most $\frac{8\cdot 2^{-\Omega(n)}}{\alpha^6}\cdot 2^{o(n)} = 2^{-\Omega(n)}$. This completes the proof. ◁