

Improved Parallel Repetition for GHZ-Supported Games via Spreadness

Yang P. Liu^{*} Shachar Lovett[†] Kunal Mittal[‡]

Abstract

We prove that for any 3-player game \mathcal{G} , whose query distribution has the same support as the GHZ game (i.e., all $x, y, z \in \{0, 1\}$ satisfying $x + y + z = 0 \pmod{2}$), the value of the n -fold parallel repetition of \mathcal{G} decays exponentially fast:

$$\text{val}(\mathcal{G}^{\otimes n}) \leq \exp(-n^c)$$

for all sufficiently large n , where $c > 0$ is an absolute constant.

We also prove a concentration bound for the parallel repetition of the GHZ game: For any constant $\epsilon > 0$, the probability that the players win at least a $(\frac{3}{4} + \epsilon)$ fraction of the n coordinates is at most $\exp(-n^c)$, where $c = c(\epsilon) > 0$ is a constant.

In both settings, our work exponentially improves upon the previous best known bounds which were only polynomially small, i.e., of the order $n^{-\Omega(1)}$. Our key technical tool is the notion of *algebraic spreadness* adapted from the breakthrough work of Kelley and Meka (FOCS '23) on sets free of 3-term progressions.

^{*}Department of Computer Science, Carnegie Mellon University. Email: yang17@andrew.cmu.edu.

[†]Department of Computer Science and Engineering, University of California, San Diego. Email: slovett@ucsd.edu. Supported by Simons Investigator Award #929894 and NSF award CCF-2425349.

[‡]Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. Email: kunal.mittal@nyu.edu. Research supported a Simons Investigator Award.

Contents

1	Introduction	3
1.1	Organization	5
2	Overview	6
2.1	General Inductive Framework and High-Level Approach	6
2.2	Step 1: Squares are Hard	7
2.3	Step 2: Approximating the Distribution	8
3	Preliminaries	9
3.1	Probability Distributions	9
3.2	Vector Spaces over $\text{GF}(2)$	9
3.3	3-Player Games and Parallel Repetition	10
3.3.1	Some Basic Results on Parallel Repetition	11
4	Algebraic Spreadness and Uniformization	11
4.1	Algebraic Spreadness	11
4.2	Uniformization	12
4.2.1	Uniformization for 1 Set	12
4.2.2	Uniformization for 2 Sets	14
4.2.3	Uniformization for 3 Sets	14
5	Uniform Square Covers	19
5.1	Combinatorial Spreadness and Graph Counts	19
5.2	Algebraic Spreadness implies Graph Counts	20
5.3	Squares inside Spread Diagonal-Product Sets	21
6	Parallel Repetition for Games with the GHZ Query Support	25
6.1	Hard Coordinate under Spread Product Events	25
6.2	Hard Coordinate under General Product Events	28
7	A Concentration Bound	30
7.1	An Improved Bound for Product Events	30
7.2	Random Sets are Hard under Parallel Repetition	32
7.3	Proof of the Concentration Bound	33
A	Marginals of a Conditional Distribution	38

1 Introduction

In a k -player game \mathcal{G} , a verifier samples a tuple of questions $(x^{(1)}, \dots, x^{(k)})$ from a distribution Q . Then, for each $j \in \{1, \dots, k\}$, the verifier gives the question $x^{(j)}$ to player j , and the player gives back an answer $a^{(j)}$, which depends only on $x^{(j)}$. The verifier now declares whether the players win or lose based on the evaluation of a predicate $V(x, a) \in \{0, 1\}$ depending on the questions $x = (x^{(1)}, \dots, x^{(k)})$ and answers $a = (a^{(1)}, \dots, a^{(k)})$. We define the game value, denoted $\text{val}(\mathcal{G})$, as the maximum winning probability (with respect to the distribution Q) over all possible player strategies; see Definitions 3.8, 3.9 for formal definitions.

A natural question that arises is: How does the value of the game behave under parallel repetition [FRS94]? The n -fold parallel repetition, denoted $\mathcal{G}^{\otimes n}$, is a game where the players play, and try to win, n independent copies of the game in parallel. More precisely, the verifier samples questions $(x_i^{(1)}, \dots, x_i^{(k)}) \sim Q$ independently for $i = 1, \dots, n$, and for each $j \in \{1, \dots, k\}$, sends questions $(x_1^{(j)}, \dots, x_n^{(j)})$ to player j , to which they give back answers $(a_1^{(j)}, \dots, a_n^{(j)})$. The verifier says the players win if $V((x_i^{(1)}, \dots, x_i^{(k)}), (a_i^{(1)}, \dots, a_i^{(k)})) = 1$ for each $i \in \{1, \dots, n\}$; see Definition 3.11 for a formal definition.

Note that for any game \mathcal{G} , it holds that $\text{val}(\mathcal{G}^{\otimes n}) \geq \text{val}(\mathcal{G})^n$, since the players can achieve this value by repeating an optimal strategy in each of the n coordinates. Although one might expect the naïve bound $\text{val}(\mathcal{G}^{\otimes n}) \leq \text{val}(\mathcal{G})^n$ to hold as well, this turns out to be false [For89, Fei91, FV02, Raz11]. Roughly speaking, this failure occurs because the players do not have to treat the n copies of the game independently, and can instead correlate their answers among different copies. Remarkably, it turns out that this failure is intimately connected to the geometry of high-dimensional Euclidean tilings [FKO07, KORW08, AK09, BM21].

Parallel repetition of 2-player games is well-understood. Raz [Raz98] showed that for any game with value less than 1, the value of the n -fold parallel repetition decays exponentially in n . Subsequent works have simplified this proof and strengthened the quantitative bounds [Hol09, BRR⁺09, Rao11, RR12, DS14, BG15]. These and related works have led to several applications in various domains, including the theory of interactive proofs [BOGKW88], PCPs and hardness of approximation [FGL⁺96, ABSS97, ALM⁺98, AS98, BGS98, Fei98, Hås01, Kho02a, Kho02b, GHS02, DGKR05, DRS05], quantum information [CHTW04, BBLV13], and communication complexity [PRW97, BBCR13, BRWY13]. The reader is referred to the survey [Raz10] for more details.

Parallel repetition of multiplayer games is much less understood. The only general bound says that for any game \mathcal{G} with value less than 1, it holds that $\text{val}(\mathcal{G}^{\otimes n}) \leq 1/\alpha(n)$, where $\alpha(n)$ is a very slowly growing inverse-Ackermann function [Ver96, FK91, Pol12]. Recent work has made substantial progress in understanding special cases of multiplayer games [DHVY17, HR20, GHM⁺21, GHM⁺22, GMRZ22, BKM23, BBK⁺24, BBK⁺25, BBK⁺26]; however, the general question remains wide open, even for 3-player games.

Proving improved parallel repetition for multiplayer games has several potential 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, which is a long-standing open problem in complexity theory [MR21]. Additionally, parallel repetition in the large answer alphabet regime is equivalent to many problems in high-dimensional extremal combinatorics, such as the density Hales-Jewett problem, and the problem of square-free

sets in finite fields [FV02, HHR16, Mit25]. Also, as stated in [DHVY17], it is believable that improved understanding of parallel repetition can lead to a better understanding of communication complexity in the number-on-forehead (NOF) model, which is intimately connected to circuit lower bounds.

The focus of this paper is the 3-player GHZ game [GHZ89], which proceeds as follows: The verifier samples questions $(x, y, z) \in \{0, 1\}^3$ uniformly at random such that $x + y + z = 0 \pmod{2}$, and the players' goal is to give answers $a, b, c \in \{0, 1\}$ respectively satisfying $a + b + c \pmod{2} = x \vee y \vee z$. The GHZ game has played a foundational role in quantum information theory, due in part to the fact that quantum strategies can win this game with probability 1, whereas any classical strategy wins with probability at most $3/4$. Moreover, the GHZ game satisfies a self-testing property, which says that all quantum strategies achieving value 1 are essentially the same; this has led to applications like entanglement testing and device-independent cryptography [SB20].

The problem of parallel repetition of the GHZ game has been discussed in several recent works. Dinur, Harsha, Venkat and Yuen [DHVY17] extend the 2-player information-theoretic techniques of Raz [Raz98] and prove parallel repetition for a specific class of multiplayer games satisfying a certain connectivity property; they identify the GHZ game as a multiplayer game which is in some sense maximally far from this class of games; regarding the difficulty of the problem, they write

“We believe that the strong correlations present in the GHZ question distribution represent the ‘hardest instance’ of the multiplayer parallel repetition problem.”

Subsequent research established polynomial decay bounds for the n -fold repetition of the GHZ game, showing $\text{val}(\text{GHZ}^{\otimes n}) \leq n^{-\Omega(1)}$ [HR20, GHM⁺21]. Notably, these results apply to any 3-player game \mathcal{G} whose query distribution Q has the same support as the GHZ game (i.e., $\{(x, y, z) \in \{0, 1\}^3 : x + y + z = 0 \pmod{2}\}$).

More recently, an exponential decay bound was proven for the GHZ game [BKM23], and then extended to all 3-player XOR games¹ satisfying a certain distributional assumption [BBK⁺25]. However, these works rely heavily on the XOR structure of the game predicate, exploiting it via sophisticated tools from Fourier analysis and additive combinatorics. For instance, in the case of the GHZ game, they observe that the predicate corresponds to a linear equation over the group $\mathbb{Z}/4\mathbb{Z}$. Consequently, these techniques do not extend to general game predicates lacking such XOR structure.

In this work, we improve upon the existing polynomial decay bounds [HR20, GHM⁺21] by establishing a stretched exponential bound for all games sharing the query support of the GHZ game. Crucially, unlike the aforementioned works on XOR games [BKM23, BBK⁺25], our result is agnostic to the answer sets and the game predicate—it solely depends on the support of the query distribution. Our technical approach is distinct as well: we utilize spreadness-based arguments from recent works [KM23, KLM24, JLL⁺25], whose application to the context of parallel repetition is novel.

Our main result is the following:

¹In a 3-player XOR game, the answers a, b, c of the three players lie in some finite Abelian group H , and the predicate is of the form $a + b + c = \varphi(x, y, z)$, for some function φ mapping questions (x, y, z) of the players to elements of H .

Theorem 1.1 (Parallel Repetition for the GHZ query support). *Let \mathcal{G} be any 3-player game with value $\text{val}(\mathcal{G}) < 1$, whose query distribution has support*

$$\{(x, y, z) \in \{0, 1\}^3 : x + y + z = 0 \pmod{2}\}.$$

Then, for all sufficiently large n ,² it holds that

$$\text{val}(\mathcal{G}^{\otimes n}) \leq \exp(-n^c),$$

where $c > 0$ is an absolute constant.

We complement this result with a concentration bound for games with the GHZ query distribution. To the best of our knowledge, the exponential decay bounds established in prior works [BKM23, BBK⁺24] do not imply such concentration. We prove the following:

Theorem 1.2 (A Concentration Bound; restated and proved as Theorem 7.1). *Let \mathcal{G} be any 3-player game with value $\text{val}(\mathcal{G}) < 1$, whose query distribution is uniform over the set*

$$\{(x, y, z) \in \{0, 1\}^3 : x + y + z = 0 \pmod{2}\}.$$

Then, for every constant $\epsilon > 0$, there exists $c = c(\epsilon) > 0$, such that for all sufficiently large n ,³ the probability that the players can win at least $\text{val}(\mathcal{G}) + \epsilon$ fraction of the n coordinates in the game $\mathcal{G}^{\otimes n}$ is at most $\exp(-n^c)$.

As a direct consequence, we obtain a concentration bound for the standard GHZ game.

Corollary 1.3. *For every constant $\epsilon > 0$, there exists $c = c(\epsilon) > 0$, such that the probability of the players winning at least $(\frac{3}{4} + \epsilon)$ n coordinates in the game $\text{GHZ}^{\otimes n}$ is at most $\exp(-n^c)$.*

It remains an interesting open problem to establish analogous concentration bounds (as in Theorem 1.2) for games over the GHZ support where the underlying distribution is not uniform.⁴

1.1 Organization

In Section 2, we give an overview of our proofs. In Section 3, we establish some preliminaries. In Section 4, we define a notion of pseudorandomness, called *algebraic spreadness*, that will be useful throughout this paper; we also show how to decompose arbitrary sets into components satisfying this spreadness condition. In Section 5, we show that any *diagonal-product* set composed of algebraically spread sets is *uniformly covered* by *squares*. Finally, in Section 6, we use the results of the previous sections and prove parallel repetition for games with the GHZ query support (Theorem 1.1). In Section 7, we prove our concentration bound (Theorem 1.2).

²i.e., $n \geq N$, where N is a constant depending on the game \mathcal{G} .

³i.e., $n \geq N$, where N is a constant depending on the game \mathcal{G} and the parameter ϵ .

⁴We remark that Theorem 1.2, combined with a reduction similar to Lemma 3.12 (also see Footnote 6), already implies the following: the probability that the players win at least $1 - \beta + \epsilon$ fraction of the n coordinates is at most $\exp(-n^{\Omega_\epsilon(1)})$, where $\beta = \beta(\mathcal{G}) > 0$ is a constant.

2 Overview

In this section, we outline the proof of our main result (Theorem 1.1).

2.1 General Inductive Framework and High-Level Approach

We begin by introducing a general inductive framework for parallel repetition. Our proof follows the basic setup established by Raz for parallel repetition of 2-player games [Raz98].

Let \mathcal{G} be a 3-player game, whose query distribution Q has support

$$\text{supp}(Q) = \{(x, y, z) \in \mathbb{F}_2^3 : x + y + z = 0\}.$$
⁵

Here, we identified $\{0, 1\}$ with the finite field \mathbb{F}_2 . For the purpose of parallel repetition, we may assume without loss of generality that Q is the uniform distribution over its support (see Lemma 3.12).⁶ Now, consider the game $\mathcal{G}^{\otimes n}$, and consider any strategy for the 3 players in this game. For each $i \in [n]$, let Win_i be the event that the players win the i^{th} coordinate. By the chain rule, for any permutation i_1, i_2, \dots, i_n of $[n]$, we can write the winning probability as

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

To bound this, we proceed inductively. Assuming the players have won a set of coordinates i_1, \dots, i_k , we aim to identify a *hard coordinate* $i \in [n]$ whose conditional winning probability is at most $1 - \Omega(1)$. Formally, we wish to prove the following condition:⁷

Inductive Step: For every product event $\mathcal{E} = E \times F \times G \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n \times \mathbb{F}_2^n$, with measure $\Pr[\mathcal{E}] \geq \alpha$, there exists a coordinate $i \in [n]$ such that $\Pr[\text{Win}_i \mid \mathcal{E}] \leq 1 - \Omega(1)$.

Establishing this condition for $\alpha = \exp(-n^{\Omega(1)})$ implies the desired bound on $\text{val}(\mathcal{G}^{\otimes n})$ via the inductive strategy above (see Lemma 3.13).

For the remainder of this overview, consider any product event $\mathcal{E} = E \times F \times G \subseteq (\mathbb{F}_2^n)^3$, with measure $\Pr[\mathcal{E}] \geq \alpha = \exp(-n^{\Omega(1)})$. Our goal is to identify a coordinate $i \in [n]$ that remains hard to win when the inputs are conditioned on \mathcal{E} . The proof proceeds in two main steps:

1. **Identify Hard Sets:** We define a class of sets called *squares*, such that if the input distribution is restricted to a square, many coordinates are hard to win.

⁵The framework applies to all k -player games, but we restrict to 3 player games with the GHZ queries.

⁶Roughly speaking, this is because any distribution Q *contains* a small copy of the uniform distribution over $\text{supp}(Q)$, and hence winning n copies under the distribution Q is at least as hard as winning $\Omega(n)$ copies under the the uniform distribution.

⁷Conditioning on the event $\text{Win}_{i_1} \wedge \dots \wedge \text{Win}_{i_k}$ can introduce complex correlations among the players' inputs. However, this event depends deterministically on the questions and answers of the players in coordinates i_1, \dots, i_k , and hence instead of conditioning on the event $\text{Win}_{i_1} \wedge \dots \wedge \text{Win}_{i_k}$, we can condition on typical questions and answers to the players in coordinates i_1, \dots, i_k . This induces a product event on the players' input.

2. **Distributional Approximation:** We show that the conditional distribution $Q^{\otimes n}|\mathcal{E}$ can be approximated (in ℓ_1 distance) by a convex combination of such square distributions.

Combining these steps establishes that $\Pr[\text{Win}_i \mid \mathcal{E}] \leq 1 - \Omega(1)$ for randomly chosen $i \in [n]$.

2.2 Step 1: Squares are Hard

Since the third player's input in $\mathcal{G}^{\otimes n}$ is fully determined by the inputs to the first two players (inputs $(x, y, z) \in \text{supp}(Q)^n \subseteq (\mathbb{F}_2^n)^3$ satisfy $z = x + y$), we can analyze the game primarily by only looking at inputs of the first two players. The restriction of a product event $\mathcal{E} = E \times F \times G$ to the first two players is captured by a *diagonal-product set*, defined as follows:

Definition 2.1 (Diagonal-Product Set). *Given sets $X, Y, Z \subseteq \mathbb{F}_2^n$, we define the corresponding diagonal-product set, denoted $S(X, Y, Z)$, as*

$$S(X, Y, Z) = \{(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n : x \in X, y \in Y, x + y \in Z\}.$$

Sampling inputs $(x, y, z) \sim Q^{\otimes n}|\mathcal{E}$ is the same as sampling $(x, y) \sim S(E, F, G)$ uniformly and setting $z = x + y$.

We now define our hard sets, that are *squares* in $\mathbb{F}_2^n \times \mathbb{F}_2^n$:

Definition 2.2 (Square). *A square $s_{x,y,w} \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$, for $x, y, w \in \mathbb{F}_2^n$, is the set*

$$s_{x,y,w} = \{(x, y), (x + w, y), (x, y + w), (x + w, y + w)\}.$$

Remark 2.3. *Given a square $s \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$, suppose we wish to represent it as $s = s_{x,y,w}$. Note that the width w is uniquely determined by the square s . If $w = 0$, the square contains a single point, in which case (x, y) is uniquely determined. If $w \neq 0$, the square $s_{x,y,w}$ has 4 different representations, given by $s_{x,y,w} = s_{x+w,y,w} = s_{x,y+w,w} = s_{x+w,y+w,w}$.*

The crucial property of a square is its local hardness: Consider a square $s = s_{x_0,y_0,w}$, and a coordinate $i \in [n]$ such that $w_i \neq 0$ (a *non-trivial* coordinate).⁸ Suppose the inputs (x, y, z) to the three players are sampled from the square s as follows: let $(x, y) \sim s$ be chosen uniformly at random, and let $z = x + y$. Under this distribution, no strategy of the players can win coordinate i with probability more than $\text{val}(\mathcal{G})$. Roughly speaking, this is true since this distribution “looks exactly the same” as the base distribution Q , and hence the players can embed a single copy of the game \mathcal{G} into coordinate i of this distribution; see Lemma 6.4 for formal details of this step.⁹

⁸A coordinate is non-trivial if and only if the points in s do not have a constant value in this coordinate.

⁹We remark that the set of inputs $\{(x, y, x + y) : (x, y) \in s\}$ was introduced in the work [GHM⁺21], where they call this set a *bow-tie*. We also note that squares correspond to the maximal *forbidden-subgraphs* inside the GHZ query distribution (in the sense that no player strategy can win on all four points of a square), and the largest size of square-free sets in $\mathbb{F}_2^n \times \mathbb{F}_2^n$ exactly captures the value of the n -fold parallel repetition of the GHZ query distribution in the large answer alphabet regime [Mit25].

2.3 Step 2: Approximating the Distribution

The second, and more technically demanding, step of the proof is to show that the distribution $Q^{\otimes n}|\mathcal{E}$ is well-approximated by a mixture of square distributions as above. Equivalently, we wish to approximate the uniform distribution over the diagonal-product set $S = S(E, F, G)$, denoted U_S , by a convex combination of uniform distributions over squares. Specifically, we analyze the distribution μ generated as follows: sample a square $s \subseteq S$ uniformly at random, and output an element uniformly at random from s .

Note that a priori it is not even clear that the set S contains squares, and that the distribution μ is well-defined. However, we utilize the concept of *algebraic spreadness*, a pseudorandomness notion introduced in the breakthrough work of Kelley and Meka on bounds for sets without 3-term arithmetic progressions [KM23]. They show that algebraically spread subsets behave like random subsets in a certain sense (see Definition 4.1 and Theorem 4.2); in our setting, we show that if all E, F, G are algebraically spread, the set S contains the expected density of squares. Our approximation argument proceeds in two stages:

Uniformization: Given arbitrary sets E, F, G , we decompose the diagonal-product set $S(E, F, G)$ into disjoint components $S(E_1, F_1, G_1) \cup S(E_2, F_2, G_2) \cdots \cup S(E_T, F_T, G_T)$ (plus a negligible remainder), such that within each component, the sets E_i, F_i, G_i are all algebraically spread. This generalizes the recent decomposition technique of [JLL⁺25], which achieved algebraic spreadness for only two out of the three sets and used it to prove bounds on corner-free sets in finite fields. We extend this to all three sets via a careful recursive argument; see Section 4.2 for more details.

Counting and Approximation: Assuming that all of E, F, G are algebraically spread, we demonstrate that they behave like random sets of the same density, and that the square-sampling distribution μ approximates the uniform distribution U_S in ℓ_1 distance. We prove this by establishing tight bounds on the ℓ_2 norm (collision probability) of the distribution μ . A key technical tool here is a recent graph-counting result of [FHHK24], based on a pseudorandomness notion called combinatorial spreadness, which we show holds in our setting. See Lemma 5.5 and Proposition 5.6 for more details on this step.

Why Algebraic Spreadness? A crucial difference between our work and previous polynomial decay bounds [GHM⁺21] lies in the choice of the pseudorandomness property for the sets E, F, G . Prior works use the concept of Fourier-uniformity (pseudorandomness against linear tests), and yield bounds when the density $\alpha = \Pr[\mathcal{E}]$ is at least polynomial, i.e., of the order $n^{-O(1)}$. The use of algebraic spreadness enables us to surpass this barrier and get bounds even when the density $\alpha = \exp(-n^{\Omega(1)})$ is exponentially smaller. However, this improvement incurs a significant technical cost. Both the decomposition and approximation steps are standard under Fourier uniformity, whereas establishing them under algebraic spreadness constitutes the main technical contribution of this work.

3 Preliminaries

Let $\mathbb{N} = \{1, 2, \dots\}$ denote the set of natural numbers. For $n \in \mathbb{N}$, we use $[n]$ to denote the set $\{1, 2, \dots, n\}$.

3.1 Probability Distributions

Let P be a distribution (over an underlying finite set Ω , which is usually clear from context). We use $\text{supp}(P) = \{\omega \in \Omega : P[\omega] > 0\}$ to denote the support of the distribution P . For an event $E \subseteq \Omega$ with $P[E] > 0$, we use $P|E$ to denote the conditional probability distribution P conditioned on E .

For distributions P and Q over a set Ω , the ℓ_1 -distance between them is defined as

$$\|P - Q\|_1 = \sum_{\omega \in \Omega} |P[\omega] - Q[\omega]|.$$

We state a useful tail bound on the sum of independent random variables:

Fact 3.1 (Chernoff Bounds, see [MU05] for reference). *Let $X_1, \dots, X_n \in \{0, 1\}$ be independent random variables each with mean μ , and let $X = \sum_{i=1}^n X_i$. Then, for all $\delta \in (0, 1)$,*

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

$$\Pr[X \geq (1 + \delta)\mu n] \leq e^{-\frac{\delta^2 \mu n}{3}}.$$

We also state another useful lemma:

Lemma 3.2. *Let $t, n \in \mathbb{N}$, $t < n$. Let μ be the uniform distribution on $[n]$, and let ν be the uniform distribution on $[n - t]$. Then, $\|\mu - \nu\|_1 = 2t/n$.*

Proof. We have $\|\mu - \nu\|_1 = \sum_{i=1}^{n-t} \left(\frac{1}{n-t} - \frac{1}{n}\right) + \sum_{i=n-t+1}^n \frac{1}{n} = \frac{2t}{n}$. □

3.2 Vector Spaces over $\text{GF}(2)$

Let \mathcal{V} be a finite dimensional vector space over \mathbb{F}_2 , with the uniform measure; often we shall have $\mathcal{V} = \mathbb{F}_2^n$ for some $n \in \mathbb{N}$.

Definition 3.3 (Inner Product). *For functions $f, g : \mathcal{V} \rightarrow \mathbb{R}$, define their inner product as*

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

where $x \sim \mathcal{V}$ denotes that x is uniformly chosen from \mathcal{V} .

Definition 3.4 (L^p norm). *For $f : \mathcal{V} \rightarrow \mathbb{R}$, and $p \geq 1$, we define*

$$\|f\|_p = \mathbb{E}[f^p]^{1/p}.$$

Definition 3.5 (Convolution). For $f, g : \mathcal{V} \rightarrow \mathbb{R}$, we define their convolution as

$$(f * g)(x) = \mathbb{E}_{y \sim \mathcal{V}}[f(y)g(x + y)].$$

We observe the following simple fact:

Fact 3.6. For functions $f, g, h : \mathcal{V} \rightarrow \mathbb{R}$,

$$\langle f * g, h \rangle = \langle f * h, g \rangle = \langle g * h, f \rangle = \mathbb{E}_{x, y \sim \mathcal{V}}[f(x)g(y)h(x + y)].$$

Definition 3.7 (Density Function). For a nonempty set $A \subseteq \mathcal{V}$, we define its density function as

$$\varphi_A = \frac{\mathbf{1}_A}{\mathbb{E}[\mathbf{1}_A]} = \frac{\mathbf{1}_A}{|A| / |\mathcal{V}|}.$$

Note that $\mathbb{E}[\varphi_A] = 1$.

3.3 3-Player Games and Parallel Repetition

We overview some basic definitions regarding multiplayer games. We shall restrict our focus to 3-player games.

Definition 3.8 (3-Player Game). A 3-player game \mathcal{G} is a tuple $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{pred})$, where the question sets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ and the answer sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are finite sets, Q is a probability distribution over $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, and $V_{pred} : (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}) \times (\mathcal{A} \times \mathcal{B} \times \mathcal{C}) \rightarrow \{0, 1\}$ is a predicate.

The game \mathcal{G} proceeds as follows: A verifier samples questions $(X, Y, Z) \sim Q$; then, the verifier sends X to player 1, Y to player 2, and Z to player 3, to which the players respond back with answers $A \in \mathcal{A}$, $B \in \mathcal{B}$, $C \in \mathcal{C}$ respectively. Finally, the verifier declares that the players win if and only if $V_{pred}((X, Y, Z), (A, B, C)) = 1$.

Definition 3.9 (Game Value). Let $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{pred})$ be a 3-player game. The value of the game \mathcal{G} , denoted $\text{val}(\mathcal{G})$, is defined as

$$\text{val}(\mathcal{G}) = \max_{f, g, h} \Pr_{(X, Y, Z) \sim Q} [V_{pred}((X, Y, Z), (f(X), g(Y), h(Z))) = 1],$$

where the maximum is over player strategies $f : \mathcal{X} \rightarrow \mathcal{A}$, $g : \mathcal{Y} \rightarrow \mathcal{B}$, $h : \mathcal{Z} \rightarrow \mathcal{C}$.

Fact 3.10. The game value is unchanged even if the players are allowed to use public and private randomness, since there always exists some optimal fixed values of the random strings.

Next, we define the parallel repetition of a 3-player game, which corresponds to playing n independent copies of the game in parallel.

Definition 3.11 (Parallel Repetition). Let $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{pred})$ be a 3-player game. We define its n -fold repetition as $\mathcal{G}^{\otimes n} = (\mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n, \mathcal{A}^n \times \mathcal{B}^n \times \mathcal{C}^n, Q^{\otimes n}, V_{pred}^{\otimes n})$. The distribution $Q^{\otimes n}$ is the n -fold product of the distribution Q with itself, i.e., $Q^{\otimes n}[(x, y, z)] = \prod_{i=1}^n Q[(x_i, y_i, z_i)]$ for each $x \in \mathcal{X}^n, y \in \mathcal{Y}^n, z \in \mathcal{Z}^n$. The predicate $V_{pred}^{\otimes n}$ is defined as $V_{pred}^{\otimes n}((x, y, z), (a, b, c)) = \bigwedge_{i=1}^n V_{pred}((x_i, y_i, z_i), (a_i, b_i, c_i))$.

3.3.1 Some Basic Results on Parallel Repetition

We state a lemma from [FV02], which shows that it suffices to prove parallel repetition in the case when the game's distribution is uniform over its support:

Lemma 3.12 ([GHM⁺22, Lemma 3.14]). *Let $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{\text{pred}})$ be a 3-player game such that $\text{val}(\mathcal{G}) < 1$. Let $\tilde{\mathcal{G}} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, U, V_{\text{pred}})$, where U is the uniform distribution over $\text{supp}(Q)$. Then, $\text{val}(\tilde{\mathcal{G}}) < 1$, and there exists a constant $c = c(\mathcal{G}) > 0$, such that for every sufficiently large $n \in \mathbb{N}$,*

$$\text{val}(\mathcal{G}^{\otimes n}) \leq 2 \cdot \text{val}(\tilde{\mathcal{G}}^{\otimes \lfloor cn \rfloor}).$$

We state an inductive parallel repetition criterion from [Raz98]:

Lemma 3.13 ([BBK⁺26, Lemma B.1]). *Let $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{\text{pred}})$ be a 3-player game, and consider its n -fold repetition $\mathcal{G}^{\otimes n} = (\mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n, \mathcal{A}^n \times \mathcal{B}^n \times \mathcal{C}^n, Q^{\otimes n}, V_{\text{pred}}^{\otimes n})$ for some sufficiently large $n \in \mathbb{N}$. Fix optimal strategies for the 3 players in this game, and for each $i \in [n]$, let Win_i be the event that this strategy wins the i^{th} coordinate of the game.*

Let $\epsilon > 0$ be a constant, and $\alpha \in (0, 1]$, $\alpha \geq 2^{-n}$ be such that the following condition holds: For every product event $E \times F \times G \subseteq \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ with $\Pr_{Q^{\otimes n}}[E \times F \times G] \geq \alpha$, there exists a coordinate $i \in [n]$ such that $\Pr[\text{Win}_i | E \times F \times G] \leq 1 - \epsilon$.

Then, for some constant $c = c(\mathcal{G})$, it holds that $\text{val}(\mathcal{G}^{\otimes n}) \leq \alpha^c$.

4 Algebraic Spreadness and Uniformization

In this section, we overview a notion of pseudorandomness called algebraic spreadness. We also show how to decompose arbitrary sets into components that satisfy this spreadness definition.

4.1 Algebraic Spreadness

Definition 4.1 (Algebraic Spreadness). *Let $\mathcal{V} \subseteq \mathbb{F}_2^n$ be an affine subspace. We say that a subset $A \subseteq \mathcal{V}$ is (r, ϵ) -algebraically spread within \mathcal{V} if for all affine subspaces $\mathcal{V}' \subseteq \mathcal{V}$ satisfying $\dim(\mathcal{V}') \geq \dim(\mathcal{V}) - r$, it holds that*

$$\frac{|A \cap \mathcal{V}'|}{|\mathcal{V}'|} \leq (1 + \epsilon) \cdot \frac{|A|}{|\mathcal{V}|}.$$

We state a useful result about spread subsets.

Theorem 4.2. *Let $d \geq 1, \epsilon \in (0, 1/4)$. Then, there exists a sufficiently large integer $r = d^8 \epsilon^{-O(1)}$, and a sufficiently small $\delta = \Omega(\epsilon)$, such that the following holds: Suppose $A, B, C \subseteq \mathbb{F}_2^n$ are sets each of size at least $2^{-d} \cdot |\mathbb{F}_2^n|$. Then,*

(i) ([KM23, Proposition 2.16]) If at least two of A, B, C are (r, δ) -algebraically spread,

$$|\langle \varphi_A * \varphi_B, \varphi_C \rangle - 1| \leq \epsilon.^{10}$$

(ii) ([KM23, Proposition 4.10]) If at least one of A, B, C is (r, δ) -algebraically spread,

$$\langle \varphi_A * \varphi_B, \varphi_C \rangle \leq 1 + \epsilon.$$

4.2 Uniformization

In this section, we describe a procedure that takes an arbitrary set $X \subseteq \mathbb{F}_2^n$ and approximately decomposes it into components that are algebraically spread. More generally, we find a simultaneous decomposition for three sets X, Y, Z , compatible with the diagonal-product:

Proposition 4.3. *Let $r \in \mathbb{N}$, $\epsilon, \eta \in (0, 1/10)$, and let $X, Y, Z \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$. Define $\alpha = |S(X, Y, Z)| / |\mathcal{V}|^2$. Then, there exists $C_\eta = (1/\eta)^{O(1)} \geq 1$ (that depends only on η), an integer $T \in \mathbb{N}$, and for each $i \in [T]$, a linear subspace $\mathcal{V}_i \subseteq \mathcal{V}$, points $x_i, y_i \in \mathcal{V}/\mathcal{V}_i$, and subsets $X_i \subseteq x_i + \mathcal{V}_i$, $Y_i \subseteq y_i + \mathcal{V}_i$, $Z_i \subseteq x_i + y_i + \mathcal{V}_i$, such that:*

1. $\dim(\mathcal{V}_i) \geq \dim(\mathcal{V}) - r\epsilon^{-3} \log_2(4/\alpha)^{C_\eta}$ for all $i \in [T]$.
2. $S(X_1, Y_1, Z_1), \dots, S(X_T, Y_T, Z_T)$ are disjoint subsets of $S(X, Y, Z)$ such that

$$|S(X, Y, Z) \setminus \cup_{i=1}^T S(X_i, Y_i, Z_i)| \leq \eta |S(X, Y, Z)|.$$

The diagonal-product sets are defined as in Definition 2.1.

3. For all $i \in [T]$, we have $|X_i|, |Y_i|, |Z_i| \geq 2^{-(\log_2(4/\alpha))^{C_\eta}} \cdot |\mathcal{V}_i|$, and X_i, Y_i, Z_i are (r, ϵ) -algebraically spread within $x_i + \mathcal{V}_i, y_i + \mathcal{V}_i, x_i + y_i + \mathcal{V}_i$ respectively.

The proof of the above proposition is based on [JLL⁺25, Section 5.3], with the exception that we work to make all the 3 sets spread (instead of just 2). As a result, we obtain a weaker dependence on the parameter η compared to their two set version; however, this is inconsequential in our setting, as we use this only for constant η . In the following subsections, we first show how to decompose one set, then two sets, and finally three sets.

4.2.1 Uniformization for 1 Set

We first show how to decompose a single set X into spread components. Before that, we show that any large set contains a spread set in it.

Lemma 4.4 (Existence of a Spread Subset). *Let $r \in \mathbb{N}$, $\epsilon \in (0, 1)$, and let $X \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$. Define $\alpha_X = |X| / |\mathcal{V}|$. Then, there exists an affine subspace $\mathcal{V}' \subseteq \mathcal{V}$ such that:*

¹⁰For example, if both A, B are algebraically spread, we have $|\langle \varphi_A * \varphi_B, \varphi_C \rangle - 1| \leq 2 \|\varphi_A * \varphi_B - 1\|_d \leq \epsilon$, where the first inequality follows from Hölder's inequality, and the second inequality follows from [KM23, Proposition 2.16].

1. $\dim(\mathcal{V}') \geq \dim(\mathcal{V}) - r\epsilon^{-1} \log_2(1/\alpha_X)$.

2. The set $X' = X \cap \mathcal{V}'$ satisfies $\frac{|X'|}{|\mathcal{V}'|} \geq \alpha_X$, and X' is (r, ϵ) -algebraically spread within \mathcal{V}' .

Proof. We proceed iteratively. Let $\mathcal{V}^{(0)} = \mathcal{V}$; for $t = 0, 1, 2, \dots$ we do the following: if $X \cap \mathcal{V}^{(t)}$ is (r, ϵ) -algebraically spread within $\mathcal{V}^{(t)}$, we stop; otherwise, there exists some subspace $\mathcal{V}^{(t+1)} \subseteq \mathcal{V}^{(t)}$ such that $\dim(\mathcal{V}^{(t+1)}) \geq \dim(\mathcal{V}^{(t)}) - r$, and

$$\frac{|X \cap \mathcal{V}^{(t+1)}|}{|\mathcal{V}^{(t+1)}|} > (1 + \epsilon) \cdot \frac{|X \cap \mathcal{V}^{(t)}|}{|\mathcal{V}^{(t)}|}.$$

If we have not stopped till step t , we have $\dim(\mathcal{V}^{(t)}) \geq \dim(\mathcal{V}) - rt$, and

$$\frac{|X \cap \mathcal{V}^{(t+1)}|}{|\mathcal{V}^{(t+1)}|} \geq (1 + \epsilon)^t \alpha_X \geq 2^{\epsilon t} \alpha_X.$$

Since this density cannot exceed 1, we must stop for some $t \leq \epsilon^{-1} \log_2(1/\alpha_X)$. \square

Using the above lemma, we show how to decompose a single set into spread components.

Lemma 4.5 (Uniformization for 1 Set). *Let $r \in \mathbb{N}$, $\epsilon, \eta \in (0, 1)$. Let $X \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$. Then, there exists $T \in \mathbb{N}$, and for each $i \in [T]$, an affine subspace $\mathcal{V}_i \subseteq \mathcal{V}$ and a subset $X_i \subseteq \mathcal{V}_i$, such that:*

1. $\dim(\mathcal{V}_i) \geq \dim(\mathcal{V}) - r\epsilon^{-1} \log_2(1/\eta)$ for all $i \in [T]$.
2. X_1, \dots, X_T are disjoint subsets of X such that $|X \setminus \cup_{i \in [T]} X_i| \leq \eta |\mathcal{V}|$.
3. For all $i \in [T]$, $|X_i| \geq \eta |\mathcal{V}_i|$ and X_i is (r, ϵ) -algebraically spread within \mathcal{V}_i .

Proof. We proceed iteratively, via the following algorithm: Let $X^{(1)} = X$, $\mathcal{V}^{(1)} = \mathcal{V}$. At any time step t , if $|X^{(t)}| \leq \eta |\mathcal{V}|$, we stop. Else, if $|X^{(t)}| \geq \eta |\mathcal{V}|$, by Lemma 4.4, we find an affine subspace $\mathcal{V}_t \subseteq \mathcal{V}$ of dimension $\dim(\mathcal{V}_t) \geq \dim(\mathcal{V}) - r\epsilon^{-1} \log_2(1/\eta)$, such that $X_t := X^{(t)} \cap \mathcal{V}_t$ satisfies $|X_t| \geq \eta |\mathcal{V}_t|$, and X_t is (r, ϵ) -algebraically spread within \mathcal{V}_t . Then, we define $X^{(t+1)} = X^{(t)} \setminus X_t$.

Finally, suppose we stop at step $T + 1$. Then, we can write X as a disjoint union $X = X^{(T+1)} \cup X_1 \cup X_2 \cup \dots \cup X_T$, with $|X^{(T+1)}| \leq \eta |\mathcal{V}|$. \square

Remarks 4.6. Note that

1. It must hold that $T \leq 2^{(1+r/\epsilon) \log_2(1/\eta)}$, since for all $i \in [T]$,

$$\frac{|X_i|}{|\mathcal{V}|} = \frac{|X_i|}{|\mathcal{V}_i|} \cdot \frac{|\mathcal{V}_i|}{|\mathcal{V}|} \geq \eta \cdot 2^{-r \log_2(1/\eta)/\epsilon} = 2^{-(1+r/\epsilon) \log_2(1/\eta)},$$

and $\sum_{i=1}^T |X_i| \leq |X| \leq |\mathcal{V}|$.

2. It is useful to think of $\eta = \eta' \cdot \frac{|X|}{|\mathcal{V}|}$, to get a multiplicative approximation for the set X .

4.2.2 Uniformization for 2 Sets

Next, we state a lemma about decomposing the product $X \times Y$ of two sets X, Y into spread components.

Lemma 4.7 (Uniformization for 2 Sets, [JLL⁺25, Lemma 5.9]). *Let $r \in \mathbb{N}$, $\epsilon, \eta \in (0, 1/10)$, and let $X, Y \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$. Define $\alpha_{XY} = |X||Y|/|\mathcal{V}|^2$. Then, there exists $T \in \mathbb{N}$, and for each $i \in [T]$, a linear subspace $\mathcal{V}_i \subseteq \mathcal{V}$, points $x_i, y_i \in \mathcal{V}/\mathcal{V}_i$, and subsets $X_i \subseteq x_i + \mathcal{V}_i$, $Y_i \subseteq y_i + \mathcal{V}_i$, such that:*

1. $\dim(\mathcal{V}_i) \geq \dim(\mathcal{V}) - O(r\epsilon^{-2} \log_2(1/\alpha_{XY})^2 \log_2(1/\eta) + r\epsilon^{-2} \log_2(1/\eta)^5)$ for all $i \in [T]$.
2. $X_1 \times Y_1, \dots, X_T \times Y_T$ are disjoint subsets of $X \times Y$ such that

$$|(X \times Y) \setminus \cup_{i \in [T]} (X_i \times Y_i)| \leq \eta |X \times Y|$$

3. For all $i \in [T]$, we have $|X_i \times Y_i| \geq \kappa \cdot |\mathcal{V}_i|^2$ for $\kappa \geq 2^{-O(\log_2(1/\eta)^2)} \cdot \alpha_{XY}$, and X_i, Y_i are (r, ϵ) -algebraically spread within $x_i + \mathcal{V}_i, y_i + \mathcal{V}_i$ respectively.

4.2.3 Uniformization for 3 Sets

We shall prove Proposition 4.3 and show how to decompose three sets into spread components, compatible with the diagonal-product. We start by stating a simple fact about the size of a diagonal-product $S(X, Y, Z)$, which along with Theorem 4.2 allows us to control this size when at least one or two of X, Y, Z are spread.

Fact 4.8. *Let $X, Y, Z \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$, and let $\alpha_X = |X|/|\mathcal{V}|$, $\alpha_Y = |Y|/|\mathcal{V}|$, $\alpha_Z = |Z|/|\mathcal{V}|$. Then, $|S(X, Y, Z)| = \alpha_X \alpha_Y \alpha_Z \cdot \langle \varphi_X * \varphi_Y, \varphi_Z \rangle$.*

First, we show a one round partitioning result for $S(X, Y, Z)$:

Lemma 4.9. *Let $\eta \in (0, 1/50)$, and let $X, Y, Z \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$. Define $\alpha = |S(X, Y, Z)|/|\mathcal{V}|^2$. Then, there exists an integer $r' = O(\log_2(1/(\eta\alpha))^{16})$, and constant $\epsilon' = \Omega(1)$, such that the following hold:*

Let $r \in \mathbb{N}$, $\epsilon \in (0, 1/10)$ satisfy $r \geq r'$, $\epsilon \leq \epsilon'$. Then, there exists $T \in \mathbb{N}$, a set $\mathcal{G} \subseteq [T]$, and for each $i \in [T]$, a linear subspace $\mathcal{V}_i \subseteq \mathcal{V}$, points $x_i, y_i \in \mathcal{V}/\mathcal{V}_i$, and subsets $X_i \subseteq x_i + \mathcal{V}_i$, $Y_i \subseteq y_i + \mathcal{V}_i$, $Z_i \subseteq x_i + y_i + \mathcal{V}_i$, such that

1. $\dim(\mathcal{V}_i) \geq \dim(\mathcal{V}) - O(r\epsilon^{-3} \log_2(1/(\eta\alpha))^6)$ for all $i \in [T]$.
2. $S(X_1, Y_1, Z_1), \dots, S(X_T, Y_T, Z_T)$ are disjoint subsets of $S(X, Y, Z)$ such that

$$|S(X, Y, Z) \setminus \cup_{i=1}^T S(X_i, Y_i, Z_i)| \leq 4\eta |S(X, Y, Z)|.$$

3. $|S(X_i, Y_i, Z_i)| \geq 2^{-O(\log_2(1/(\eta\alpha))^2)} \cdot |\mathcal{V}_i|^2$ for all $i \in [T]$.
4. For all $i \in \mathcal{G}$, the sets X_i, Y_i, Z_i are (r, ϵ) -algebraically spread within $x_i + \mathcal{V}_i, y_i + \mathcal{V}_i, x_i + y_i + \mathcal{V}_i$ respectively.

$$5. \sum_{i \in \mathcal{G}} |S(X_i, Y_i, Z_i)| \geq \frac{1}{10} \cdot |S(X, Y, Z)|.$$

Proof. Let $\kappa = 2^{-O(\log_2(1/(\eta\alpha))^2)} \cdot \alpha_{XY} \geq 2^{-O(\log_2(1/(\eta\alpha))^2)}$ be as in the third item in Lemma 4.7, when applied with parameter $\eta\alpha$; we may assume $\kappa \leq \eta\alpha/2$. Let $r' = O(\log_2(1/\kappa)^8)$, $\epsilon' = \Omega(1)$ be as in Theorem 4.2 with density $\kappa/2$ and error $1/10$ (that is, with $d \leftarrow \log_2(2/\kappa)$ and $\epsilon \leftarrow 1/10$).

Let $r_0 = r + \lceil r\epsilon^{-1} \log_2(1/(\eta\alpha)) \rceil$. By Lemma 4.7 applied to $X \times Y$, with parameters $r_0, \epsilon/10, \eta\alpha$, we may find $T \in \mathbb{N}$, and for each $t \in [T]$ a linear subspace \mathcal{V}_t of dimension $\dim(\mathcal{V}_t) \geq \dim(\mathcal{V}) - O(r_0\epsilon^{-2} \log_2(1/(\eta\alpha))^5)$, points $x_t, y_t \in \mathcal{V}/\mathcal{V}_t$, and subsets $X_t \subseteq x_t + \mathcal{V}_t, Y_t \subseteq y_t + \mathcal{V}_t$ of size $|X_t| |Y_t| / |\mathcal{V}_t|^2 \geq \kappa$, such that X_t, Y_t are $(r_0, \epsilon/10)$ -algebraically spread inside $x_t + \mathcal{V}_t, y_t + \mathcal{V}_t$ respectively. Also, we have that $X_1 \times Y_1, \dots, X_T \times Y_T$ are disjoint subsets of $X \times Y$, with $XY = (X \times Y) \setminus \bigcup_{t \in [T]} (X_t \times Y_t)$ satisfying $|XY| \leq \eta\alpha |X \times Y|$. In particular, it also holds that $S(X_1, Y_1, Z), \dots, S(X_T, Y_T, Z)$ are disjoint subsets of $S(X, Y, Z)$.

Now, consider any $t \in [T]$. Observe that $S(X_t, Y_t, Z) = S(X_t, Y_t, Z_t)$, where $Z_t = Z \cap (x_t + y_t + \mathcal{V}_t)$. We shall now further decompose $S(X_t, Y_t, Z_t)$ so as to make the third set spread. By Lemma 4.5 applied to Z_t ,¹¹ with parameters $r, \epsilon, \eta\alpha$, we can find $T_t \in \mathbb{N}$, and for each $t' \in [T_t]$, a linear subspace $\mathcal{V}_{t,t'} \subseteq \mathcal{V}_t$ of dimension $\dim(\mathcal{V}_{t,t'}) \geq \dim(\mathcal{V}_t) - r\epsilon^{-1} \log_2(1/(\eta\alpha))$, a point $z_{t,t'} \in \mathcal{V}_t/\mathcal{V}_{t,t'}$, and a subset $Z_{t,t'} \subseteq x_t + y_t + z_{t,t'} + \mathcal{V}_{t,t'}$ of size $|Z_{t,t'}| \geq \eta\alpha |\mathcal{V}_{t,t'}|$ that is (r, ϵ) -algebraically spread. Also, we may write Z_t as a disjoint union $Z_t = Z_{t,1} \cup \dots \cup Z_{t,T_t} \cup \tilde{Z}_t$, with $|\tilde{Z}_t| \leq \eta\alpha |\mathcal{V}_t|$.

Note that for all $t \in [T], t' \in [T_t]$, we have $S(X_t, Y_t, Z_{t,t'}) = \bigcup_{z \in \mathcal{V}_t/\mathcal{V}_{t,t'}} S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})$, where $X_{t,t',z} = X_t \cap (x_t + z + \mathcal{V}_{t,t'})$, $Y_{t,t',z} = Y_t \cap (y_t + z + \mathcal{V}_{t,t'})$; recall $Z_{t,t'} \subseteq x_t + y_t + z_{t,t'} + \mathcal{V}_{t,t'}$.

Finally, we define the sets X_i, Y_i, Z_i in the lemma statement as all $X_{t,t',z}, Y_{t,t',z}, Z_{t,t'}$, for $t \in [T], t' \in [T_t], z \in \mathcal{V}_t/\mathcal{V}_{t,t'}$ that satisfy

$$|S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})| \geq \eta^2 \alpha^2 \kappa \cdot |\mathcal{V}_{t,t'}|^2.$$

Also, we define \mathcal{G} to contain all $X_{t,t',z}, Y_{t,t',z}, Z_{t,t'}$ satisfying

$$\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|}, \quad \frac{|Y_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|Y_t|}{|\mathcal{V}_t|}.$$

We will show later (in the last item below) that the pieces in the set \mathcal{G} also satisfy the condition $|S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})| \geq \eta^2 \alpha^2 \kappa \cdot |\mathcal{V}_{t,t'}|^2$. Now, we verify each of the conclusions in the lemma statement:

1. We have for all t, t' that

$$\begin{aligned} \dim(\mathcal{V}_{t,t'}) &\geq \dim(\mathcal{V}_t) - r\epsilon^{-1} \log_2(1/(\eta\alpha)) \\ &\geq \dim(\mathcal{V}) - O(r_0\epsilon^{-2} \log_2(1/(\eta\alpha))^5) \\ &\geq \dim(\mathcal{V}) - O(r\epsilon^{-3} \log_2(1/(\eta\alpha))^6). \end{aligned}$$

¹¹Formally, we shift the set Z_t so it lies in the linear subspace \mathcal{V}_t , then apply the lemma and shift back.

2. By construction, it holds that the sets $S(X_i, Y_i, Z_i)$ are disjoint subsets of $S(X, Y, Z)$. Let \mathcal{I} to denote all (t, t', z) such that $|S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})| < \eta^2 \alpha^2 \kappa \cdot |\mathcal{V}_{t,t'}|^2$. Then, the quantity $|S(X, Y, Z)| - \sum_i |S(X_i, Y_i, Z_i)|$ equals

$$|\widetilde{XY} \cap \{(x, y) : x + y \in Z\}| + \sum_t |S(X_t, Y_t, \tilde{Z}_t)| + \sum_{(t,t',z) \in \mathcal{I}} |S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})|.$$

We bound each of these one by one.

- (a) The first term is at most $|\widetilde{XY}| \leq \eta \alpha |X| |Y| \leq \eta \alpha |\mathcal{V}|^2$.
- (b) For any $t \in [T]$, we have that X_t, Y_t are $(r_0, \epsilon/10)$ -algebraically spread, and hence also (r, ϵ) -algebraically spread. Thus, by Theorem 4.2, we have $|S(X_t, Y_t, \tilde{Z}_t)| \leq 1.1 \cdot |X_t| |Y_t| \cdot \eta \alpha$. Here, we used that X_t, Y_t have density at least κ , and also that for this calculation we may assume that \tilde{Z}_t has density at least κ (which is at most $\eta \alpha / 2$) by possibly adding more elements to \tilde{Z}_t . Hence, the second term is at most

$$\sum_{t \in [T]} |S(X_t, Y_t, \tilde{Z}_t)| \leq 1.1 \cdot \eta \alpha \cdot \sum_{t \in [T]} |X_t| |Y_t| \leq 1.1 \cdot \eta \alpha \cdot |X| |Y| \leq 1.1 \cdot \eta \alpha |\mathcal{V}|^2.$$

- (c) The third term is at most

$$\begin{aligned} \sum_{(t,t',z) \in \mathcal{I}} |S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})| &\leq \sum_{(t,t',z) \in \mathcal{I}} \eta^2 \alpha^2 \kappa \cdot |\mathcal{V}_{t,t'}|^2 \leq \sum_{\substack{t,t' \\ z \in \mathcal{V}_t / \mathcal{V}_{t,t'}}} \eta^2 \alpha^2 \kappa \cdot |\mathcal{V}_{t,t'}|^2 \\ &= \sum_{t,t'} \eta \alpha \kappa \cdot (\eta \alpha |\mathcal{V}_{t,t'}|) \cdot |\mathcal{V}_t| \leq \sum_{t,t'} \eta \alpha \kappa \cdot |Z_{t,t'}| \cdot |\mathcal{V}_t| \\ &\leq \sum_t \eta \alpha \cdot (\kappa |\mathcal{V}_t|^2) \leq \sum_t \eta \alpha \cdot |X_t| |Y_t| \leq \eta \alpha |\mathcal{V}|^2. \end{aligned}$$

Combining, we get the bound $3.1 \cdot \eta \alpha |\mathcal{V}|^2 \leq 4\eta |S(X, Y, Z)|$.

3. By construction, we have $|S(X_i, Y_i, Z_i)| \geq \eta^2 \alpha^2 \kappa |\mathcal{V}_i|^2 \geq 2^{-O(\log_2(1/(\eta \alpha))^2)} \cdot |\mathcal{V}_i|^2$.
4. Each $Z_{t,t'}$ is (r, ϵ) -algebraically spread within $x_t + y_t + z_{t,t'} + \mathcal{V}_{t,t'}$.

Now, consider any $X_{t,t',z}$ such that $\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|}$. We show that it is (r, ϵ) -algebraically spread within $x_t + z + \mathcal{V}_{t,t'}$. For this, let $\mathcal{W} \subseteq x_t + z + \mathcal{V}_{t,t'}$ be any affine subspace satisfying $\dim(\mathcal{W}) \geq \dim(\mathcal{V}_{t,t'}) - r \geq \dim(\mathcal{V}_t) - r_0$. Then, as X_t is $(r_0, \epsilon/10)$ -algebraically spread within $x_t + \mathcal{V}_t$, we get

$$\frac{|X_{t,t',z} \cap \mathcal{W}|}{|\mathcal{W}|} = \frac{|X_t \cap \mathcal{W}|}{|\mathcal{W}|} \leq \left(1 + \frac{\epsilon}{10}\right) \cdot \frac{|X_t|}{|\mathcal{V}_t|} \leq \frac{\left(1 + \frac{\epsilon}{10}\right)}{\left(1 - \frac{4\epsilon}{10}\right)} \cdot \frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \leq (1 + \epsilon) \cdot \frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|}.$$

A similar argument shows that $Y_{t,t',z}$ is (r, ϵ) -algebraically spread within $y_t + z + z_{t,t'} + \mathcal{V}_{t,t'}$, assuming that $\frac{|Y_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|Y_t|}{|\mathcal{V}_t|}$.

5. Consider any piece $X_{t,t',z}, Y_{t,t',z}, Z_{t,t'}$ in \mathcal{G} . This satisfies

$$\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|} \geq \frac{\kappa}{2}, \quad \frac{|Y_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|Y_t|}{|\mathcal{V}_t|} \geq \frac{\kappa}{2}, \quad \frac{|Z_{t,t'}|}{|\mathcal{V}_{t,t'}|} \geq \eta\alpha \geq \frac{\kappa}{2}.$$

Since all three are (r, ϵ) -algebraically spread, with $r \geq r', \epsilon \leq \epsilon'$, by Theorem 4.2,

$$\begin{aligned} |S(X_{t,t',z}, Y_{t,t',z}, Z_{t,t'})| &\geq \frac{9}{10} \cdot \frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \cdot \frac{|Y_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \cdot \frac{|Z_{t,t'}|}{|\mathcal{V}_{t,t'}|} \cdot |\mathcal{V}_{t,t'}|^2 \\ &\geq \frac{9}{10} \cdot \left(1 - \frac{4\epsilon}{10}\right)^2 \cdot \frac{|X_t|}{|\mathcal{V}_t|^2} \cdot \frac{|Y_t|}{|\mathcal{V}_t|^2} \cdot \frac{|Z_{t,t'}|}{|\mathcal{V}_{t,t'}|} \cdot |\mathcal{V}_{t,t'}|^2 \\ &\geq \frac{1}{2} \cdot \frac{|X_t|}{|\mathcal{V}_t|^2} \cdot |Z_{t,t'}| \cdot |\mathcal{V}_{t,t'}|. \end{aligned} \quad (1)$$

In particular, the above is at least $1/2 \cdot \kappa \cdot \eta\alpha \cdot |\mathcal{V}_{t,t'}|^2 \geq \eta^2 \alpha^2 \kappa \cdot |\mathcal{V}_{t,t'}|^2$, which shows that the pieces in \mathcal{G} are a subset of the pieces X_i, Y_i, Z_i that were not discarded.

Now, fix any $t \in [T], t' \in [T_t]$. We show that

$$\Pr_{z \sim \mathcal{V}_t/\mathcal{V}_{t,t'}} \left[\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|} \right] \geq 4/5.$$

For this, let p denote the probability on the left hand side. Observe that for every $z \in \mathcal{V}_t/\mathcal{V}_{t,t'}$, as X_t is $(r_0, \epsilon/10)$ -algebraically spread within $x_t + \mathcal{V}_t$, we have the upper bound $\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \leq (1 + \epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|}$. This implies that

$$\frac{|X_t|}{|\mathcal{V}_t|} = \mathbb{E}_{z \sim \mathcal{V}_t/\mathcal{V}_{t,t'}} \left[\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \right] \leq (1 - p) \cdot (1 - 4\epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|} + p \cdot (1 + \epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|},$$

which gives $p \geq 4/5$.

Thus, for any $t \in [T], t' \in [T_t]$, by the above, and by using a similar bound for $Y_{t,t',z}$,

$$\Pr_{z \sim \mathcal{V}_t/\mathcal{V}_{t,t'}} \left[\frac{|X_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|X_t|}{|\mathcal{V}_t|}, \quad \frac{|Y_{t,t',z}|}{|\mathcal{V}_{t,t'}|} \geq (1 - 4\epsilon/10) \cdot \frac{|Y_t|}{|\mathcal{V}_t|} \right] \geq 3/5.$$

Combining with the earlier bound (Equation 1), we get

$$\begin{aligned} \sum_{i \in \mathcal{G}} |S(X_i, Y_i, Z_i)| &\geq \sum_{t \in [T], t' \in [T_t]} \frac{1}{2} \cdot \frac{|X_t|}{|\mathcal{V}_t|^2} \cdot |Z_{t,t'}| \cdot |\mathcal{V}_{t,t'}| \cdot \left(\frac{3}{5} \cdot \frac{|\mathcal{V}_t|}{|\mathcal{V}_{t,t'}|} \right) \\ &= \sum_{t \in [T], t' \in [T_t]} \frac{3}{10} \cdot \frac{|X_t|}{|\mathcal{V}_t|} \cdot |Z_{t,t'}| \\ &\geq \sum_{t \in [T]} \frac{3}{10} \cdot \frac{|X_t|}{|\mathcal{V}_t|} \cdot (|Z_t| - \eta\alpha |\mathcal{V}_t|) \\ &= \sum_{t \in [T]} \frac{3}{10} \cdot \frac{|X_t|}{|\mathcal{V}_t|} \cdot |Z_t| - \eta\alpha \sum_{t \in [T]} |X_t| \\ &\geq \sum_{t \in [T]} \frac{3}{10} \cdot \frac{|X_t|}{|\mathcal{V}_t|} \cdot |Z_t| \cdot \mathbf{1}[|Z_t| \geq \eta\alpha |\mathcal{V}_t|] - \eta\alpha |\mathcal{V}|^2. \end{aligned}$$

Now, as X_t, Y_t are $(r_0, \epsilon/10)$ -algebraically spread within $x_t + \mathcal{V}_t, y_t + \mathcal{V}_t$ respectively, by the choice of r', ϵ' , and by Theorem 4.2, we have $|S(X_t, Y_t, Z_t)| \leq 1.1 \cdot \frac{|X_t||Y_t||Z_t|}{|\mathcal{V}_t|}$ whenever $|Z_t| \geq \eta\alpha |\mathcal{V}_t| \geq \kappa |\mathcal{V}_t|$. Hence,

$$\begin{aligned}
\sum_{i \in \mathcal{G}} |S(X_i, Y_i, Z_i)| &\geq \sum_{t \in [T]} \frac{3}{10} \cdot \frac{1}{1.1} \cdot |S(X_t, Y_t, Z_t)| \cdot \mathbb{1}[|Z_t| \geq \eta\alpha |\mathcal{V}_t|] - \eta\alpha |\mathcal{V}|^2 \\
&\geq 0.25 \cdot \sum_{t \in [T]} |S(X_t, Y_t, Z_t)| - \sum_{t \in [T]} |S(X_t, Y_t, Z_t)| \cdot \mathbb{1}[|Z_t| < \eta\alpha |\mathcal{V}_t|] - \eta\alpha |\mathcal{V}|^2 \\
&\geq 0.25 \cdot |S(X, Y, Z)| - \eta\alpha |\mathcal{V}|^2 - 1.1 \cdot \sum_{t \in [T]} |X_t| |Y_t| \eta\alpha - \eta\alpha |\mathcal{V}|^2 \\
&\geq 0.25 \cdot |S(X, Y, Z)| - 3.1\eta\alpha |\mathcal{V}|^2 \\
&= (0.25 - 3.1\eta) \cdot |S(X, Y, Z)| \geq 0.1 \cdot |S(X, Y, Z)|.
\end{aligned}$$

For the middle term $|S(X_t, Y_t, Z_t)| \cdot \mathbb{1}[|Z_t| < \eta\alpha]$, we used that X_t, Y_t are $(r_0, \epsilon/10)$ -algebraically spread within $x_t + \mathcal{V}_t, y_t + \mathcal{V}_t$ respectively, and also for the purpose of calculating an upper bound assumed that $|Z_t|/|\mathcal{V}_t| \geq \kappa$ which can be obtained by possibly adding more elements to Z_t . \square

We write down a simpler version of the above lemma, with the assumption on r, ϵ removed:

Corollary 4.10. *Let $r \in \mathbb{N}$, $\epsilon, \eta \in (1/10)$, and let $X, Y, Z \subseteq \mathcal{V}$ for a linear subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$. Define $\alpha = |S(X, Y, Z)| / |\mathcal{V}|^2$. Then, there exists $T \in \mathbb{N}$, a set $\mathcal{G} \subseteq [T]$, and for each $i \in [T]$, a linear subspace $\mathcal{V}_i \subseteq \mathcal{V}$, points $x_i, y_i \in \mathcal{V}/\mathcal{V}_i$, and subsets $X_i \subseteq x_i + \mathcal{V}_i, Y_i \subseteq y_i + \mathcal{V}_i, Z_i \subseteq x_i + y_i + \mathcal{V}_i$, such that*

$$1. \dim(\mathcal{V}_i) \geq \dim(\mathcal{V}) - O(r\epsilon^{-3} \log_2(1/(\eta\alpha))^6 + \epsilon^{-3} \log_2(1/(\eta\alpha))^{22}) \text{ for all } i \in [T].$$

$$2. S(X_1, Y_1, Z_1), \dots, S(X_T, Y_T, Z_T) \text{ are disjoint subsets of } S(X, Y, Z) \text{ such that}$$

$$|S(X, Y, Z) \setminus \cup_{i=1}^T S(X_i, Y_i, Z_i)| \leq \eta |S(X, Y, Z)|.$$

$$3. |S(X_i, Y_i, Z_i)| \geq 2^{-O(\log_2(1/(\eta\alpha))^2)} \cdot |\mathcal{V}_i|^2 \text{ for all } i \in [T].$$

$$4. \text{ For all } i \in \mathcal{G}, \text{ the sets } X_i, Y_i, Z_i \text{ are } (r, \epsilon)\text{-algebraically spread within } x_i + \mathcal{V}_i, y_i + \mathcal{V}_i, x_i + y_i + \mathcal{V}_i \text{ respectively.}$$

$$5. \sum_{i \in \mathcal{G}} |S(X_i, Y_i, Z_i)| \geq \frac{1}{10} \cdot |S(X, Y, Z)|.$$

Proof. We plug in the parameter $\eta/5$ in Lemma 4.9 to get $r' = O(\log_2(1/(\eta\alpha))^{16})$, and $\epsilon' = \Omega(1)$. Then, we use the lemma with parameters $r_0 = r + r'$ and $\epsilon_0 = \epsilon \cdot \epsilon'$, noting that any set which is (r_0, ϵ_0) -algebraically spread is also (r, ϵ) -algebraically spread. \square

Finally, we complete the main result of this section:

Proof of Proposition 4.3. We perform the following recursive process, given sets $X, Y, Z \subseteq \mathcal{V}$: Let $T \in \mathbb{N}$, $\mathcal{G} \subseteq [T]$, and subsets $X_i \subseteq x_i + \mathcal{V}_i, Y_i \subseteq y_i + \mathcal{V}_i, Z_i \subseteq x_i + y_i + \mathcal{V}_i$ be as in Corollary 4.10, with parameters $r, \frac{\epsilon}{10}, \frac{\eta^2}{100}$. Now, for each $i \in [T] \setminus \mathcal{G}$, we invoke Corollary 4.10 recursively on X_i, Y_i, Z_i ,¹² with parameters $r, \frac{\epsilon}{10}, \frac{\eta^2}{100}$. We terminate when the recursion depth is $L = \lceil 20 \log_2(1/\eta) \rceil$, and throw away all the remaining pieces at this point.

Note that for any X_i, Y_i, Z_i lying in affine shifts of a linear subspace $\mathcal{V}_i \subseteq \mathcal{V}$, that the algorithm sees at some point but does not throw away, it holds that

$$\frac{|S(X_i, Y_i, Z_i)|}{|\mathcal{V}_i|^2} \geq 2^{-(\log_2(4/\alpha))^{\text{poly}(1/\eta)}}.$$

This follows by Corollary 4.10, as the measure reduces from β to $2^{-O(\log_2(1/(\eta\beta))^2)}$ in a single step, and the number of steps is $L = O(\log_2(1/\eta))$. With this observation, we verify each of the conclusions in the statement of the proposition:

1. At each step, the dimension reduces by at most $r\epsilon^{-3} \log_2(4/\alpha)^{\text{poly}(1/\eta)}$, and the number of steps is $L = O(\log_2(1/\eta))$.
2. The disjointness of the pieces follows by construction.

At each recursive layer, the total size of pieces thrown out is at most $\frac{\eta^2}{100} \cdot |S(X, Y, Z)|$.

Additionally, at any layer the total size of the pieces being recursed on is at most 9/10 times the total size of pieces in the previous layer; in particular, the total size of pieces at layer L is at most $(9/10)^L \cdot |S(X, Y, Z)| \leq \eta^2 |S(X, Y, Z)|$. Hence, the total fraction of pieces thrown out is at most

$$\frac{\eta^2}{100} \cdot 21 \log_2(1/\eta) + \eta^2 \leq \eta.$$

3. The size bound for X_i, Y_i, Z_i is as above, and spreadness follows by Corollary 4.10. \square

5 Uniform Square Covers

In this section, we show that any *diagonal-product* set (see Definition 2.1) composed of algebraically spread sets (see Definition 4.1) is *uniformly covered by squares* (see Definition 2.2). In Section 5.1, we introduce a combinatorial notion of spreadness that will be useful for us; in Section 5.2, we show that algebraically spread sets satisfy this definition (in a specific way); finally, in Section 5.3, we prove Proposition 5.6, our square covering result.

5.1 Combinatorial Spreadness and Graph Counts

We state some useful definitions from [KLM24]:

¹²Formally, we shift the sets so they all in the linear subspace \mathcal{V}_i , then apply the corollary and shift back.

Definition 5.1 (Combinatorial Spreadness). *Let X, Y be finite sets, and let $f : X \times Y \rightarrow \{0, 1\}$. We say that f is (r, ϵ) -combinatorially spread if for every $S \subseteq X$, $T \subseteq Y$ satisfying $|S \times T| \geq 2^{-r} |X \times Y|$, it holds that*

$$\mathbb{E}_{(x,y) \sim S \times T} [f(x, y)] \leq (1 + \epsilon) \mathbb{E}[f],$$

where $\mathbb{E}[f] = \mathbb{E}_{(x,y) \sim X \times Y} [f(x, y)]$.

Definition 5.2 (Lower Bounded Left-Marginals). *Let X, Y be finite sets, and let $f : X \times Y \rightarrow \{0, 1\}$. We say that f has (r, ϵ) -lower bounded left-marginals if*

$$\Pr_{x \sim X} \left[\mathbb{E}_{y \sim Y} [f(x, y)] \leq (1 - \epsilon) \cdot \mathbb{E}[f] \right] \leq 2^{-r}.$$

We state a result about graph counts in spread sets:

Theorem 5.3 ([FHHK24, Theorem 2.1]). *For any $k \in \mathbb{N}$, $\epsilon \in (0, 1)$, there exists sufficiently small $\gamma = \gamma(\epsilon, k) > 0$, and sufficiently large $C = C(\epsilon, k) \in \mathbb{N}$, such that the following holds:*

Let $H = ([k], E)$ be an oriented graph, where $(i, j) \in E$ implies $i < j$. Let X_1, \dots, X_k be finite sets; let $d \geq 1$, and for each $(i, j) \in E$, let $f_{ij} : X_i \times X_j \rightarrow \{0, 1\}$ be a function satisfying:

1. $\mathbb{E}[f_{ij}] \geq 2^{-d}$,
2. f_{ij} is (Cd^2, γ) -combinatorially spread, and
3. f_{ij} has (Cd, γ) -lower bounded left-marginals.

Then, it holds that

$$\left| \mathbb{E}_{x_1 \sim X_1, \dots, x_k \sim X_k} \left[\prod_{(i,j) \in E} f_{ij}(x_i, x_j) \right] - \prod_{(i,j) \in E} \mathbb{E}[f_{ij}] \right| \leq \epsilon \cdot \prod_{(i,j) \in E} \mathbb{E}[f_{ij}].$$

5.2 Algebraic Spreadness implies Graph Counts

We show that a certain function corresponding to algebraically spread sets is both combinatorially spread and has lower bounded left-marginals.

Lemma 5.4. *Let $d, r \geq 1, \epsilon \in (0, 1/4)$. Then, there exists a sufficiently large integer $s = (r^8 + d^8) \cdot \epsilon^{-O(1)}$, and a sufficiently small $\delta = \Omega(\epsilon)$, such that the following holds:*

Let $X, Y, Z \subseteq \mathbb{F}_2^n$ be subsets each of density at least 2^{-d} , and such that Y, Z are (s, δ) -algebraically spread. Then, the function $f : X \times Y \rightarrow \{0, 1\}$, given by $f(x, y) = \mathbf{1}[x + y \in Z]$, satisfies:

1. $|\mathbb{E}[f] - \alpha_Z| \leq \epsilon \cdot \alpha_Z$, where $\alpha_Z = |Z|/2^n$,
2. f is (r, ϵ) -combinatorially spread, and

3. f has (r, ϵ) -lower bounded left-marginals.

Proof. Let s, δ be such that Theorem 4.2 holds with parameters $r + d, \epsilon/4$.

First, we have

$$\mathbb{E}[f] = \Pr_{x \sim X, y \sim Y}[x + y \in Z] = \langle \varphi_X * \varphi_Y, \mathbb{1}_Z \rangle = \alpha_Z \cdot \langle \varphi_X * \varphi_Y, \varphi_Z \rangle.$$

Hence, by Theorem 4.2(i), it holds that

$$|\mathbb{E}[f] - \alpha_Z| \leq \frac{\epsilon}{4} \cdot \alpha_Z.$$

Next, we show combinatorial spreadness. Let $S \subseteq X$, $T \subseteq Y$ be such that $|S \times T| \geq 2^{-r} |X \times Y|$. We can write

$$\mathbb{E}_{(x,y) \sim S \times T}[f(x, y)] = \alpha_Z \cdot \langle \varphi_S * \varphi_T, \varphi_Z \rangle.$$

Since $|X| \geq 2^{-d} \cdot 2^n$, we have $|S| \cdot |Y| \geq |S \times T| \geq 2^{-r} \cdot |X \times Y| \geq 2^{-r} \cdot 2^{n-d} \cdot |Y|$, and so $|S| \geq 2^{-(r+d)} \cdot 2^n$. Similarly, $|T| \geq 2^{-(r+d)} \cdot 2^n$, and hence, by Theorem 4.2(ii), we get

$$\mathbb{E}_{(x,y) \sim S \times T}[f(x, y)] \leq \left(1 + \frac{\epsilon}{4}\right) \alpha_Z \leq \left(1 + \frac{\epsilon}{4}\right) \cdot \left(1 - \frac{\epsilon}{4}\right)^{-1} \mathbb{E}[f] \leq (1 + \epsilon) \cdot \mathbb{E}[f].$$

Finally, we show lower bounded left-marginals. Define $A \subseteq X$ by

$$A = \left\{ x \in X : \mathbb{E}_{y \sim Y}[f(x, y)] \leq (1 - \epsilon) \cdot \mathbb{E}[f] \right\}.$$

Suppose, for the sake of contradiction, that $|A| > 2^{-r} |X|$; then, $|A| \geq 2^{-(r+d)} \cdot 2^n$. By the definition of A , we have

$$\mathbb{E}_{x \sim A, y \sim Y}[f(x, y)] \leq (1 - \epsilon) \mathbb{E}[f].$$

On the other hand, by Theorem 4.2(i), we have

$$\mathbb{E}_{x \sim A, y \sim Y}[f(x, y)] = \alpha_Z \cdot \langle \varphi_A * \varphi_Y, \varphi_Z \rangle \geq \left(1 - \frac{\epsilon}{4}\right) \alpha_Z \geq \left(1 - \frac{\epsilon}{4}\right) \cdot \left(1 + \frac{\epsilon}{4}\right)^{-1} \mathbb{E}[f] > (1 - \epsilon) \mathbb{E}[f],$$

which is a contradiction. \square

5.3 Squares inside Spread Diagonal-Product Sets

In this subsection, we prove our square covering result. We start by proving the following counting lemma:

Lemma 5.5. *Let $\epsilon \in (0, 1/4)$ be a constant, and let $d \geq 1$. Then, there exists a sufficiently large integer $r = O_\epsilon(d^{16})$ and sufficiently small $\delta = \delta(\epsilon) > 0$ such that the following holds:*

Let $X, Y, Z \subseteq \mathbb{F}_2^n$ be each of density at least 2^{-d} , and such that all of X, Y, Z are (r, δ) -algebraically spread. Let $S = S(X, Y, Z)$ be as in Definition 2.1, and let $\Gamma : \mathbb{F}_2^n \times \mathbb{F}_2^n \rightarrow \mathbb{R}$ be the function mapping (x, y) to the normalized number of squares (see Definition 2.2) in S containing (x, y) , i.e.,

$$\Gamma(x, y) = \mathbb{E}_{w \sim \mathbb{F}_2^n} [\mathbb{1}[s_{x,y,w} \subseteq S]].$$

Define $\alpha_X = \mathbb{E}[\mathbb{1}_X]$, $\alpha_Y = \mathbb{E}[\mathbb{1}_Y]$, $\alpha_Z = \mathbb{E}[\mathbb{1}_Z]$. Then, it holds that

$$1. \quad ||S| \cdot 2^{-2n} - \alpha_X \alpha_Y \alpha_Z| \leq \epsilon \cdot \alpha_X \alpha_Y \alpha_Z.$$

$$2. \quad ||\Gamma||_1 - \alpha_X^2 \alpha_Y^2 \alpha_Z^2 \leq \epsilon \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z^2.$$

$$3. \quad ||\Gamma||_2^2 - \alpha_X^3 \alpha_Y^3 \alpha_Z^3 \leq \epsilon \cdot \alpha_X^3 \alpha_Y^3 \alpha_Z^3.$$

Proof. Let $f : X \times Y \rightarrow \{0, 1\}$, $g : Y \times Z \rightarrow \{0, 1\}$, $h : X \times Z \rightarrow \{0, 1\}$ be given by

$$f(x, y) = \mathbb{1}[x + y \in Z], \quad g(y, z) = \mathbb{1}[y + z \in X], \quad h(x, z) = \mathbb{1}[x + z \in Y].$$

Let $\gamma = \gamma(\frac{\epsilon}{10}, 4) \leq \frac{\epsilon}{10}$, $C = C(\frac{\epsilon}{10}, 4)$ be as in Theorem 5.3. Then, by Lemma 5.4 and the choice of parameters (r, δ) , we may assume that f (and similarly g, h) is such that $|\mathbb{E}[f] - \alpha_Z| \leq \gamma \alpha_Z \leq \frac{\epsilon}{10} \alpha_Z$, and f is (Cd^2, γ) -combinatorially spread, and f has (Cd, γ) -lower bounded left-marginals.

First, observe that

$$|S| \cdot 2^{-2n} = \mathbb{E}_{x, y \sim \mathbb{F}_2^n} [\mathbb{1}[x \in X, y \in Y, x + y \in Z]] = \alpha_X \alpha_Y \Pr_{x \sim X, y \sim Y} [x + y \in Z] = \alpha_X \alpha_Y \mathbb{E}[f].$$

The result now follows, as $|\mathbb{E}[f] - \alpha_Z| \leq \gamma \alpha_Z \leq \frac{\epsilon}{10} \alpha_Z$.

Next, we note that

$$\begin{aligned} ||\Gamma||_1 &= \mathbb{E}_{x, y, w \sim \mathbb{F}_2^n} [\mathbb{1}[s_{x, y, w} \subseteq S]] \\ &= \mathbb{E}_{x, y, w \sim \mathbb{F}_2^n} [\mathbb{1}_X(x) \mathbb{1}_X(x + w) \mathbb{1}_Y(y) \mathbb{1}_Y(y + w) \mathbb{1}_Z(x + y) \mathbb{1}_Z(x + y + w)] \\ &= \mathbb{E}_{x, y, z \sim \mathbb{F}_2^n} [\mathbb{1}_X(x) \mathbb{1}_X(y + z) \mathbb{1}_Y(y) \mathbb{1}_Y(x + z) \mathbb{1}_Z(x + y) \mathbb{1}_Z(z)] \quad (\text{replace } w = x + y + z) \\ &= \alpha_X \alpha_Y \alpha_Z \mathbb{E}_{x \sim X, y \sim Y, z \sim Z} [f(x, y) g(y, z) h(x, z)]. \end{aligned}$$

Now, by Theorem 5.3 (with the oriented graph $\{(1, 2), (1, 3), (2, 3)\}$, with 1, 2, 3 labelled by X, Y, Z respectively), we get

$$\begin{aligned} (1 - \epsilon) \cdot \alpha_X \alpha_Y \alpha_Z &\leq \left(1 - \frac{\epsilon}{10}\right)^4 \cdot \alpha_X \alpha_Y \alpha_Z \\ &\leq \left(1 - \frac{\epsilon}{10}\right) \cdot \mathbb{E}[f] \mathbb{E}[g] \mathbb{E}[h] \\ &\leq \mathbb{E}_{x \sim X, y \sim Y, z \sim Z} [f(x, y) g(y, z) h(x, z)] \\ &\leq \left(1 + \frac{\epsilon}{10}\right) \cdot \mathbb{E}[f] \mathbb{E}[g] \mathbb{E}[h] \\ &\leq \left(1 + \frac{\epsilon}{10}\right)^4 \cdot \alpha_X \alpha_Y \alpha_Z \leq (1 + \epsilon) \cdot \alpha_X \alpha_Y \alpha_Z, \end{aligned}$$

and hence

$$||\Gamma||_1 - \alpha_X^2 \alpha_Y^2 \alpha_Z^2 \leq \epsilon \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z^2.$$

Finally, we can write

$$\begin{aligned}
\|\Gamma\|_2^2 &= \mathbb{E}_{x,y,w,w' \sim \mathbb{F}_2^n} [\mathbb{1}[s_{x,y,w} \subseteq S, s_{x,y,w'} \subseteq S]] \\
&= \mathbb{E}_{x,y,z,z' \sim \mathbb{F}_2^n} [\mathbb{1}[s_{x,y,x+y+z} \subseteq S, s_{x,y,x+y+z'} \subseteq S]] \quad (\text{replace } w = x + y + z, w' = x + y + z') \\
&= \mathbb{E}_{x,y,z,z' \sim \mathbb{F}_2^n} [\mathbb{1}_X(x) \mathbb{1}_X(y+z) \mathbb{1}_X(y+z') \mathbb{1}_Y(y) \mathbb{1}_Y(x+z) \mathbb{1}_Y(x+z') \mathbb{1}_Z(x+y) \mathbb{1}_Z(z) \mathbb{1}_Z(z')] \\
&= \alpha_X \alpha_Y \alpha_Z^2 \mathbb{E}_{x \sim X, y \sim Y, z, z' \sim Z} [f(x, y) g(y, z) g(y, z') h(x, z) h(x, z')].
\end{aligned}$$

Now, by Theorem 5.3 (with the oriented graph $\{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4)\}$, with 1, 2, 3, 4 labelled by X, Y, Z, Z respectively), we get

$$\begin{aligned}
(1 - \epsilon) \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z &\leq \left(1 - \frac{\epsilon}{10}\right)^6 \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z \\
&\leq \left(1 - \frac{\epsilon}{10}\right) \cdot \mathbb{E}[f] \mathbb{E}[g]^2 \mathbb{E}[h]^2 \\
&\leq \mathbb{E}_{x \sim X, y \sim Y, z, z' \sim Z} [f(x, y) g(y, z) g(y, z') h(x, z) h(x, z')] \\
&\leq \left(1 + \frac{\epsilon}{10}\right) \cdot \mathbb{E}[f] \mathbb{E}[g]^2 \mathbb{E}[h]^2 \\
&\leq \left(1 + \frac{\epsilon}{10}\right)^6 \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z \leq (1 + \epsilon) \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z,
\end{aligned}$$

and hence

$$|\|\Gamma\|_2^2 - \alpha_X^3 \alpha_Y^3 \alpha_Z^3| \leq \epsilon \cdot \alpha_X^3 \alpha_Y^3 \alpha_Z^3. \quad \square$$

With the above, we are ready to prove the main result of this section.

Proposition 5.6. *Let $\epsilon \in (0, 1/4)$ be a constant, and let $d \geq 1$. Then, there exists a sufficiently large integer $r = O_\epsilon(d^{16})$ and sufficiently small $\delta = \delta(\epsilon) > 0$ such that the following holds:*

Let $X, Y, Z \subseteq \mathbb{F}_2^n$, be each of density at least 2^{-d} , and such that all of X, Y, Z are (r, δ) -algebraically spread. Let $S = S(X, Y, Z)$ be as in Definition 2.1, and let \mathcal{T} denote the set of all squares in S (see Definition 2.2); formally,

$$\mathcal{T} = \{(x, y, w) \in (\mathbb{F}_2^n)^3 : s_{x,y,w} \subseteq S\}.$$

Define $\alpha_X = \mathbb{E}[\mathbb{1}_X], \alpha_Y = \mathbb{E}[\mathbb{1}_Y], \alpha_Z = \mathbb{E}[\mathbb{1}_Z]$. Then, it holds that

1. $||S| \cdot 2^{-2n} - \alpha_X \alpha_Y \alpha_Z| \leq \epsilon \cdot \alpha_X \alpha_Y \alpha_Z.$
2. $||\mathcal{T}| \cdot 2^{-3n} - \alpha_X^2 \alpha_Y^2 \alpha_Z^2| \leq \epsilon \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z^2.$
3. *Let μ be the distribution on S obtained as follows: Pick a random square in S , by picking uniformly at random $(x, y, w) \sim \mathcal{T}$; then, output a uniformly random element from the set $s_{x,y,w}$. It holds that*

$$\|\mu - U_S\|_1 \leq \epsilon,$$

where U_S denotes the uniform distribution over S .

Proof. Let (r, δ) be such that Lemma 5.5 holds with parameters d and $\epsilon^2/5 := \gamma$, and let Γ be the function as in the statement of Lemma 5.5. By Lemma 5.5, it directly holds that

$$||S| \cdot 2^{-2n} - \alpha_X \alpha_Y \alpha_Z| \leq \gamma \cdot \alpha_X \alpha_Y \alpha_Z \leq \epsilon \cdot \alpha_X \alpha_Y \alpha_Z.$$

Now, observe that

$$|\mathcal{T}| = \sum_{x,y,w \in \mathbb{F}_2^n} \mathbb{1}[s_{x,y,w} \subseteq S] = 2^{3n} \|\Gamma\|_1,$$

and hence by Lemma 5.5,

$$||\mathcal{T}| \cdot 2^{-3n} - \alpha_X^2 \alpha_Y^2 \alpha_Z^2| \leq \gamma \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z^2 \leq \epsilon \cdot \alpha_X^2 \alpha_Y^2 \alpha_Z^2.$$

Finally, we recall Remark 2.3: every square $s_{x,y,w} \subseteq S$ with $w \neq 0$ occurs exactly 4 times in \mathcal{T} , as $s_{x,y,w} = s_{x+w,y,w} = s_{x,y+w,w} = s_{x+w,y+w,w}$. Also, when $w = 0$, we have $s_{x,y,0} = \{(x,y)\}$. Thus, for any $(x,y) \in S$,

$$\mu[(x,y)] = \frac{1}{|\mathcal{T}|} \cdot \left(\sum_{w \in \mathbb{F}_2^n, w \neq 0} 4 \cdot \mathbb{1}[s_{x,y,w} \subseteq S] \cdot \frac{1}{4} + \mathbb{1}[s_{x,y,0} \subseteq S] \cdot 1 \right) = \frac{\Gamma(x,y) \cdot 2^n}{|\mathcal{T}|}. \quad (2)$$

Hence, we can compute the ℓ_2 -norm of μ as

$$\begin{aligned} \|\mu\|_2^2 &= \sum_{(x,y) \in S} \mu[(x,y)]^2 \\ &= \frac{2^{2n}}{|\mathcal{T}|^2} \cdot \sum_{(x,y) \in S} \Gamma(x,y)^2 \\ &= \frac{2^{2n}}{|\mathcal{T}|^2} \cdot \sum_{x,y \in \mathbb{F}_2^n} \Gamma(x,y)^2 \quad (\text{as } \Gamma(x,y) = 0 \text{ for } (x,y) \notin S) \\ &= \frac{2^{4n}}{|\mathcal{T}|^2} \cdot \|\Gamma\|_2^2. \end{aligned}$$

Now, by Lemma 5.5, we obtain

$$\|\mu\|_2^2 \leq \frac{2^{4n}}{|\mathcal{T}|^2} \cdot (1 + \gamma) \cdot \alpha_X^3 \alpha_Y^3 \alpha_Z^3,$$

and so,

$$\begin{aligned} \|\mu\|_2^2 \cdot |S| &\leq 2^{4n} \cdot (1 + \gamma) \cdot \alpha_X^3 \alpha_Y^3 \alpha_Z^3 \cdot \frac{|S|}{|\mathcal{T}|^2} \\ &\leq 2^{4n} \cdot (1 + \gamma) \cdot \alpha_X^3 \alpha_Y^3 \alpha_Z^3 \cdot \frac{(1 + \gamma) \cdot \alpha_X \alpha_Y \alpha_Z \cdot 2^{2n}}{(1 - \gamma)^2 \cdot \alpha_X^4 \alpha_Y^4 \alpha_Z^4 \cdot 2^{6n}} \\ &= \frac{(1 + \gamma)^2}{(1 - \gamma)^2} \leq 1 + 5\gamma. \quad (\text{as } \gamma \leq 0.1) \end{aligned}$$

Finally, by Cauchy-Schwarz, we get

$$\|\mu - U_S\|_1 \leq (|S| \cdot \|\mu - U_S\|_2^2)^{1/2} = \left(|S| \cdot \left(\|\mu\|_2^2 - \frac{1}{|S|} \right) \right)^{1/2} \leq \sqrt{5\gamma} = \epsilon. \quad \square$$

6 Parallel Repetition for Games with the GHZ Query Support

This section is devoted to the proof of Theorem 1.1. By Lemma 3.12, it suffices to consider the case where Q is the uniform distribution over $\text{supp}(Q)$. We prove the following:

Theorem 6.1. *Let $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{\text{pred}})$ be any 3-player game with $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathbb{F}_2$, and with Q the uniform distribution over*

$$\text{supp}(Q) = \{(x, y, z) \in \mathbb{F}_2^3 : x + y + z = 0\},$$

and such that $\text{val}(\mathcal{G}) < 1$. Then, for all sufficiently large n ,² it holds that

$$\text{val}(\mathcal{G}^{\otimes n}) \leq \exp(-n^c),$$

where $c > 0$ is an absolute constant.

The remainder of this section is devoted to the proof of the above theorem. Towards this, we fix a 3-player game \mathcal{G} as in the statement of the theorem; in particular, we must have that $\text{val}(\mathcal{G}) \leq 3/4$.

The main step of the proof is to prove that the game $\mathcal{G}^{\otimes n}$ has a hard coordinate when the inputs to the players are conditioned on a product event with sufficiently large measure. We carry this out in two parts: first, in Section 6.1, we prove this statement assuming that each of the three sets in the product event are well spread (as in Definition 4.1) inside some large subspace of \mathbb{F}_2^n ; then, in Section 6.2, we prove the statement for general product events, using our uniformization strategy.

6.1 Hard Coordinate under Spread Product Events

We consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n, \mathcal{A}^n \times \mathcal{B}^n \times \mathcal{C}^n, Q^{\otimes n}, V_{\text{pred}}^{\otimes n})$ for some sufficiently large $n \in \mathbb{N}$. Fix any strategies for the 3 players in this game, and for each $i \in [n]$, let $\text{Win}_i \subseteq \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ be the event that this strategy wins the i^{th} coordinate of the game.

Lemma 6.2. *For any $1 \leq d \leq o(n)$, and constant $\epsilon \in (0, 1/4)$, there exists a (sufficiently large) integer $r = O_\epsilon(d^{16})$, and a (sufficiently small) constant $0 < \delta = \delta(\epsilon) < 1$, such that the following holds:*

Let $\mathcal{V} \subseteq \mathbb{F}_2^n$ be a linear subspace of dimension $\dim(\mathcal{V}) \geq n - o(n)$, and let $E, F, G \subseteq \mathcal{V}$ be such that each of them is (r, δ) -algebraically spread within \mathcal{V} (see Definition 4.1), and such that each of the densities $\frac{|E|}{|\mathcal{V}|}, \frac{|F|}{|\mathcal{V}|}, \frac{|G|}{|\mathcal{V}|} \geq 2^{-d}$. Then,

$$\mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] \leq \frac{1}{2} + \frac{\text{val}(\mathcal{G})}{2} + \epsilon.$$

The remainder of this subsection is devoted to the proof of the Lemma 6.2. We fix some $1 \leq d \leq o(n)$, and constant $\epsilon \in (0, 1/4)$, and let $r = O_\epsilon(d^{16})$, $\delta = \delta(\epsilon) > 0$ be so that Proposition 5.6 holds with the parameters $d, \frac{\epsilon}{2}$. Let $\mathcal{V} \subseteq \mathbb{F}_2^n$ be a linear subspace of

dimension $\dim(\mathcal{V}) \geq n - o(n)$, and let $E, F, G \subseteq \mathcal{V}$ be each (r, δ) -algebraically spread within \mathcal{V} , and such that each of the densities $\alpha_E = \frac{|E|}{|\mathcal{V}|}$, $\alpha_F = \frac{|F|}{|\mathcal{V}|}$, $\alpha_G = \frac{|G|}{|\mathcal{V}|}$ is at least 2^{-d} .

We start by proving a lemma that demonstrates the usefulness of squares. It says that the no strategy for the game $\mathcal{G}^{\otimes n}$ can win on non-trivial coordinates of squares.

Definition 6.3 (Non-trivial coordinates). *For a square $s = s_{x,y,w} \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$, as in Definition 2.2, we define its set of non-trivial coordinates as*

$$\mathcal{I}_s = \{i \in [n] : w_i \neq 0\}.$$

This is well-defined since w is uniquely determined by the square s (see Remark 2.3).

Lemma 6.4. *Let s be a square, and let $i \in \mathcal{I}_s$ be a non-trivial coordinate in s . Then,*

$$\Pr_{(x,y) \sim s} [(x, y, x+y) \in \text{Win}_i] \leq \text{val}(\mathcal{G}) < 1.$$

Proof. Let $x_0, y_0, w \in \mathbb{F}_2^n$ be such that $s = s_{x_0, y_0, w}$, and $w_i \neq 0$. We may assume that $(x_0)_i = 0$ (by possibly replacing x_0 with $x_0 + w$) and similarly $(y_0)_i = 0$. Define the points $x_1 = x_0 + w$, $y_1 = y_0 + w$, $z_0 = x_0 + y_0$, $z_1 = z_0 + w$. Also, define

$$\begin{aligned} \tilde{s} &= \{(x, y, x+y) : (x, y) \in s\} \\ &= \{(x_0, y_0, z_0), (x_0, y_1, z_1), (x_1, y_0, z_1), (x_1, y_1, z_0)\} \\ &= \{(x_\alpha, y_\beta, z_\gamma) : (\alpha, \beta, \gamma) \in \text{supp}(Q)\} \end{aligned}$$

Note that $\tilde{s} \subseteq \text{supp}(Q)^n$ and all points in \tilde{s} are valid inputs to the players in the game $\mathcal{G}^{\otimes n}$. Moreover, these satisfy $(x_\alpha, y_\beta, z_\gamma)_i = (\alpha, \beta, \gamma)$ for each $(\alpha, \beta, \gamma) \in \text{supp}(Q)$.

Observe that the distribution of $(x, y, x+y)$ for $(x, y) \sim s$ is simply the uniform distribution over the 4 points in \tilde{s} . Hence, it suffices to show that no strategies $f : \mathbb{F}_2^n \rightarrow \mathcal{A}$, $g : \mathbb{F}_2^n \rightarrow \mathcal{B}$, $h : \mathbb{F}_2^n \rightarrow \mathcal{C}$ for the i^{th} coordinate in $\mathcal{G}^{\otimes n}$ can win on more than $\text{val}(\mathcal{G})$ fraction of the points \tilde{s} . For this, consider any such functions f, g, h . Using these, we construct a strategy for the base game \mathcal{G} as follows:

1. The verifier chooses $(\alpha, \beta, \gamma) \sim Q$, and gives them to the three players respectively.
2. Player 1 outputs $f(x_\alpha)$, Player 2 outputs $g(y_\beta)$, and Player 3 outputs $h(z_\gamma)$.

On any input $(\alpha, \beta, \gamma) \in \text{supp}(Q)$, as observed before, the i^{th} coordinate of the vector $(x_\alpha, y_\beta, z_\gamma) \in \tilde{s}$ equals (α, β, γ) . Hence, probability that the above strategy wins the game \mathcal{G} equals the fraction of points in \tilde{s} on which f, g, h win the i^{th} coordinate of $\mathcal{G}^{\otimes n}$. By the definition of game value, this is at most $\text{val}(\mathcal{G})$. \square

Next, we show that the spreadness of $E, F, G \subseteq \mathcal{V}$ guarantees that $S = S(E, F, G)$ (see Definition 2.1) is covered uniformly by squares. For this, let \mathcal{T} denote the set of all squares contained in S ; formally,

$$\mathcal{T} = \{(x, y, w) \in \mathcal{V}^3 : s_{x,y,w} \subseteq S\}.$$

Let U_S denote the uniform distribution over S . Also, let μ be the distribution on S obtained as follows: Pick a random square in S , by picking uniformly at random $(x, y, w) \sim \mathcal{T}$; then, output a uniformly random element from the set $s_{x,y,w}$.

Lemma 6.5. *We have that*

1. $\left| \frac{|S|}{|\mathcal{V}|^2} - \alpha_E \alpha_F \alpha_G \right| \leq \frac{\epsilon}{2} \cdot \alpha_E \alpha_F \alpha_G.$
2. $\left| \frac{|\mathcal{T}|}{|\mathcal{V}|^3} - \alpha_E^2 \alpha_F^2 \alpha_G^2 \right| \leq \frac{\epsilon}{2} \cdot \alpha_E^2 \alpha_F^2 \alpha_G^2.$
3. $\|\mu - U_S\|_1 \leq \frac{\epsilon}{2}.$

Proof. This follows directly by Proposition 5.6, applied with respect to $\mathcal{V} \cong \mathbb{F}_2^{\dim(\mathcal{V})}$. Note that the definition of squares is *coordinate-free*, and hence the proposition is applicable. \square

Before completing the proof of Lemma 6.2, we show that a random square inside $S = S(E, F, G)$ contains many non-trivial coordinates.

Lemma 6.6. *Let $(x, y, w) \sim \mathcal{T}$, and let $s = s_{x,y,w}$. Then,*

$$\mathbb{E} |\mathcal{I}_s| \geq \frac{n}{2} \cdot (1 - o(1)).$$

Proof. For any $0 < \theta < 1$, we have

$$\begin{aligned} \Pr \left[|\mathcal{I}_s| \leq (1 - \theta) \cdot \frac{n}{2} \right] &= \Pr_{(x,y,w) \sim \mathcal{T}} [|\{i : w_i \neq 0\}| \leq (1 - \theta) \cdot n/2] \\ &= \frac{1}{|\mathcal{T}|} \sum_{(x,y,w) \in \mathcal{T}} \mathbb{1} [|\{i : w_i \neq 0\}| \leq (1 - \theta) \cdot n/2] \\ &\leq \frac{1}{|\mathcal{T}|} \sum_{x,y,w \in \mathbb{F}_2^n} \mathbb{1} [|\{i : w_i \neq 0\}| \leq (1 - \theta) \cdot n/2] \\ &= \frac{2^{3n}}{|\mathcal{T}|} \Pr_{w \sim \mathbb{F}_2^n} [|\{i : w_i \neq 0\}| \leq (1 - \theta) \cdot n/2]. \end{aligned}$$

By Fact 3.1, we have, $\Pr_{w \sim \mathbb{F}_2^n} [|\{i : w_i \neq 0\}| \leq (1 - \theta) \cdot n/2] \leq 2^{-\theta^2 n/4}$, and by Lemma 6.5, we have

$$\frac{|\mathcal{T}|}{2^{3n}} \geq \frac{|\mathcal{V}|^3}{2^{3n}} \cdot (1 - \epsilon/2) \cdot \alpha_E^2 \alpha_F^2 \alpha_G^2 \geq 2^{-3 \cdot (n - \dim(\mathcal{V}))} \cdot \frac{1}{2} \cdot 2^{-6d} \geq 2^{-\kappa n},$$

for some $\kappa = o(1)$. Combining everything, we get

$$\Pr \left[|\mathcal{I}_s| \leq (1 - \theta) \cdot \frac{n}{2} \right] \leq 2^{\kappa n} \cdot 2^{-\theta^2 n/4} = 2^{-(\theta^2 - 4\kappa)n/4}.$$

Now, we choose $\theta = \max \{3\sqrt{\kappa}, n^{-0.1}\} \leq o(1)$, and so the above probability is $o(1)$. Hence,

$$\mathbb{E} |\mathcal{I}_s| \geq \frac{n}{2} \cdot (1 - \theta) \cdot (1 - o(1)) \geq \frac{n}{2} \cdot (1 - o(1)). \quad \square$$

Finally, we complete the proof.

Proof of Lemma 6.2. For any $i \in [n]$, and $S = S(E, F, G)$, we can write

$$\begin{aligned} \Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] &= \Pr_{x, y \sim \mathbb{F}_2^n}[(x, y, x + y) \in \text{Win}_i \mid (x, y, x + y) \in E \times F \times G] \\ &= \Pr_{x, y \sim \mathbb{F}_2^n}[(x, y, x + y) \in \text{Win}_i \mid (x, y) \in S] \\ &= \Pr_{(x, y) \sim S}[(x, y, x + y) \in \text{Win}_i]. \end{aligned}$$

By Lemma 6.5, this can be bounded as

$$\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \leq \mathbb{E}_{(x_0, y_0, w) \sim \mathcal{T}} \mathbb{E}_{(x, y) \sim s_{x_0, y_0, w}} [\mathbb{1}[(x, y, x + y) \in \text{Win}_i]] + \frac{\epsilon}{2}. \quad (3)$$

Observe that for a fixed square $(x_0, y_0, w) \in \mathcal{T}$, $s = s_{x_0, y_0, w}$, by Lemma 6.4, we can bound

$$\mathbb{E}_{(x, y) \sim s_{x_0, y_0, w}} [\mathbb{1}[(x, y, x + y) \in \text{Win}_i]] \leq 1 - (1 - \text{val}(\mathcal{G})) \cdot \mathbb{1}[i \in \mathcal{I}_s].$$

Plugging this into the above, and taking expectation over $i \sim [n]$, we obtain

$$\begin{aligned} \mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] &\leq \mathbb{E}_{i \sim [n]} \mathbb{E}_{(x_0, y_0, w) \sim \mathcal{T}} [1 - (1 - \text{val}(\mathcal{G})) \cdot \mathbb{1}[i \in \mathcal{I}_{s_{x_0, y_0, w}}]] + \frac{\epsilon}{2} \\ &= 1 - \frac{(1 - \text{val}(\mathcal{G}))}{n} \mathbb{E}_{(x_0, y_0, w) \sim \mathcal{T}} |\mathcal{I}_{s_{x_0, y_0, w}}| + \frac{\epsilon}{2}. \\ &\leq 1 - (1 - \text{val}(\mathcal{G})) \cdot \frac{1}{2} \cdot (1 - o(1)) + \frac{\epsilon}{2} \quad (\text{by Lemma 6.6}) \\ &\leq \frac{1}{2} + \frac{\text{val}(\mathcal{G})}{2} + \frac{\epsilon}{2} + o(1) \leq \frac{1}{2} + \frac{\text{val}(\mathcal{G})}{2} + \epsilon. \quad \square \end{aligned}$$

6.2 Hard Coordinate under General Product Events

Now, we work with general product events. Consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n, \mathcal{A}^n \times \mathcal{B}^n \times \mathcal{C}^n, Q^{\otimes n}, V_{\text{pred}}^{\otimes n})$ for some sufficiently large $n \in \mathbb{N}$. Fix any strategies for the 3 players in this game, and for each $i \in [n]$, let $\text{Win}_i \subseteq \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ be the event that this strategy wins the i^{th} coordinate of the game. We prove that:

Lemma 6.7. *For any constant $\epsilon \in (0, 1/4)$, there exists a constant $c = c(\epsilon) > 0$, such that the following holds: For any sets $E, F, G \subseteq \mathbb{F}_2^n$ with $\Pr_{Q^{\otimes n}}[E \times F \times G] \geq 2^{-n^c}$, it holds that*

$$\mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] \leq \frac{1}{2} + \frac{\text{val}(\mathcal{G})}{2} + \epsilon.$$

Proof. Let $d = n^{1/20}$, and let $r = O_\epsilon(d^{16}) \in \mathbb{N}$ and constant $0 < \delta = \delta(\epsilon) < 1/10$ satisfy the statement of Lemma 6.2, with parameters $d, \epsilon/2$.

Let $\eta = \epsilon/3$, and let $C = C_\eta$ be the constant as in Proposition 4.3. Let $d_0 = d^{1/C} = n^{1/(20C)}$ and let $E, F, G \subseteq \mathbb{F}_2^n$ be such that $\alpha = \Pr_{Q^{\otimes n}}[E \times F \times G] \geq 4 \cdot 2^{-d_0}$. Note that the lemma statement will hold with $c(\epsilon) = 1/(40C)$.

Let $S = S(E, F, G)$ be as in Definition 2.1; this satisfies $|S| = 2^{2n}\alpha$. Now, by Proposition 4.3, (with parameters r, δ, η , and the space \mathbb{F}_2^n), we find an integer $T \in \mathbb{N}$, and for each $t \in [T]$, a linear subspace $\mathcal{V}_t \subseteq \mathbb{F}_2^n$, points $e_t, f_t \in \mathbb{F}_2^n$, and subsets $E_t \subseteq e_t + \mathcal{V}_t$, $F_t \subseteq f_t + \mathcal{V}_t$, $G_t \subseteq e_t + f_t + \mathcal{V}_t$, such that:

1. For all $t \in [T]$, we have $\dim(\mathcal{V}_t) \geq n - r\delta^{-3} \log_2(4/\alpha)^C = n - O(d^{17}) \geq n - o(n)$.
2. $S(E_1, F_1, G_1), \dots, S(E_T, F_T, G_T)$ are disjoint subsets of S such that

$$|S \setminus \cup_{t=1}^T S(E_t, F_t, G_t)| \leq \eta |S| = \eta \alpha \cdot 2^{2n}.$$

3. For all $t \in [T]$, we have $|E_t|, |F_t|, |G_t| \geq 2^{-\log_2(4/\alpha)^C} |\mathcal{V}_t| \geq 2^{-d} |\mathcal{V}_t|$, and E_t, F_t, G_t are (r, δ) -algebraically spread within $e_t + \mathcal{V}_t, f_t + \mathcal{V}_t, e + f_t + \mathcal{V}_t$ respectively.

Now, for any $i \in [n]$, we may write

$$\begin{aligned} \Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] &= \alpha^{-1} \cdot \Pr_{Q^{\otimes n}}[\text{Win}_i \cap E \times F \times G] \\ &= \alpha^{-1} \cdot \Pr_{x, y \sim \mathbb{F}_2^n}[(x, y, x + y) \in \text{Win}_i \cap (x, y) \in S(E, F, G)] \\ &\leq \eta + \sum_{t=1}^T \alpha^{-1} \cdot \Pr_{x, y \sim \mathbb{F}_2^n}[(x, y, x + y) \in \text{Win}_i \cap (x, y) \in S(E_t, F_t, G_t)] \\ &= \eta + \sum_{t=1}^T \alpha^{-1} \cdot \Pr_{Q^{\otimes n}}[\text{Win}_i \cap E_t \times F_t \times G_t] \\ &= \eta + \sum_{t=1}^T \alpha^{-1} \cdot \Pr_{Q^{\otimes n}}[\text{Win}_i \mid E_t \times F_t \times G_t] \cdot \frac{|S(E_t, F_t, G_t)|}{2^{2n}}. \end{aligned}$$

Also, by Lemma 6.2, we have for each $t \in [T]$, that

$$\mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E_t \times F_t \times G_t] \right] \leq \frac{1 + \text{val}(\mathcal{G}) + \epsilon}{2}.$$

Here, we note that when the inputs are drawn conditioned on $E_t \times F_t \times G_t$, the three players can shift them by $e_t, f_t, e_t + f_t \in \mathbb{F}_2^n$ respectively; now the inputs to each player lie in spread subsets of the linear subspace \mathcal{V}_t , and the lemma is applicable. Finally, combining the above two equations, and recalling $\eta = \epsilon/3$, we get

$$\begin{aligned} \mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] &\leq \eta + \sum_{t=1}^T \alpha^{-1} \cdot \mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E_t \times F_t \times G_t] \right] \cdot \frac{|S(E_t, F_t, G_t)|}{2^{2n}} \\ &\leq \eta + \sum_{t=1}^T \alpha^{-1} \cdot \left(\frac{1 + \text{val}(\mathcal{G}) + \epsilon}{2} \right) \cdot \frac{|S(E_t, F_t, G_t)|}{2^{2n}} \\ &= \eta + \alpha^{-1} \cdot \left(\frac{1 + \text{val}(\mathcal{G}) + \epsilon}{2} \right) \cdot \frac{\sum_{t=1}^T |S(E_t, F_t, G_t)|}{2^{2n}} \\ &\leq \eta + \alpha^{-1} \cdot \left(\frac{1 + \text{val}(\mathcal{G}) + \epsilon}{2} \right) \cdot \frac{|S|}{2^{2n}} = \eta + \frac{1 + \text{val}(\mathcal{G}) + \epsilon}{2} \\ &\leq \frac{1}{2} + \frac{\text{val}(\mathcal{G})}{2} + \epsilon. \end{aligned} \quad \square$$

With the above, the proof of our main theorem is direct:

Proof of Theorem 6.1. The theorem follows by combining Lemma 3.13 with Lemma 6.7, applied with the parameter choice $\epsilon = 0.1$, recalling that $\text{val}(\mathcal{G}) \leq 3/4$. \square

Proof of Theorem 1.1. The theorem follows by combining Theorem 6.1 and Lemma 3.12. \square

7 A Concentration Bound

We also prove a concentration bound for the GHZ game:

Theorem 7.1. *Let $\mathcal{G} = (\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{A} \times \mathcal{B} \times \mathcal{C}, Q, V_{\text{pred}})$ be any 3-player game with $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathbb{F}_2$, and with Q the uniform distribution over*

$$\text{supp}(Q) = \{(x, y, z) \in \mathbb{F}_2^3 : x + y + z = 0\},$$

and such that $\text{val}(\mathcal{G}) < 1$. Then, for every constant $\epsilon \in (0, 1/4)$, there exists a constant $c = c(\epsilon) > 0$, such that the following holds for all sufficiently large n .³

Consider the game $\mathcal{G}^{\otimes n}$, and consider any strategies for the 3 players in this game. Then, if Z denotes the number of coordinates that the players win, it holds that

$$\Pr[Z \geq (\text{val}(\mathcal{G}) + \epsilon) \cdot n] \leq \exp(-n^c).$$

The remainder of this section is devoted to the proof of the above theorem. Towards this, we fix a game \mathcal{G} as in the statement of the theorem. Consider the game $\mathcal{G}^{\otimes n} = (\mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n, \mathcal{A}^n \times \mathcal{B}^n \times \mathcal{C}^n, Q^{\otimes n}, V_{\text{pred}}^{\otimes n})$ for some sufficiently large $n \in \mathbb{N}$. Also, fix any strategies for the 3 players in this game, and for each $i \in [n]$, let $\text{Win}_i \subseteq \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ be the event that this strategy wins the i^{th} coordinate of the game. For each $S \subseteq [n]$, we use Win_S to denote the event $\bigwedge_{i \in S} \text{Win}_i$.

7.1 An Improved Bound for Product Events

The first step in our proof is to prove an improved version of Lemma 6.7, where we don't lose a "factor of 2." We prove the following:

Lemma 7.2. *For any constant $\epsilon \in (0, 1/4)$, there exists $c = c(\epsilon) > 0$, such that the following holds: For any sets $E, F, G \subseteq \mathbb{F}_2^n$ with $\Pr_{Q^{\otimes n}}[E \times F \times G] \geq 2^{-n^c}$, it holds that*

$$\mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] \leq \text{val}(\mathcal{G}) + \epsilon.$$

The proof of the above lemma exactly follows the proof in Section 6, so we only mention how to improve it. We note that the only place where we lost a factor of 2 in the bound was in the proof of Lemma 6.2; the proof of Lemma 6.7, using uniformization, goes through as is if an improved version of Lemma 6.2 is proved. Now, in Lemma 6.2, we proceed in exactly the same manner upto Equation (3), which says (after taking expectation over $i \sim [n]$)

$$\mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] \leq \mathbb{E}_{i \sim [n]} \mathbb{E}_{(x_0, y_0, w) \sim \mathcal{T}} \mathbb{E}_{(x, y) \sim s_{x_0, y_0, w}} [\mathbb{1}[(x, y, x + y) \in \text{Win}_i]] + \frac{\epsilon}{2}.$$

The factor 2 loss now occurred because when $(x_0, y_0, w) \sim \mathcal{T}$, the event $w_i = 1$ occurs with probability roughly $1/2$ (see Lemma 6.6), and we can only apply Lemma 6.4 when this holds. To improve upon this bound, we will replace the distribution with one where we are already conditioning on the event $w_i = 1$, and only then apply Lemma 6.4; this requires proving a stronger version of Lemma 6.6, which says that the distribution of (x, y) in the above statement remains unaffected even when the square is sampled conditioned on $w_i = 1$. Formally, we prove:

Lemma 7.3. *Recall the distribution μ on S obtained as follows: choose $(x_0, y_0, w) \sim \mathcal{T}$, and let $(x, y) \sim s_{x_0, y_0, w}$. For any $i \in [n]$, define the distribution ν_i as follows: choose $(x_0, y_0, w) \sim \mathcal{T}$ conditioned on $w_i = 1$, and let $(x, y) \sim s_{x_0, y_0, w}$. Then, we have*

$$\mathbb{E}_{i \sim [n]} \|\mu - \nu_i\|_1 \leq o(1).$$

In the proof, we will need the following technical lemma, whose proof is deferred to Appendix A.

Lemma 7.4. *Let $\mathcal{T} \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n \times \mathbb{F}_2^n$ be any set of size at least $2^{3n-o(n)}$. Let $(X, Y, W) \sim \mathcal{T}$ be chosen uniformly at random. Then,*

$$\mathbb{E}_{i \sim [n]} \|P_{X,Y|W_i=1} - P_{X,Y}\|_1 \leq o(1).$$

Here, $P_{X,Y}$ denotes the marginal distribution of (X, Y) , and $P_{X,Y|W_i=1}$ denotes the marginal distribution of (X, Y) conditioned on the event $W_i = 1$.

Proof of Lemma 7.3. First, observe the distribution μ is the same as the following distribution: choose $(x_0, y_0, w) \sim \mathcal{T}$, and output (x_0, y_0) . This can be seen from Equation (2): essentially, this follows by Remark 2.3, as every square with $w \neq 0$ occurs 4 times in \mathcal{T} , and each point in such a square is chosen with probability $1/4$ after the square is chosen; also, each square with $w = 0$ occurs only once, however the point within this square is chosen with probability 1 after such a square is chosen. By a similar reasoning, for every $i \in [n]$, the distribution ν_i is the same as the following: choose $(x_0, y_0, w) \sim \mathcal{T} \mid w_i = 1$, and output (x_0, y_0) .

Now, observe that by Lemma 6.5, we have

$$|\mathcal{T}| \geq |\mathcal{V}|^3 \cdot (1 - \epsilon/2) \cdot \alpha_E^2 \alpha_F^2 \alpha_G^2 \geq 2^{3n - (3(n - \dim(\mathcal{V})) + 1 + 6d)} \geq 2^{3n - o(n)}.$$

The result now follows by Lemma 7.4. □

Finally, we complete the proof of the improved bound:

Proof of Lemma 7.2. As observed before, it suffices to improve Lemma 6.2, and we have

$$\mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}}[\text{Win}_i \mid E \times F \times G] \right] \leq \mathbb{E}_{i \sim [n]} \mathbb{E}_{(x_0, y_0, w) \sim \mathcal{T}} \mathbb{E}_{(x, y) \sim s_{x_0, y_0, w}} [\mathbb{1}[(x, y, x + y) \in \text{Win}_i]] + \frac{\epsilon}{2}.$$

By Lemma 7.3 and Lemma 6.4, the above is at most

$$\begin{aligned} \mathbb{E}_{i \sim [n]} \mathbb{E}_{(x_0, y_0, w) \sim \mathcal{T} \mid w_i = 1} \mathbb{E}_{(x, y) \sim s_{x_0, y_0, w}} [\mathbb{1}[(x, y, x + y) \in \text{Win}_i]] + \frac{\epsilon}{2} + o(1) &\leq \text{val}(\mathcal{G}) + \frac{\epsilon}{2} + o(1) \\ &\leq \text{val}(\mathcal{G}) + \epsilon. \end{aligned} \quad \square$$

7.2 Random Sets are Hard under Parallel Repetition

We prove a strengthening of Theorem 6.1, by using an argument similar to the proof of Lemma 3.13.

Lemma 7.5. *For every constant $\epsilon \in (0, 1/4)$, there exists $c = c(\epsilon) > 0$, such that for every integer $1 \leq t \leq n^c$, it holds that*

$$\mathbb{E}_{|S|=t} \left[\Pr_{Q^{\otimes n}} [\text{Win}_S] \right] \leq 2 \cdot (\text{val}(\mathcal{G}) + \epsilon)^t,$$

where $\mathbb{E}_{|S|=t}$ denotes that the expectation is over a uniformly chosen subset $S \subseteq [n]$ of size t .

Proof. Let $\epsilon \in (0, 1/4)$ be any constant, and let $0 < c \leq 1$ be as in Lemma 7.2 with respect to the parameter $\epsilon/2$. We show that the lemma holds with constant $c/2$.

More formally, by induction on t , we show that for every $1 \leq t \leq \lfloor n^{c/2} \rfloor$,

$$\mathbb{E}_{|S|=t} \left[\Pr_{Q^{\otimes n}} [\text{Win}_S] \right] \leq (\text{val}(\mathcal{G}) + \epsilon)^t + t \cdot 2^{-n^{c/2}} \leq 2 \cdot (\text{val}(\mathcal{G}) + \epsilon)^t.$$

Note that the second inequality holds as we may assume that $\text{val}(\mathcal{G}) \geq 1/4$ (or else $\text{val}(\mathcal{G}) = 0$ and there is nothing to prove) in which case $t \cdot 2^{-n^{c/2}} \leq (1/4)^{n^{c/2}} \leq (1/4)^t \leq (\text{val}(\mathcal{G}) + \epsilon)^t$.

The base case $t = 1$ holds trivially as $\Pr_{Q^{\otimes n}} [\text{Win}_i] \leq \text{val}(\mathcal{G})$ for each $i \in [n]$. For the inductive step, consider any $1 \leq t < \lfloor n^{c/2} \rfloor$, and suppose that $\mathbb{E}_{|T|=t} [\Pr_{Q^{\otimes n}} [\text{Win}_T]] \leq (\text{val}(\mathcal{G}) + \epsilon)^t + t \cdot 2^{-n^{c/2}}$. Then, we have

$$\begin{aligned} \mathbb{E}_{|S|=t+1} \left[\Pr_{Q^{\otimes n}} [\text{Win}_S] \right] &= \mathbb{E}_{|T|=t} \mathbb{E}_{i \sim [n] \setminus T} \left[\Pr_{Q^{\otimes n}} [\text{Win}_T \wedge \text{Win}_i] \right] \\ &= \mathbb{E}_{|T|=t} \left[\Pr_{Q^{\otimes n}} [\text{Win}_T] \cdot \mathbb{E}_{i \sim [n] \setminus T} \left[\Pr_{Q^{\otimes n}} [\text{Win}_i \mid \text{Win}_T] \right] \right]. \end{aligned}$$

With the above expression in mind, we fix any set $T \subseteq [n]$ of size t , and calculate an upper bound for $\mathbb{E}_{i \sim [n] \setminus T} [\Pr [\text{Win}_i \mid \text{Win}_T]]$. Let $R \in (\mathbb{F}_2 \times \mathbb{F}_2 \times \mathbb{F}_2 \times \mathcal{A} \times \mathcal{B} \times \mathcal{C})^T$ be the random variable consisting of the questions and answers of the players in coordinates given by set T ; note that this random variable takes at most $(8|\mathcal{A}||\mathcal{B}||\mathcal{C}|)^t \leq 2^{O(t)}$ possible values. Also, observe that the event Win_T is a deterministic function of R . Let \mathcal{R} be the set of all values r of the random variable R that satisfy the event Win_T , and let $\mathcal{R}' \subseteq \mathcal{R}$ be the set of all $r \in \mathcal{R}$ such that $\Pr_{Q^{\otimes n}} [R = r] \geq 2^{-n^c}$. Note that by Lemma 7.2, we have that for each $r \in \mathcal{R}'$,

$$\mathbb{E}_{i \sim [n]} \Pr_{Q^{\otimes n}} [\text{Win}_i \mid R = r] \leq \text{val}(\mathcal{G}) + \frac{\epsilon}{2},$$

since the event $R = r$ is a product event with respect to the three players.

Hence, we get

$$\begin{aligned}
\mathbb{E}_{i \sim [n] \setminus T} \left[\Pr_{Q^{\otimes n}} [\text{Win}_i \mid \text{Win}_T] \right] &\leq \mathbb{E}_{i \sim [n]} \left[\Pr_{Q^{\otimes n}} [\text{Win}_i \mid \text{Win}_T] \right] + \frac{2t}{n} \quad (\text{Lemma 3.2}) \\
&= \sum_{r \in \mathcal{R}} \frac{\mathbb{E}_{i \sim [n]} [\Pr_{Q^{\otimes n}} [\text{Win}_i \mid R = r]] \cdot \Pr_{Q^{\otimes n}} [R = r]}{\Pr_{Q^{\otimes n}} [\text{Win}_T]} + o(1) \\
&\leq \sum_{r \in \mathcal{R}'} \left(\text{val}(\mathcal{G}) + \frac{\epsilon}{2} \right) \cdot \frac{\Pr_{Q^{\otimes n}} [R = r]}{\Pr_{Q^{\otimes n}} [\text{Win}_T]} + \sum_{r \in \mathcal{R} \setminus \mathcal{R}'} \frac{1 \cdot 2^{-n^c}}{\Pr_{Q^{\otimes n}} [\text{Win}_T]} + o(1) \\
&\leq \left(\text{val}(\mathcal{G}) + \frac{\epsilon}{2} \right) + \frac{(8|\mathcal{A}||\mathcal{B}||\mathcal{C}|)^t \cdot 2^{-n^c}}{\Pr_{Q^{\otimes n}} [\text{Win}_T]} + o(1) \\
&\leq \text{val}(\mathcal{G}) + \epsilon + \frac{2^{O(n^{c/2}) - n^c}}{\Pr_{Q^{\otimes n}} [\text{Win}_T]} \leq \text{val}(\mathcal{G}) + \epsilon + \frac{2^{-n^c/2}}{\Pr_{Q^{\otimes n}} [\text{Win}_T]}.
\end{aligned}$$

Combining this with the above, and using the inductive hypothesis, we get

$$\begin{aligned}
\mathbb{E}_{|S|=t+1} \left[\Pr_{Q^{\otimes n}} [\text{Win}_S] \right] &= \mathbb{E}_{|T|=t} \left[\Pr_{Q^{\otimes n}} [\text{Win}_T] \cdot \left(\text{val}(\mathcal{G}) + \epsilon + \frac{2^{-n^c/2}}{\Pr_{Q^{\otimes n}} [\text{Win}_T]} \right) \right] \\
&\leq \mathbb{E}_{|T|=t} \left[\Pr_{Q^{\otimes n}} [\text{Win}_T] \right] \cdot (\text{val}(\mathcal{G}) + \epsilon) + 2^{-n^c/2} \\
&\leq ((\text{val}(\mathcal{G}) + \epsilon)^t + t \cdot 2^{-n^c/2}) \cdot (\text{val}(\mathcal{G}) + \epsilon) + 2^{-n^c/2} \\
&\leq (\text{val}(\mathcal{G}) + \epsilon)^{t+1} + (t+1) \cdot 2^{-n^c/2}. \quad \square
\end{aligned}$$

7.3 Proof of the Concentration Bound

Finally, we complete the proof of our concentration bound, using standard arguments [Rao11]:

Proof of Theorem 7.1. Consider any constant $\epsilon \in (0, 1/4)$, let $0 < c < 1$ be as in Lemma 7.5 with the parameter $\epsilon/2$, and let $t = \lfloor n^c \rfloor$.

Let Z be the random variable denoting the number of coordinates won by the players. Whenever $Z \geq (\text{val}(\mathcal{G}) + \epsilon) \cdot n$, we pick a uniformly random subset S of size t from the coordinates that the players won. Note that for any fixed subset $T \subseteq [n]$ of size t , the probability that S equals T is at most $\binom{(\text{val}(\mathcal{G}) + \epsilon) \cdot n}{t}^{-1}$. Hence, we have

$$\begin{aligned}
\Pr [Z \geq (\text{val}(\mathcal{G}) + \epsilon) \cdot n] &\leq \sum_{T \subseteq [n], |T|=t} \Pr[S = T] \cdot \Pr[\text{Win}_T] \\
&\leq \sum_{T \subseteq [n], |T|=t} \left(\binom{(\text{val}(\mathcal{G}) + \epsilon) \cdot n}{t} \right)^{-1} \cdot \Pr[\text{Win}_T].
\end{aligned}$$

Now, using Lemma 7.5, we get

$$\begin{aligned}
\Pr[Z \geq (\text{val}(\mathcal{G}) + \epsilon) \cdot n] &\leq \binom{n}{t} \cdot \left(\frac{(\text{val}(\mathcal{G}) + \epsilon) \cdot n}{t} \right)^{-1} \cdot 2 \cdot (\text{val}(\mathcal{G}) + \epsilon/2)^t \\
&\leq \left(\frac{n}{(\text{val}(\mathcal{G}) + \epsilon) \cdot n - t} \right)^t \cdot 2 \cdot (\text{val}(\mathcal{G}) + \epsilon/2)^t \\
&\leq 2 \cdot \left(\frac{\text{val}(\mathcal{G}) + \epsilon/2}{\text{val}(\mathcal{G}) + \epsilon - o(1)} \right)^t \\
&= 2 \cdot \left(1 - \frac{\epsilon/2 - o(1)}{\text{val}(\mathcal{G}) + \epsilon - o(1)} \right)^t \leq 2 \cdot (1 - \epsilon/4)^t \\
&\leq 2^{1-\epsilon t/4} \leq 2^{-n^{c/2}}. \quad \square
\end{aligned}$$

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. System Sci.*, 54(2, part 2):317–331, 1997. (also in FOCS 1993). [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. (also in STOC 2010). [3](#)
- [BBK⁺24] Amey Bhangale, Mark Braverman, Subhash Khot, Yang P. Liu, and Dor Minzer. Parallel repetition of k -player projection games. In *APPROX-RANDOM*, pages Art. No. 54, 16, 2024. [3](#), [5](#)
- [BBK⁺25] Amey Bhangale, Mark Braverman, Subhash Khot, Yang Liu, and Dor Minzer. Parallel repetition for 3-player XOR games. In *STOC*, pages 104–110, 2025. [3](#), [4](#)
- [BBK⁺26] Amey Bhangale, Mark Braverman, Subhash Khot, Yang P. Liu, Dor Minzer, and Kunal Mittal. An analytical approach to parallel repetition via CSP inverse theorems. In *STOC*, 2026. (To appear). [3](#), [11](#)

- [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 *STOC*, pages 335–340, 2015. [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. (also in FOCS 1995). [3](#)
- [BKM23] Mark Braverman, Subhash Khot, and Dor Minzer. Parallel repetition for the GHZ game: exponential decay. In *FOCS*, pages 1337–1341, 2023. [3](#), [4](#), [5](#)
- [BM21] Mark Braverman and Dor Minzer. Optimal tiling of the Euclidean space using permutation-symmetric bodies. In *CCC*, pages Art. No. 5, 48, 2021. [3](#)
- [BOGKW88] 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. [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 *FOCS*, pages 746–755, 2013. [3](#)
- [CHTW04] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *CCC*, pages 236–249, 2004. [3](#)
- [CT06] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience, second edition, 2006. [38](#)
- [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. (also in STOC 2003). [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. [3](#), [4](#)
- [DRS05] Irit Dinur, Oded Regev, and Clifford Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005. [3](#)
- [DS14] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *STOC*, pages 624–633, 2014. [3](#)

- [Fei91] Uriel Feige. On the success probability of the two provers in one-round proof systems. In *Structure in Complexity Theory Conference*, pages 116–123, 1991. [3](#)
- [Fei98] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45(4):634–652, 1998. (also in STOC 1996). [3](#)
- [FGL⁺96] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996. [3](#)
- [FHHK24] Yuval Filmus, Hamed Hatami, Kaave Hosseini, and Esty Kelman. Sparse graph counting and Kelley-Meka bounds for binary systems. In *FOCS*, pages 1559–1578, 2024. [8](#), [20](#)
- [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 *CCC*, pages 179–192, 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. *Theoret. 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](#), [11](#)
- [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. [3](#), [4](#), [7](#), [8](#)
- [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. [3](#), [11](#)
- [GHS02] Venkatesan Guruswami, Johan Håstad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM J. Comput.*, 31(6):1663–1686, 2002. (also in FOCS 2000). [3](#)
- [GHZ89] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. Going beyond bell’s theorem. In *Bell’s Theorem, Quantum Theory and Conceptions of the Universe*, pages 69–72. Springer Netherlands, 1989. [4](#)
- [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. [3](#)

- [Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. (also in STOC 1997). 3
- [HHR16] Jan Håzla, 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). 3
- [HR20] Justin Holmgren and Ran Raz. A parallel repetition theorem for the GHZ game. *CoRR*, abs/2008.05059, 2020. Available at <https://arxiv.org/abs/2008.05059>. 3, 4
- [JLL⁺25] Michael Jaber, Yang P. Liu, Shachar Lovett, Anthony Ostuni, and Mehtaab Sawhney. Quasipolynomial bounds for the corners theorem. In *FOCS*, 2025. 4, 8, 12, 14
- [Kho02a] Subhash Khot. Hardness results for approximate hypergraph coloring. In *STOC*, pages 351–359, 2002. 3
- [Kho02b] Subhash Khot. Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *FOCS*, pages 23–32, 2002. 3
- [KLM24] Zander Kelley, Shachar Lovett, and Raghu Meka. Explicit separations between randomized and deterministic Number-on-Forehead communication. In *STOC*, pages 1299–1310, 2024. 4, 19
- [KM23] Zander Kelley and Raghu Meka. Strong bounds for 3-progressions. In *FOCS*, pages 933–973, 2023. 4, 8, 12
- [KORW08] Guy Kindler, Ryan O’Donnell, Anup Rao, and Avi Wigderson. Spherical cubes and rounding in high dimensions. In *FOCS*, pages 189–198, 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, 7
- [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. 3
- [MU05] Michael Mitzenmacher and Eli Upfal. *Probability and computing*. Cambridge University Press, Cambridge, 2005. Randomized algorithms and probabilistic analysis. 9
- [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 *STOC*, pages 363–372, 1997. [3](#)
- [Rao11] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011. (also in STOC 2008). [3](#), [33](#)
- [Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. (also in STOC 1995). [3](#), [4](#), [6](#), [11](#)
- [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](#)
- [ŠB20] Ivan Šupić and Joseph Bowles. Self-testing of quantum systems: a review. *Quantum*, 4:337, 2020. [4](#)
- [Ver96] Oleg Verbitsky. Towards the parallel repetition conjecture. *Theoret. Comput. Sci.*, 157(2):277–282, 1996. [3](#)

A Marginals of a Conditional Distribution

In this section, we prove Lemma [7.4](#). Our proof will use some basic information theory, and the reader is referred to [\[CT06\]](#) for an excellent introduction to information theory. Recall that for a random variable X over a finite set Ω , its entropy is defined as

$$H(X) = - \sum_{x \in \Omega} \Pr[X = x] \cdot \log_2(\Pr[X = x]).$$

We have the following simple fact:

Lemma A.1. *Let $X \in \{0, 1\}$ be a binary valued random variable. Then,*

$$\left| \Pr[X = 1] - \frac{1}{2} \right|^2 \leq 1 - H(X).$$

Proof. Let $\delta = \left| \Pr[X = 1] - \frac{1}{2} \right|$. The lemma statement is equivalent to the inequality

$$\delta^2 \leq 1 - H(X) = 1 + \left(\frac{1}{2} + \delta \right) \log_2 \left(\frac{1}{2} + \delta \right) + \left(\frac{1}{2} - \delta \right) \log_2 \left(\frac{1}{2} - \delta \right),$$

which is true for all $\delta \in [0, 1/2]$. □

Next, we prove the main lemma of this section:

Proof of Lemma 7.4. Suppose that $|\mathcal{T}| \geq 2^{(3-\delta)n}$ for some $\delta = o(1)$. Then, we have

$$(3 - \delta)n \leq H(X, Y, W) = H(X, Y) + H(W | X, Y) \leq 2n + \sum_{i=1}^n H(W_i | X, Y),$$

which implies

$$\mathbb{E}_{i \sim [n]} [H(W_i | X, Y)] \geq 1 - \delta.$$

Define the set

$$\mathcal{G} = \left\{ i \in [n] : H(W_i | X, Y) \geq 1 - \sqrt{\delta} \right\}.$$

Then, by Markov's inequality, we have $\Pr_{i \sim [n]} [i \notin \mathcal{G}] \leq \sqrt{\delta}$.

Now, consider any $i \in \mathcal{G}$. By Cauchy-Schwarz and Lemma A.1, we have

$$\begin{aligned} \left(\mathbb{E}_{(x,y) \sim P_{X,Y}} \left| \Pr[W_i = 1 | X = x, Y = y] - \frac{1}{2} \right| \right)^2 &\leq \mathbb{E}_{(x,y) \sim P_{X,Y}} \left| \Pr[W_i = 1 | x, y] - \frac{1}{2} \right|^2 \\ &\leq \mathbb{E}_{(x,y) \sim P_{X,Y}} [1 - H(W_i | x, y)] \\ &= 1 - H(W_i | X, Y) \leq \sqrt{\delta}. \end{aligned}$$

Hence, by the triangle inequality, we have

$$\left| \Pr[W_i = 1] - \frac{1}{2} \right| \leq \mathbb{E}_{(x,y) \sim P_{X,Y}} \left| \Pr[W_i = 1 | x, y] - \frac{1}{2} \right| \leq \delta^{1/4}.$$

By the above, we get

$$\begin{aligned} \|P_{X,Y|W_i=1} - P_{X,Y}\|_1 &= \sum_{x,y} |\Pr[x, y | W_i = 1] - \Pr[x, y]| \\ &= \mathbb{E}_{(x,y) \sim P_{X,Y}} \left| \frac{\Pr[W_i = 1 | x, y]}{\Pr[W_i = 1]} - 1 \right| \\ &\leq \mathbb{E}_{(x,y) \sim P_{X,Y}} \left| \frac{\Pr[W_i = 1 | x, y]}{1/2} - 1 \right| \\ &\quad + \mathbb{E}_{(x,y) \sim P_{X,Y}} [\Pr[W_i = 1 | x, y]] \cdot \left| \frac{1}{\Pr[W_i = 1]} - \frac{1}{1/2} \right| \\ &\leq 2 \cdot \delta^{1/4} + \Pr[W_i = 1] \cdot \frac{\delta^{1/4}}{\Pr[W_i = 1] \cdot 1/2} \\ &\leq 4\delta^{1/4}. \end{aligned}$$

Hence,

$$\mathbb{E}_{i \sim [n]} \|P_{X,Y|W_i=1} - P_{X,Y}\|_1 \leq 4\delta^{1/4} + 2 \cdot \Pr_{i \sim [n]} [i \notin \mathcal{G}] \leq 4\delta^{1/4} + 2\delta^{1/2} \leq o(1). \quad \square$$