

# Parallel Repetition for 3-Player XOR Games

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

## Abstract

In a 3-XOR game  $\mathcal{G}$ , the verifier samples a challenge  $(x, y, z) \sim \mu$  where  $\mu$  is a probability distribution over  $\Sigma \times \Gamma \times \Phi$ , and a map  $t: \Sigma \times \Gamma \times \Phi \rightarrow \mathcal{A}$  for a finite Abelian group  $\mathcal{A}$  defining a constraint. The verifier sends the questions  $x, y$  and  $z$  to the players Alice, Bob and Charlie respectively, receives answers  $f(x), g(y)$  and  $h(z)$  that are elements in  $\mathcal{A}$  and accepts if  $f(x) + g(y) + h(z) = t(x, y, z)$ . The value,  $\text{val}(\mathcal{G})$ , of the game is defined to be the maximum probability the verifier accepts over all players' strategies.

We show that if  $\mathcal{G}$  is a 3-XOR game with value strictly less than 1, whose underlying distribution over questions  $\mu$  does not admit Abelian embeddings into  $(\mathbb{Z}, +)$ , then the value of the  $n$ -fold repetition of  $\mathcal{G}$  is exponentially decaying. That is, there exists  $c = c(\mathcal{G}) > 0$  such that  $\text{val}(\mathcal{G}^{\otimes n}) \leq 2^{-cn}$ . This extends a previous result of [Braverman-Khot-Minzer, FOCS 2023] showing exponential decay for the GHZ game. Our proof combines tools from additive combinatorics and tools from discrete Fourier analysis.

---

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

†Department of Computer Science at Princeton University. Supported by the National Science Foundation under the Alan T. Waterman Award, Grant No. 1933331.

‡Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. Supported by the NSF Award CCF-1422159, NSF CCF award 2130816, and the Simons Investigator Award.

§School of Mathematics, Institute for Advanced Study, Princeton, NJ. This material is based upon work supported by the National Science Foundation under Grant No. DMS-1926686

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

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	3-XOR Games, Embeddings and Connectedness . . . . .	2
1.2	Our Main Result . . . . .	3
1.3	$\mathbb{Z}$ -Embeddings, Pairwise-Connectivity and Beyond 3-XOR Games . . . . .	3
<b>2</b>	<b>Overview of Proof of Theorem 1</b>	<b>3</b>
2.1	High-Level Structure of Argument . . . . .	3
2.2	Proof of Main Analytic Results . . . . .	7
<b>3</b>	<b>Preliminaries</b>	<b>9</b>
3.1	General Notation . . . . .	9
3.2	Efron-Stein Decomposition . . . . .	9
3.3	Random Restrictions . . . . .	10
<b>4</b>	<b>No Complex Solutions: Reductions</b>	<b>10</b>
4.1	Path Tricks . . . . .	11
4.2	Merging Symbols . . . . .	13
4.3	The Relaxed Base Case . . . . .	15
4.4	Reducing to High-Degree Functions . . . . .	17
<b>5</b>	<b>Main Induction: No Complex Embeddings</b>	<b>19</b>
5.1	Finding a Singular Value Gap . . . . .	20
5.2	Cauchy-Schwarz and Simple Observations . . . . .	20
5.3	Handling the Case $ R  <  S  T $ . . . . .	22
5.4	Handling the Case $ R  \geq  S  T $ . . . . .	24
<b>6</b>	<b>Complex Embeddings: Reductions</b>	<b>26</b>
6.1	Master Group and Embedding . . . . .	26
6.2	Modest Noise Operator and Modest Influences . . . . .	27
6.3	Saturating the Master Embedding via Path Tricks . . . . .	27
6.4	Merging Symbols . . . . .	30
6.5	Establishing the Master Relaxed Base Case . . . . .	30
6.6	Applying the Modest Noise Operator . . . . .	32
<b>7</b>	<b>Main Induction: Complex Embeddings</b>	<b>34</b>
7.1	Finding a Singular Value Gap . . . . .	35
7.2	Cauchy-Schwarz and Simple Observations . . . . .	35
7.3	Handling the Case $ R  \geq  S  T $ . . . . .	36
<b>8</b>	<b>Additive Combinatorics</b>	<b>37</b>
<b>A</b>	<b>Proof of Lemma 4.12</b>	<b>42</b>

# 1 Introduction

In a  $k$ -player game  $\mathcal{G}$ , a verifier chooses a question  $x := (x^1, x^2, \dots, x^k) \sim \mu$ , for some distribution  $\mu$  over  $\Sigma_1 \times \Sigma_2 \times \dots \times \Sigma_k$ , for finite alphabets  $\Sigma_1, \dots, \Sigma_k$ . The verifier sends  $x^i$  to player  $i$ , who responds with an answer  $f^i(x^i)$ . The verifier accepts based on the evaluation of some known predicate  $V(x, \{f^1(x^1), \dots, f^k(x^k)\})$ . We define the *value*  $\text{val}(\mathcal{G})$  of the game  $\mathcal{G}$  as the maximum probability of winning such a game for the players.

A natural question that arises is how the value of a game decays under parallel repetitions. The  $n$ -fold parallel repetition of  $\mathcal{G}$ , which we denote as  $\mathcal{G}^{\otimes n}$ , is the game where the verifier samples  $n$  questions  $(x_1, \dots, x_n) \sim \mu^{\otimes n}$ , and sends the  $i$ -th coordinate of all of  $(x_1, \dots, x_n)$  to player  $i$  simultaneously. The players win the repeated game if they win each one of the  $n$  instances.

Parallel repetition of 2-player games is well-understood by now. Originally suggested by [FRS94] as a way to amplify the advantage in interactive protocols, it is now known that the value of 2-player games decays exponentially if the value of the game is strictly smaller than 1. This result was first proved by Raz [Raz98] using information theoretic techniques. Subsequent works simplified the proofs and improved it quantitatively [Hol09, Rao11, DS14, BG15]. We remark that most of these works, with the exception of [DS14], follow the information theoretic proof of Raz. The work [DS14] suggests an analytical approach to the problem in the case the game  $\mathcal{G}$  is a projection game.<sup>1</sup> While it is natural to suspect that the (naïve) bound  $\text{val}(\mathcal{G}^{\otimes n}) \leq \text{val}(\mathcal{G})^n$  should hold, it turns out to be false [Fei91, FV02]. Even this failure is well understood, and it is known to be related to the geometry of tilings of high dimensional Euclidean space (see [FKO07, KORW08, AK09, BM21]), and more specifically to the fact that there are bodies of volume 1 and surface area  $\Theta(\sqrt{n})$  that tile the Euclidean space  $\mathbb{R}^n$ . Parallel repetition of 2-player games has many applications in the theory of interactive proofs [BGKW88], communication complexity [PRW97, BBCR13, BRWY13], quantum information [CHTW04, BBLV13], and hardness of approximation [FGL<sup>+</sup>96, ABSS97, ALM<sup>+</sup>98, AS98, BGS98, Fei98, Has01, Kho02a, Kho02b, GHS02, DGKR05, DRS05].

The situation for  $k$ -player games, where  $k \geq 3$ , is much more poorly understood. Even for  $k = 3$ , the best bound known general bound for  $\text{val}(\mathcal{G}^{\otimes n})$  is inverse Ackermann, which follows from quantitative versions of the Density Hales-Jewett theorem [Ver96, Pol12]. Since proving effective bounds on the value of parallel repetition of  $k$ -player games seems like a very challenging task, researchers tried to focus their attention on more limited classes of games. In [DHVY17], the authors consider the class of *connected games*, and prove that an exponential decay holds for them. Their proof uses information-theoretic techniques similar to those used in the context of 2-player version. Here, a game is connected if the following graph on  $\text{supp}(\mu)$  is connected: connect  $x, x' \in \text{supp}(\mu)$  with an edge if  $x, x'$  differ in a single coordinate. More quantitatively, one considers a weighted version of this graph, where the weights correspond to re-sampling a coordinate in the distribution  $\mu$ ; with this in mind, connectedness is concerned with the second eigenvalue of this graph being bounded away from 1. The authors of [DHVY17] then suggested that to make more progress on multi-player parallel repetition one must tackle games that are not connected, and suggested the GHZ game as an example of such game.

The GHZ game is a special 3-player game that has its origin in physics (where it is known as quantum pseudo-telepathy). In the GHZ game the verifier samples a challenge  $(x, y, z) \in \mathbb{F}_2^3$  such that  $x + y + z = 0$ , sends  $x, y$  and  $z$  to Alice, Bob and Charlie respectively, receives answers  $f(x), g(y), h(z) \in \mathbb{F}_2$ , and accepts if  $f(x) + g(y) + h(z) = x \vee y \vee z$ . The GHZ game is easily seen to have classical value  $3/4$ , but all existing techniques seemed to fail to prove an effective upper bounds on the value of  $\text{GHZ}^{\otimes n}$ . Holmgren and Raz [HR20] were the first to tackle this problem,

---

<sup>1</sup>We remark that in the majority of applications, and especially in the area of probabilistically checkable proofs, one is often only concerned with projection games.

and they showed that  $\text{val}(\text{GHZ}^{\otimes n}) \leq n^{-\Omega(1)}$  using a combination of information theory and Fourier analysis. Later works simplified the proof for the GHZ game [GHM<sup>+</sup>21] and used these techniques to prove polynomial bounds for all 3-player games with binary questions [GHM<sup>+</sup>22, GMRZ22], namely games where  $\text{supp}(\mu) \subseteq \{0, 1\}^3$ . Finally, in [BKM23c] it was shown that  $\text{val}(\text{GHZ}^{\otimes n}) \leq 2^{-\Omega(n)}$  using techniques from additive combinatorics. We defer the reader to Section 2.1.2 for further discussion.

The primary goal of this work is to study parallel repetition for a very natural class of 3-XOR which includes within it the GHZ game. We show that the value of a 3-XOR games decays exponentially provided that the underlying distribution of question does not admit any Abelian embeddings into  $(\mathbb{Z}, +)$ , and in particular we recover the result of [BKM23c]. Our proof uses new techniques from Fourier analysis [BKM22, BKM23a, BKM23b, BKM24] as well as powerful results from additive combinatorics [Gow01].

## 1.1 3-XOR Games, Embeddings and Connectedness

To state our main result formally we require some setup. Suppose that  $\Sigma, \Gamma, \Phi$  are finite alphabets, and let  $\mu$  be a distribution on  $\Sigma \times \Gamma \times \Phi$ . Let  $t : \text{supp}(\mu) \rightarrow \mathcal{A}$  be a target function, where  $\mathcal{A}$  is a finite Abelian group. The pair  $(\mu, t)$  defines a 3-XOR game  $\mathcal{G}$ , wherein the verifier samples a challenge  $(x, y, z) \sim \mu$ , sends  $x, y$  and  $z$  to Alice, Bob and Charlie respectively, receives answers  $f(x), g(y), h(z) \in \mathcal{A}$  and accepts if  $f(x) + g(y) + h(z) = t(x, y, z)$ .

The *value* of an  $n$ -fold repeated game under this setup is defined as the maximum over all  $f : \Sigma^n \rightarrow \mathcal{A}^n, g : \Gamma^n \rightarrow \mathcal{A}^n, h : \Phi^n \rightarrow \mathcal{A}^n$  of

$$\Pr_{(x,y,z) \sim \mu^{\otimes n}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \quad \forall i = 1, 2, \dots, n]. \quad (1)$$

Let  $\mathcal{G}^{\otimes n}$  be the  $n$ -fold repeated game, with value  $\text{val}(\mathcal{G}^{\otimes n})$ . We are not able to establish parallel repetition for *all* 3-player XOR games, and we need to make some assumptions on the structure of the distribution  $\mu$ . We discuss why removing these assumptions may prove to be quite challenging in Section 1.3.

**Definition 1.1** (Embeddings). An embedding of a distribution  $\mu$  on  $\Sigma \times \Gamma \times \Phi$  into an abelian group  $A$  consists of three maps  $\sigma : \Sigma \rightarrow A, \gamma : \Gamma \rightarrow A, \phi : \Phi \rightarrow A$  such that  $\sigma(x) + \gamma(y) + \phi(z) = 0$  for all  $(x, y, z) \in \text{supp}(\mu)$ .

We say that  $\mu$  has no  $\mathbb{Z}$ -embeddings if all embeddings of  $\mu$  into  $\mathbb{Z}$  are *trivial* (in the sense that  $\sigma, \gamma, \phi$  must be constant maps). Throughout, we will assume that  $\mu$  has no non-trivial  $\mathbb{Z}$ -embeddings. If  $\mu$  has no  $\mathbb{Z}$ -embeddings, then it also satisfies a mild connectivity property.

**Definition 1.2** (Pairwise-connected). We say that a distribution  $\mu$  on  $\Sigma \times \Gamma \times \Phi$  is  $(x, y)$ -*pairwise-connected* if the graph on  $\Sigma \times \Gamma$  with an edge between  $(x, y)$  if there is some  $z$  with  $(x, y, z) \in \text{supp}(\mu)$  is connected. We say that  $\mu$  is pairwise-connected if it is  $(x, y)$ ,  $(y, z)$ , and  $(x, z)$  pairwise-connected.

**Lemma 1.3.** *If  $\mu$  has no  $\mathbb{Z}$ -embeddings, then  $\mu$  is pairwise-connected.*

*Proof.* If  $\mu$  is not  $(x, y)$ -connected, then there are partitions  $\Sigma \rightarrow \Sigma' \cup \Sigma''$  and  $\Gamma = \Gamma' \cup \Gamma''$  such that if  $(x, y, z) \in \text{supp}(\mu)$ , then  $(x, y) \in \Sigma' \times \Gamma'$  or  $(x, y) \in \Sigma'' \times \Gamma''$ . Thus, we can define the embedding  $\sigma(x) = \gamma(y) = 1$  for  $x \in \Sigma', y \in \Gamma', \sigma(x) = \gamma(y) = -1$  otherwise, and  $\phi(z) = 0$  for all  $z$ .  $\square$

We note that this notion of pairwise-connectivity is much weaker than connectivity in the context of [DHVY17], or even player-wise connectivity as defined in [GHM<sup>+</sup>22]. In particular, the GHZ game is pairwise-connected but neither connected or player-wise connected. This explains why the techniques for showing parallel repetition for the GHZ game differ significantly from the analyses of (player-wise) connected games, which are information-theoretic in nature.

## 1.2 Our Main Result

The main result of work is that games  $\mathcal{G}$  with XOR predicates over input distributions  $\mu$  without  $\mathbb{Z}$ -embeddings, have exponential decay under parallel repetition.

**Theorem 1.** *For a 3-player game  $\mathcal{G}$  over distribution  $\mu$  with XOR predicates, if  $\text{val}(\mathcal{G}) < 1$  and  $\mu$  has no  $\mathbb{Z}$ -embeddings, then there is a constant  $c := c(\mathcal{G}, \mu) > 0$  such that  $\text{val}(\mathcal{G}^{\otimes n}) \leq 2^{-cn}$ .*

Beyond the result itself, our proof method shows a connection between parallel repetition, at least in the case of XOR games, and techniques/questions in additive combinatorics. Specifically, our proof leverages an analytical stability result from [BKM24] for 3-wise correlations over pairwise-connected distributions with no  $\mathbb{Z}$ -embeddings. Thus, we (very speculatively) suspect that better understanding of these stability results (or extensions to higher arities  $k > 3$ ) may lead to more general 3-player or  $k$ -player parallel repetition results. We defer a more detailed discussion of our techniques to [Section 2](#).

## 1.3 $\mathbb{Z}$ -Embeddings, Pairwise-Connectivity and Beyond 3-XOR Games

The current paper only deals with distributions  $\mu$  that do not admit  $(\mathbb{Z}, +)$ -embeddings, and this is due to the technical limitations in our analytical machinery. In a future work, we show how to prove upper bounds for the  $n$ -fold repeated game under the weaker assumption that  $\mu$  is pairwise connected. In this case though, we can only show that the value of the  $n$ -fold repeated game decays as  $2^{-n^\varepsilon}$  for some  $\varepsilon > 0$  (as opposed to a purely exponential bound as in [Theorem 1](#)). In fact, this result holds for a class of games that strictly contains the class of 3-XOR games.

Proving results for 3-XOR games on distributions  $\mu$  that are not pairwise connected however seems to require additional ideas, since the type of analytical tools we use do not seem to have analogs for in such distributions.

## 2 Overview of Proof of [Theorem 1](#)

In this section we give an overview of the proof of [Theorem 1](#).

### 2.1 High-Level Structure of Argument

By the fundamental theorem of finite Abelian groups, to show [Theorem 1](#) it suffices to consider the case where  $\mathcal{A}$  is a cyclic group of prime power order, say  $\mathcal{A} = \mathbb{Z}_m$  where  $m = p^k$  for a prime  $p$ . For intuition throughout this section, the reader can think of the game as being over  $\mathbb{Z}_2$ , though all our proofs work over other cyclic groups. Even if the original game is over  $\mathbb{Z}_2$ , our proofs ultimately require defining functions over larger cyclic groups (e.g.,  $\mathbb{Z}_{2^k}$ ), just as the analysis of the GHZ game of [BKM23c] looked at functions over  $\mathbb{Z}_4$ .

#### 2.1.1 Step 1: Reducing to an Analytic Statement and Nonembeddable Targets

The first step of the proof is to arithmetize the probability that a set of strategies  $f, g, h$  win  $\mathcal{G}^{\otimes n}$  as an analytic quantity. Throughout, let  $\omega$  be the complex  $m$ th root of unity.

**Lemma 2.1.** *Let  $f : \Sigma^n \rightarrow \mathbb{Z}_m^n, g : \Gamma^n \rightarrow \mathbb{Z}_m^n, h : \Phi^n \rightarrow \mathbb{Z}_m^n$  be strategies. For a vector  $S \in \mathbb{Z}_m^n$  define  $F_S(x) = \omega^{\sum_{i \in [n]} S_i f(x)_i}$ , and  $G_S, H_S$  analogously, where  $\omega$  is the primitive root of unity of*

order  $m$ . Define  $T_S(x, y, z) = \omega^{-\sum_{i \in [n]} S_i t(x_i, y_i, z_i)}$ . The winning probability of the strategies  $f, g, h$  is exactly given by:

$$\mathbb{E}_{S \sim \mathbb{Z}_m^n} \mathbb{E}_{(x, y, z) \sim \mu^{\otimes n}} [F_S(x) G_S(y) H_S(z) T_S(x, y, z)].$$

**Lemma 2.1** follows from simply expanding out the definitions of  $F_S, G_S, H_S, T_S$  and reversing the order of the expectations.

From here, it is beneficial to consider 2-player XOR games, and give a simple and direct proof of parallel repetition. For the following result, let  $\mathbb{D}$  denote the unit disk in the complex plane, and  $T(x, y) = \omega^{-t(x, y)}$ .

**Lemma 2.2** (Two-dimensional case). *For a connected two player XOR game  $\mathcal{G}$  with distribution  $\mu$  over  $\Sigma \times \Gamma$  and target function  $t : \text{supp}(\mu) \rightarrow \mathbb{Z}_m$  with value less than 1, there is constant  $c := c(\mathcal{G}, \mu) > 0$  such that for any  $F : \Sigma^n \rightarrow \mathbb{D}, G : \Gamma^n \rightarrow \mathbb{D}$  we have*

$$\left| \mathbb{E}_{(x, y) \sim \mu^{\otimes n}} \left[ F(x) G(y) \prod_{i=1}^n T(x_i, y_i) \right] \right| \leq 2^{-cn}.$$

The intuition is that the quantity we wish to bound is an inner product of  $F, G$  with respect to  $T^{\otimes n}$  (where  $T$  is thought of as a matrix), and we know that singular value of matrices tensorize.

*Proof.* Let  $\mathbf{D}_x = \text{diag}(\mu_x)$  and  $\mathbf{D}_y = \text{diag}(\mu_y)$ . Note that  $\|(\mathbf{D}_x^{1/2})^{\otimes n} F\|_2, \|(\mathbf{D}_y^{1/2})^{\otimes n} G\|_2 \leq 1$ . Thus, it suffices to show that for all  $A : \Sigma \rightarrow \mathbb{D}$  and  $B : \Gamma \rightarrow \mathbb{D}$  with  $\|\mathbf{D}_x^{1/2} A\|_2 = \|\mathbf{D}_y^{1/2} B\|_2 = 1$ , that

$$|\mathbb{E}_{(x, y) \sim \mu} [A(x) B(y) T(x, y)]| < 1.$$

This implies then that the maximum singular value of  $\mathbf{D}_x^{-1/2} T \mathbf{D}_y^{-1/2}$  is less than 1, and the lemma follows because the maximum singular value tensorizes. Note that by the Cauchy-Schwarz inequality,

$$|\mathbb{E}_{(x, y) \sim \mu} [A(x) B(y) T(x, y)]| \leq \|\mathbf{D}_x^{1/2} A\|_2 \|\mathbf{D}_y^{1/2} B\|_2 = 1,$$

with equality only if  $|A(x)| = |B(y)|$  and  $A(x) B(y) T(x, y) = 1$  for all  $(x, y) \in \text{supp}(\mu)$ . By connectivity of  $\mu$ , the former condition and  $\|\mathbf{D}_x^{1/2} A\|_2 = 1$  implies that  $|A(x)| = 1$  for all  $x \in \Sigma$ , and  $|B(y)| = 1$  for all  $y \in \Gamma$ . Let  $A(x) = e^{2\pi i a(x)}, B(y) = e^{2\pi i b(y)}$ . Then the latter condition gives that  $a(x) + b(y) = t(x, y)/m \pmod{1}$  for all  $(x, y) \in \text{supp}(\mu)$ . By shifting  $a(x), b(y)$  to be multiples of  $1/m$ , this implies that  $t$  has value 1, a contradiction.  $\square$

This implies parallel repetition for 2-player XOR games. Indeed, we can assume the game is connected. Using **Lemma 2.2**, for any  $S \in \mathbb{Z}_m^n$  we have  $|\mathbb{E}_{(x, y) \sim \mu^{\otimes n}} [F_S(x) G_S(y) T_S(x, y)]| \leq 2^{-c|S|}$ , and as  $|S| \geq \Omega(n)$  on average over the choice  $S \sim \mathbb{Z}_m^n$  we get from **Lemma 2.1** an exponential decay for the value of the 2-player XOR game.

Can we carry out such an argument for 3-player games? The first piece we would need is an analogue of **Lemma 2.2**. In fact, the naïve generalization of **Lemma 2.2** to 3-functions is *false*, as we discuss in more detail in **Section 2.1.2**. The condition we actually require on  $t$  and  $T$  comes from looking at the end of the proof of **Lemma 2.2**, where we argued that if  $A(x) B(y) = T(x, y)$  for all  $(x, y) \in \text{supp}(\mu)$ , then  $t$  has a strategy of value 1. This is false in the 3-player case. Instead, for the 3-player setting, the assumption we require on  $T$  is exactly that there are no functions  $A : \Sigma \rightarrow \mathbb{D}, B : \Gamma \rightarrow \mathbb{D}, C : \Phi \rightarrow \mathbb{D}$  such that  $A(x) B(y) C(z) T(x, y, z) = 1$  for all  $(x, y, z) \in \text{supp}(\mu)$ . In this case, we say that  $T$  is *nonembeddable over  $\mathbb{D}$* . In the case that  $T$  is nonembeddable over  $\mathbb{D}$ , we are able to show that this statement tensorizes by following the work of [BKM23a] (though some adaptations are needed to get a proper exponential decay):

**Theorem 2.** *Let  $T : \text{supp}(\mu) \rightarrow \mathbb{D}$  be pairwise-connected and nonembeddable over  $\mathbb{D}$ . Then there is a constant  $c := c(T, \mu) > 0$  such that for all functions  $F : \Sigma^n \rightarrow \mathbb{D}, G : \Gamma^n \rightarrow \mathbb{D}, H : \Phi^n \rightarrow \mathbb{D}$ ,*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} \left[ F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i) \right] \right| \leq 2^{-cn}.$$

The proof of [Theorem 2](#) is substantially more challenging than its two-dimensional counterpart stated in [Lemma 2.2](#), and we overview the ideas in [Section 2.2](#). Combining this with [Lemma 2.1](#) shows that for target functions  $t$  that are nonembeddable over  $\mathbb{D}$ , parallel repetition with exponential decay holds.

### 2.1.2 Step 2: Barriers to Analytic Argument, and Linear Structure

Next, we discuss a case where  $T$  is embeddable over  $\mathbb{D}$ , but the game  $\mathcal{G}$  still has value less than 1. In this case, we are not able to apply [Theorem 2](#), and we must resort to a different set of techniques.

The most natural example of a 3-XOR game in this case is the GHZ game, defined as follows. The distribution  $\mu$  over  $\mathbb{Z}_2^3$  is uniform over  $\{(x, y, z) : x + y + z = 0\}$ , and strategies  $f, g, h$  win on  $x, y, z$  if  $f(x) + g(y) + h(z) = x \vee y \vee z \pmod{2}$ . It is easy to check that the value of this game is  $3/4$ . The tensor  $T$  defined by  $T(x, y, z) = (-1)^{x \vee y \vee z}$ , and has values

$$T(0, 0, 0) = 1, T(1, 1, 0) = -1, T(1, 0, 1) = -1, T(0, 1, 1) = -1.$$

This is embeddable over  $\mathbb{D}$  by taking  $A = B = C$ , and  $A(x) = i^x = e^{\frac{1}{2}\pi i \cdot x}$ . Another way of viewing this is that

$$\frac{1}{2}(x + y + z) = x \vee y \vee z = t(x, y, z) \pmod{2} \quad \forall (x, y, z) \in \text{supp}(\mu).$$

Thus,  $2t(x, y, z) = x + y + z \pmod{4}$ . Thus, the winning condition  $f(x) + g(y) + h(z) = t(x, y, z) \pmod{2}$  can be rewritten  $(2f(x) - x) + (2g(y) - y) + (2h(z) - z) = 0 \pmod{4}$ . This is the key observation in [\[BKM23c\]](#), and combining it with powerful tools from additive combinatorics (which we overview in [Section 2.1.4](#)) yields exponential decay for parallel repetition of the GHZ game.

Our argument in the general case that  $T$  is embeddable follows a similar strategy and uses additive combinatorics as well. Note that if  $A(x)B(y)C(z)T(x, y, z) = 1$  for

$$A(x) = e^{2\pi i a'(x)}, B(y) = e^{2\pi i b'(y)}, C(z) = e^{2\pi i c'(z)}, T(x, y, z) = e^{-2\pi i t(x, y, z)/m},$$

then we get that  $a'(x) + b'(y) + c'(z) = t(x, y, z)/m \pmod{1}$ . Assuming that  $a(x), b(y), c(z)$  are all rational with denominator dividing  $Nm < \infty$  (which we can assume) we get that (writing  $a'(x) = a(x)/(Nm)$  and similar for  $b'(y), c'(z)$ ),

$$a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{Nm},$$

for functions  $a : \Sigma \rightarrow \mathbb{Z}_{Nm}, b : \Gamma \rightarrow \mathbb{Z}_{Nm}, c : \Phi \rightarrow \mathbb{Z}_{Nm}$ . Thus the winning condition  $f(x) + g(y) + h(z) = t(x, y, z) \pmod{m}$  can be rewritten as:

$$(N \cdot f(x) - a(x)) + (N \cdot g(y) - b(y)) + (N \cdot h(z) - c(z)) = 0 \pmod{Nm}.$$

This is a very similar setup to the GHZ game we described above, with one critical difference. In the GHZ game, the distribution  $\mu$  itself was a subspace of  $\mathbb{Z}_2^3$  consisting of points with sum 0, while our input distribution  $\mu$  is still somewhat arbitrary. This arbitrariness prevents us from applying tools from additive combinatorics, and to remedy the situation, we must gain more control over the structure of the distribution  $\mu$ . In the next section, we discuss how to use ideas from [\[BKM24\]](#) to reduce to the case where  $\mu$  is uniform over  $\{(x, y, z) \in \mathcal{A}^3 : x + y + z = 0\}$ , for some (possibly very large) finite Abelian group  $\mathcal{A}$ .

### 2.1.3 Step 3: The Master Embedding and Reducing to Functions Constant on the Master Embedding

The group  $\mathcal{A}$  discussed at the end of the previous section will be called the *master group*, which we define now. Fix a sufficient large integer  $r$ , and let  $(\sigma_1, \gamma_1, \phi_1), \dots, (\sigma_s, \gamma_s, \phi_s)$  be all embeddings of  $\mu$  into cyclic groups  $\mathcal{A}_1, \dots, \mathcal{A}_s$  of (prime power) order at most  $r$ , with  $\sigma_i(x^*) = \gamma_i(y^*) = \phi_i(z^*) = 0$  for some fixed  $x^* \in \Sigma, y^* \in \Gamma, z^* \in \Phi$ , which we can assume by shifting each embedding. Then, we define the master embedding as  $\sigma_m(x) := (\sigma_1(x), \dots, \sigma_s(x))$  (Definition 6.1), and set  $\mathcal{A} \subseteq \mathcal{A}_1 \times \dots \times \mathcal{A}_s$  as the subgroup generated by  $\{\sigma_m(x) : x \in \Sigma\}$ . The key property of  $\sigma_m$  and  $\mathcal{A}$  is that if  $\mu$  has no  $\mathbb{Z}$ -embeddings, then for  $r$  sufficiently large the following holds. For all  $x, x' \in \Sigma$ ,  $\sigma(x) = \sigma(x')$  for all embeddings  $\sigma$  if and only if  $\sigma_m(x) = \sigma_m(x')$  (Lemma 6.2), and analogously for  $\gamma_m, \phi_m$ .

Recall that at the end of Section 2.1.2, we rewrote the winning strategy of the game as  $\tilde{f}(x) + \tilde{g}(y) + \tilde{h}(z) = 0$  for  $(x, y, z) \in \text{supp}(\mu)$ , where  $\tilde{f}(x) = N \cdot f(x) - a(x)$ ,  $\tilde{g}(y) = N \cdot g(y) - b(y)$ , and  $\tilde{h}(z) = N \cdot h(z) - c(z)$ . Thus,  $f, g, h$  win the game exactly when  $\tilde{f}, \tilde{g}, \tilde{h}$  act like an embedding on  $\mu$ . The general philosophy of the master embedding asserts that to get such behaviour, the values of the functions  $\tilde{f}(x)$ ,  $\tilde{g}(y)$  and  $\tilde{h}(z)$  must morally depend only on  $\sigma_m(x)$ ,  $\gamma_m(y)$  and  $\phi_m(z)$  respectively. Indeed, we are able to transform the functions  $\tilde{f}$ ,  $\tilde{g}$  and  $\tilde{h}$  to functions with similar winning probabilities that are constant on the master embeddings. Abusing of notations we assume that  $\tilde{f}$ ,  $\tilde{g}$  and  $\tilde{h}$  are such functions to begin with, namely  $\tilde{f}(x) = \tilde{f}(x')$  if  $\sigma_m(x) = \sigma_m(x')$ , and analogously for  $\tilde{g}$  and  $\tilde{h}$ . This way, regardless of the original distribution  $\mu$ , we can imagine that we are working with the set of challenges  $\{(\sigma_m(x), \gamma_m(y), \phi_m(z)) \mid (x, y, z) \in \text{supp}(\mu)\}$ , and by changing variables we may assume that to begin with we have that

$$\text{supp}(\mu) \subseteq \{(x, y, z) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A} : x + y + z = 0\},$$

(where  $x + y + z = 0$  because the master embeddings satisfies this equation). Formalizing this intuition is much more challenging, and is the main content of Section 6. In particular, it requires a version of our analytic result Theorem 2 in this setting (Theorem 5). We remark that a version of this was shown in [BKM24], but in order to achieve our tighter quantitative bounds of  $\exp(-cn)$  for parallel repetition, we need to give a tighter analysis and a slightly different proof adapted to our setting.

In the above paragraph we gave intuition on how to reduce to the case where  $\text{supp}(\mu)$  is a subset of  $(x, y, z) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}$  with  $x + y + z = 0$ . However, we actually wish to have  $\text{supp}(\mu)$  to be exactly this set. To achieve this, we use path tricks (discussed below in Section 2.2) to enlarge the input alphabets and distribution  $\mu$ , without changing the master group  $\mathcal{A}$ , and without reducing the winning probability of the game by much. This idea of using path tricks to enlarge the distribution was exactly used in [BKM24] in the setting of showing analytic theorems, and an identical approach works for parallel repetition too.

### 2.1.4 Step 4: Applying an Additive Combinatorics Argument

Finally, we use techniques from additive combinatorics to solve the case where  $\text{supp}(\mu)$  is exactly  $\{(x, y, z) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A} : x + y + z = 0\}$ . Our main theorem is the following.

**Theorem 3.** *Let  $\mathcal{A}$  be a finite abelian group and  $t : \mathcal{A}^3 \rightarrow \mathbb{Z}_{p^k}$ . Let  $N > 1$  be minimal so that there are functions  $a, b, c : \mathcal{A} \rightarrow \mathbb{Z}_{p^k N}$  with  $a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{p^k N}$  for all  $x + y + z = 0$ . Then there is a constant  $c = c(t, \mathcal{A}) > 0$  such that for any  $f, g, h : \mathcal{A}^n \rightarrow \mathbb{Z}_{p^k}$ ,*

$$\Pr_{\substack{(x, y, z) \in (\mathcal{A}^n)^3 \\ x + y + z = 0}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \pmod{p^k}, i = 1, 2, \dots, n] \leq 2^{-cn}.$$



A simpler version was given for the GHZ game in [BKM23c]. At a high level (using the notation from the previous sections), the argument goes as follows. If  $\tilde{f}(x) + \tilde{g}(y) + \tilde{h}(z) = 0$  with probability at least  $\varepsilon$  over  $x+y+z = 0$ ,  $(x, y, z) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}$ , then using tools from additive combinatorics such as the Balog-Szemerédi-Gowers lemma, we deduce significant structural facts about  $\tilde{f}, \tilde{g}, \tilde{h}: \mathcal{A} \rightarrow \mathbb{Z}_{Nm}$  (Lemma 8.1). We ultimately combine this with the fact that  $\tilde{f}$  is not completely arbitrary, as it must take the form  $\tilde{f}(x) = N \cdot f(x) - a(x)$ , to arrive at a contradiction.

## 2.2 Proof of Main Analytic Results

### 2.2.1 Changing the distribution via random restrictions

In this section overview the tools needed to show our analytic results, focusing on Theorem 2 (our other analytical result, Theorem 5 involves similar ideas). A very useful trick throughout our proofs will be that we can alter the distribution over questions  $\mu$  (without changing its support), at the cost of reducing the number of repetitions  $n$  by a constant factor. For example, we can change the distribution to uniform over its support. Formally, we can show:

**Lemma 2.3.** *Let  $\mathcal{G}_\mu$  be a game with question distribution  $\mu = \delta\mu' + (1 - \delta)\nu$  for distributions  $\mu', \nu$ . Let  $\mathcal{G}_{\mu'}$  be the game with the same questions over distribution  $\mu'$ . Then*

$$\text{val}(\mathcal{G}_\mu^{\otimes n}) \leq \text{val}(\mathcal{G}_{\mu'}^{\otimes \delta n/2}) + 2^{-\Omega(\delta n)}.$$

*Proof.* Deferred to Section 3.3. □

### 2.2.2 Path Tricks

A natural approach towards showing Theorem 2 is to induct on  $n$ , where the base case  $n = 1$  follows exactly because  $T$  is nonembeddable over  $\mathbb{D}$ . However, we don't know how to induct on such a base case. We instead first work towards establishing a base case for functions  $F, G, H$  that are  $\ell_2$ -bounded, instead of  $\ell_\infty$ -bounded. To achieve this, the idea of [BKM23a] was to use a *path trick* (Section 4.1), which changes the distribution  $\mu$  to a larger distribution  $\mu^+$  over  $\Sigma' \times \Gamma \times \Phi$ , where  $\mu_{yz}^+$  has full support. In fact, we can assume the support is uniform, by applying the random restriction trick (Lemma 2.3).

Formally, the path trick takes an integer  $r$ , sets  $\Sigma^+ \subseteq \Sigma^{2^{r+1}-1}$  and for all  $y_1, z_1, y_2, z_2, \dots, y_{2^r}, z_{2^r}$  such that  $(y_i, z_i), (z_i, y_{i+1}) \in \text{supp}(\mu_{yz})$ , adds the following to  $\text{supp}(\mu^+)$ . Let  $(x_i, y_i, z_i) \in \text{supp}(\mu)$  for  $i = 1, \dots, 2^r$ , and  $(x'_i, y_{i+1}, z_i) \in \text{supp}(\mu)$  for  $i = 1, 2, \dots, 2^r - 1$ . Then,

$$(\{x_1, x'_1, x_2, \dots, x_{2^r}\}, y_1, z_{2^r}) \in \Sigma^+ \times \Gamma \times \Phi$$

is added to  $\text{supp}(\mu^+)$ . For sufficiently larger  $r$ , if  $\mu_{yz}$  was initially pairwise-connected, then  $\mu_{yz}^+$  has full support. Also, we can show that if there were functions  $F, G, H$  with

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} \left[ F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i) \right] \right| \geq \varepsilon,$$

then there are functions  $\tilde{F}: (\Sigma^+)^n \rightarrow \mathbb{D}$ ,  $\tilde{G}: \Gamma^n \rightarrow \mathbb{D}$ , and  $\tilde{H}: \Phi^n \rightarrow \mathbb{D}$  with

$$\left| \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} \left[ \tilde{F}(x)\tilde{G}(y)\tilde{H}(z) \prod_{i=1}^n T^+(x_i, y_i, z_i) \right] \right| \geq \varepsilon^{O_r(1)},$$

for some tensor  $T^+$  that is still nonembeddable over  $\mathbb{D}$ , i.e., we only incur polynomial loss. We prove this in [Lemma 4.6](#).

At this point, it is worth discussing how the path trick formally helps with enlarging the distribution  $\mu$ , towards eventually ensuring that  $\text{supp}(\mu)$  is all  $x + y + z = 0$  for  $(x, y, z) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}$ . Towards this, we first calculate how embeddings evolve under the path trick. Given an embedding  $(\sigma, \gamma, \phi)$  of  $\mu$ , one can easily check that it uniquely induces an embedding on  $\mu^+$  of the form  $(\sigma^+, \gamma, \phi)$ , where

$$\sigma^+(\{x_1, x'_1, x_2, \dots, x_{2r}\}) = \sigma(x_1) - \sigma(x'_1) + \dots - \sigma(x'_{2r-1}) + \sigma(x_{2r}).$$

Because  $(\{x, x, \dots, x\}, y, z) \in \text{supp}(\mu^+)$ , and  $\sigma^+(\{x, \dots, x\}) = \sigma(x)$ , we know that the image of  $\sigma^+$  on  $\mu$  is nondecreasing. We can also show that by carefully applying a sequence of path tricks, that we can increase its size until the image is a subgroup ([Lemma 6.8](#)).

### 2.2.3 Establishing an Inductive Hypothesis

When  $\mu_{yz}$  is a product distribution, we can establish a weak base case via the Cauchy-Schwarz inequality:

$$\begin{aligned} & |\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)T(x,y,z)]| \\ & \leq (\mathbb{E}_{(x,y,z) \sim \mu} |F(x)|^2 \mathbb{E}_{(x,y,z) \sim \mu} |G(y)|^2 |H(z)|^2)^{1/2} = \|F\|_2 \|G\|_2 \|H\|_2, \end{aligned} \quad (2)$$

where the final equality critically used that  $\mu_{yz}$  is a product distribution. If the right hand side was instead  $(1 - \delta)\|F\|_2 \|G\|_2 \|H\|_2$ , then we would be able to actually perform the induction on  $n$  via singular value decomposition (SVD) on  $F, G, H$ , thinking of  $F : \Sigma^n \rightarrow \mathbb{C}$  as a matrix in  $\mathbb{C}^{\Sigma \times \Sigma^{n-1}}$ . Thus, we want to say something like: if  $T$  is nonembeddable over  $\mathbb{D}$ , then equality in (2) cannot hold. Equality in (2) holds only if  $\overline{F(x)} = \theta G(y)H(z)T(x,y,z)$  for some  $\theta \in \mathbb{C}$ , for all  $(x,y,z) \in \text{supp}(\mu)$ . This looks quite close to  $T$  having an embedding over  $\mathbb{D}$ , except that  $F(x), G(y), H(z)$  may be 0.

In [\[BKM23a\]](#), this issue (which they called the *Horn-SAT obstruction*), was handled as follows. First, perform more path tricks to the distribution to transform it to a larger distribution (which we relabel  $\mu$  for convenience) satisfying a *relaxed base case* (see [Definition 4.10](#)). Informally, the distribution  $\mu$  over  $\Sigma \times \Gamma \times \Phi$  satisfies the relaxed base case, if for some subset  $\Sigma' \subseteq \Sigma$ , any function  $F$  with nontrivial variance on  $\Sigma'$ , and  $G, H$ , satisfy the base case, i.e.,

$$|\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)T(x,y,z)]| \leq (1 - \delta)\|F\|_2 \|G\|_2 \|H\|_2.$$

We too need to establish a relaxed base case, and we do so by analogous ideas.

### 2.2.4 High and Low Degree Parts

With the relaxed base case in hand, we are able to perform the induction on functions  $F$  where during the SVD, the the resulting functions all have nontrivial variance on  $\Sigma'$ . To handle general functions  $F$ , we decompose them via an Efron-Stein decomposition ([Section 3.2](#)), into functions that have equal degrees on  $\Sigma'$ . The reader can think of an Efron-Stein as a version of a Fourier transform, on a set  $\Sigma$  without additive structure. Just as with a Fourier transform, high-degree functions have high influence/variance on each coordinate. Thus, starting from a high-degree function, one show that while performing SVDs in the induction, it remains high-degree for many iterations, allowing us to perform the induction. This allows us to handle the high-degree parts of  $F$ .

For the lower-degree parts of  $F$ , their variance on  $A'$  is low, so intuitively  $F$  does not change much under the noise operator  $\mathcal{T}^{\text{eff}}$  which mixes inputs in  $A'$ . Applying  $\mathcal{T}^{\text{eff}}$  (and taking random restrictions) makes the input distribution symmetric over  $A'$ . This case ends up reducing to the two-variable case on  $(y, z)$ , where an exponential decay bound is known (see [Lemma 2.2](#)).

### 2.2.5 The Case with Complex Embeddings

The above discussion was for showing [Theorem 2](#), when  $T$  was nonembeddable over  $\mathbb{D}$ . As discussed in [Section 2.1.3](#), we require an alternate analytic theorem which reduces to the case where  $\mu$  is uniform over  $x + y + z = 0$ ,  $(x, y, z) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}$ . We describe the idea for adapting to this case. The approach for this case is largely the same as for proving [Theorem 2](#) as discussed above: establish a variant of the relaxed base case, and split into high and low-degree parts. Here, high and low-degree are with respect to the noise operator (which we call  $\mathcal{T}^{\text{modest}}$ ) which mixes over  $x, x'$  with the same master embedding, i.e.,  $\sigma_m(x) = \sigma_m(x')$ . Once again, the contribution of the high-degree parts vanishes, and the low-degree parts intuitively correspond to applying  $\mathcal{T}^{\text{modest}}$ . This is exactly what we wanted to show: that  $x, x'$  with  $\sigma_m(x) = \sigma_m(x')$  are (approximately) indistinguishable from the perspective of strategies  $f$ , and thus we can reduce the input alphabets to  $\mathcal{A}$  themselves.

## 3 Preliminaries

### 3.1 General Notation

Let  $[n] = \{1, \dots, n\}$ . For a real number  $x$  we write  $\|x\|_{\mathbb{R}/\mathbb{Z}} := \min_{z \in \mathbb{Z}} |x - z|$ . We use  $a \lesssim b$  to denote  $a = O(b)$  and  $a \gtrsim b$  to denote  $a = \Omega(b)$ . We let  $\Phi$  denote complex numbers, and let  $\mathbb{D}$  denote the unit disk. We let  $\text{supp}(\mu)$  denote the support of a distribution  $\mu$ .

**Matrices.** We briefly use matrices in a few places, when we are working with “two-player” like games, such as merging symbols ([Lemma 4.7](#)). We will denote matrices using boldface. For a matrix  $\mathbf{A}$ , we let  $\mathbf{A}^{\otimes n}$  denote its  $n$ -th Kronecker power. For a vector  $v$ , we let  $\text{diag}(v)$  denote the diagonal matrix whose diagonal entries are  $v_i$ .

### 3.2 Efron-Stein Decomposition

We will use the standard Efron-Stein decomposition on the product space  $L_2(\Sigma^n, \mu_x^{\otimes n})$ . For a function  $F \in L_2(\Sigma^n, \mu_x^{\otimes n})$  and  $S \subseteq [n]$ , we let  $F^{=S}$  be the homogeneous function of degree  $|S|$  that depends on the coordinates in  $S$ . Let  $F^{=i} = \sum_{|S|=i} F^{=S}$ .

**Effective degree and influences.** In our relaxed base case ([Lemma 4.11](#)), we will consider the variance on a subset  $\Sigma' \subseteq \Sigma$ . Thus, we define *effective* degrees and influences of a function  $f$ , which are the influence and degrees that  $\Sigma'$  contributes.

Create an orthonormal basis  $B$  for  $L_2(\Sigma, \mu_x)$  of the form  $B = B_1 \cup B_2$ , where  $B_1$  is an orthonormal basis for functions that are constant on  $\Sigma'$ , and  $B_2$  is an orthonormal basis for functions supported on  $\Sigma'$  that are orthogonal to the constant function.

For any function  $F \in L_2(\Sigma^n, \mu_x^{\otimes n})$  we can write  $F = \sum_{\chi \in B^{\otimes n}} \widehat{F}(\chi)\chi$ , where  $\widehat{F}(\chi) = \langle F, \chi \rangle$ . Let  $\chi_{\text{const}}$  be the constant function on  $\Sigma$ , which we assume is in  $B_1$ . For a character  $\chi \in B^{\otimes n}$ , we define its degree and effective degree as

$$\text{deg}(\chi) = |\{i \in [n] : \chi_i \neq \chi_{\text{const}}\}| \quad \text{and} \quad \text{effdeg}(\chi) = |\{i \in [n] : \chi_i \in B_2\}|.$$

Note that  $\text{deg}(\chi) \geq \text{effdeg}(\chi)$ .

This allows us to define the decomposition into effective degree homogeneous parts. Let

$$F^{\text{eff}=i} = \sum_{\chi \in B^{\otimes n}, \text{effdeg}(\chi)=i} \widehat{F}(\chi)\chi.$$

We now define the effective influence.

$$I_{\text{eff}}[F] = \sum_{i \in [n]} I_{i, \text{eff}}[F] \quad \text{for} \quad I_{i, \text{eff}}[F] = \mathbb{E}_{(x, y) \in \Sigma^n} [|F(x) - F(y)|^2],$$

where  $(x, y)$  are sampled as follows. Sample  $x$  distributed as  $\mu_x^{\otimes n}$ , and set  $y_j = x_j$  for all  $j \neq i$ . If  $x_i \in \Sigma \setminus \Sigma'$  set  $y_i = x_i$ , and if  $x_i \in \Sigma'$  then sample  $y_i \in \Sigma'$  independently.

Analogous to the case of standard influences, we know that

$$I_{i, \text{eff}}[F] = \sum_{\chi \in B^{\otimes n}, \chi_i \in B_2} |\widehat{F}(\chi)|^2 \quad \text{and} \quad I_{\text{eff}}[F] = \sum_{\chi \in B^{\otimes n}} \text{effdeg}(\chi) |\widehat{F}(\chi)|^2 = \sum_{0 \leq i \leq n} i \|F^{\text{eff}=i}\|_2^2.$$

We finally define the effective noise operator, which we denote as  $\mathcal{T}_{1-\rho}^{\text{eff}}$ . Define  $\mathcal{T}_{1-\rho}^{\text{eff}} F(x) = \mathbb{E}_{y \sim N_{1-\rho}^{\text{eff}}(x)} [F(y)]$  where  $y \sim N_{1-\rho}^{\text{eff}}(x)$  is distributed as follows. For each  $i \in [n]$ , with probability  $1 - \rho$  set  $y_i = x_i$ . With probability  $\rho$ , if  $x_i \in \Sigma \setminus \Sigma'$ , then set  $y_i = x_i$ . Otherwise if  $x_i \in \Sigma'$ , then  $y_i$  is sampled independently according to  $\mu_x$  conditioned on being in  $\Sigma'$ . As with the standard noise operator, we have

$$\mathcal{T}_{1-\rho}^{\text{eff}} F = \sum_{0 \leq i \leq n} (1 - \rho)^i F^{\text{eff}=i}.$$

### 3.3 Random Restrictions

Consider distributions  $\mu, \mu', \nu$  on  $\Sigma \times \Gamma \times \Phi$  with  $\mu = \alpha\mu' + (1 - \alpha)\nu$ . We can take a random restriction of functions  $F, G, H$  on  $\Sigma^n, \Gamma^n, \Phi^n$  as follows. Let  $I \subseteq [n]$ , where each  $i \in I$  with probability  $1 - \alpha$ . Sample  $p := (x(p), y(p), z(p)) \in \nu^{\otimes I}$ . We may define the restricted functions  $F_{I \rightarrow p} : \Sigma^{[n] \setminus I} \rightarrow \mathbb{C}$  as  $F_{I \rightarrow p}(x) = F(x, x(p))$ , and similar for  $G, H$ . Note that over the distribution of  $(I, p)$ , and  $x \sim (\mu')^{\otimes [n] \setminus I}$ , the distribution of  $(x, x(p))$  is identical to  $\mu_x^{\otimes n}$ . At a high level, this lets us move from a distribution  $\mu$  on  $[n]$  to a distribution  $\mu'$  on about  $\alpha n$  coordinates, as long as  $\alpha\mu' \leq \mu$  pointwise. This was exactly the content of [Lemma 2.3](#), which we now prove.

*Proof of Lemma 2.3.* Take a random restriction of  $\mu$  as described above. The probability that  $|[n] \setminus I| < \delta n/2$  is at most  $2^{-\Omega(\delta n)}$ , by a Chernoff bound. Thus, for a strategy  $f, g, h$  achieving  $\text{val}(\mathcal{G}_\mu^{\otimes n})$  we know that

$$\begin{aligned} \text{val}(\mathcal{G}_\mu^{\otimes n}) &= \Pr_{(x, y, z) \sim \mu^{\otimes n}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \quad \forall i = 1, \dots, n] \\ &= \Pr_{\substack{I \sim_{1-\delta} [n] \\ (x(p), y(p), z(p)) \sim \nu^{\otimes I}}} \Pr_{(x, y, z) \sim (\mu')^{\otimes [n] \setminus I}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \quad \forall i = 1, \dots, n] \\ &\leq \Pr_{\substack{I \sim_{1-\delta} [n] \\ (x(p), y(p), z(p)) \sim \nu^{\otimes I}}} \Pr_{(x, y, z) \sim (\mu')^{\otimes [n] \setminus I}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \quad \forall i \in [n] \setminus I] \\ &\leq \mathbb{E}_{I \sim_{1-\delta} [n]} \text{val}(\mathcal{G}_{\mu'}^{\otimes |I|}) \leq \text{val}(\mathcal{G}_{\mu'}^{\otimes \delta n/2}) + 2^{-\Omega(\delta n)}. \quad \square \end{aligned}$$

## 4 No Complex Solutions: Reductions

The main goal of the following two sections is to show [Theorem 2](#), restated below.

**Theorem 2.** *Let  $T : \text{supp}(\mu) \rightarrow \mathbb{D}$  be pairwise-connected and nonembeddable over  $\mathbb{D}$ . Then there is a constant  $c := c(T, \mu) > 0$  such that for all functions  $F : \Sigma^n \rightarrow \mathbb{D}, G : \Gamma^n \rightarrow \mathbb{D}, H : \Phi^n \rightarrow \mathbb{D}$ ,*

$$\left| \mathbb{E}_{(x, y, z) \sim \mu^{\otimes n}} \left[ F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i) \right] \right| \leq 2^{-cn}.$$

Combining this with [Lemma 4.2](#) directly implies an exponential decay bound on 3-player XOR games where the target function  $t$  does not have linear structure over any abelian group.

**Corollary 4.1.** *Let  $\mathcal{G}$  be a game given by  $\mu, t$ . If  $\mu$  is pairwise-connected, and  $T(x, y, z) = \omega^{t(x, y, z)}$  for  $\omega = \exp(2\pi i/m)$  is nonembeddable over  $\mathbb{D}$  then there is a constant  $c := c(\mathcal{G}, \mu) > 0$  with  $\text{val}(\mathcal{G}^{\otimes n}) \leq 2^{-cn}$ .*

*Proof.* For  $S \in \mathbb{Z}_m^n$ , let  $1(S) = |\{i : S_i = 1\}|$ . Then by [Theorem 2](#) and the triangle inequality, the winning probability as written in [Lemma 2.1](#) at most

$$\mathbb{E}_{S \sim \mathbb{Z}_m^n} 2^{-c1(S)} \leq 2^{-c'n}. \quad \square$$

In the remainder of the section and the next we show [Theorem 2](#). Before starting, we make the trivial observation that if  $|T(x, y, z)| < 1$  for any  $(x, y, z) \in \text{supp}(\mu)$ , then the result is clear. So for all  $T$  encountered in the series of reductions we do later, we assume  $|T(x, y, z)| = 1$  for all  $(x, y, z) \in \text{supp}(\mu)$ .

**Lemma 4.2.** *If  $|T(x, y, z)| < 1$  for some  $(x, y, z) \in \text{supp}(\mu)$ , then there is  $c := c(\mu, T) > 0$  such*

$$\left| \mathbb{E}_{(x, y, z) \sim \mu^{\otimes n}} \left[ F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i) \right] \right| \leq 2^{-cn}.$$

*Proof.* This follows from the fact that  $\mathbb{E}_{(x, y, z) \sim \mu} [|T(x, y, z)|] < 1$ . □

## 4.1 Path Tricks

We start by defining what it means to do a path trick on  $x$ . The effect of this will be changing  $\Sigma$  to the larger  $\Sigma^+$ , the distribution to  $\mu^+$  over  $\Sigma^+ \times \Gamma \times \Phi$ , and the tensor to  $t : \text{supp}(\mu^+) \rightarrow [-1, 1]$ . The goal is to ensure that  $\mu^+$  has full support over  $(y, z)$ .

We define  $\mu^+$  as follows.

**Definition 4.3** (Path trick). The following procedure takes a distribution  $\mu$  on  $\Sigma \times \Gamma \times \Phi$  and generates  $\mu^+$  on  $\Sigma^+ \times \Gamma \times \Phi$ , which we call applying an “ $x$  path trick with  $r$  steps” to  $\mu$ .

1. Sample  $y_1 \sim \mu_y$ .
2. Sample  $(x_1, z_1)$  from  $\mu$ , conditioned on  $y = y_1$ .
3. Sample  $(x'_1, y_2)$  from  $\mu$ , conditioned on  $z = z_1$ .
4. Repeat, and in the final step sample  $(x_{2^r}, y_{2^{r-1}+1})$  from  $\mu$  conditioned on  $z = z_{2^{r-1}}$ .

Now let  $\Sigma^+ \subseteq \Sigma^{2^r-1}$ .  $\mu^+$  is over the sequences  $((x_1, x'_1, \dots, x'_{2^{r-1}-1}, x_{2^r}), y_1, z_{2^{r-1}})$  generated by the process above.

Throughout, we will let  $\vec{x} = (x_1, x'_1, \dots, x'_{2^{r-1}-1}, x_{2^r})$ .

Now we define  $T^+(\vec{x}, y, z)$  as follows:

$$T^+(\vec{x}, y, z) = \mathbb{E}_{\substack{x_1, x'_1, \dots, x'_{2^{r-1}-1}, x_{2^r} \\ y_1, \dots, y_{2^{r-1}} \\ z_1, \dots, z_{2^{r-1}}}} \left[ \prod_{i=1}^{2^r-1} T(x_i, y_i, z_i) \prod_{i=1}^{2^{r-1}-1} \overline{T(x'_i, y_{i+1}, z_i)} \right] \Bigg|_{\vec{x} = (x_1, x'_1, \dots, x'_{2^{r-1}-1}, x_{2^r}), y = y_1, z = z_{2^{r-1}}}$$

Note that  $T^+$  is still nonembeddable over  $\mathbb{D}$ .

**Lemma 4.4.** *As defined above, if  $T$  is nonembeddable over  $\mathbb{D}$  then  $T^+$  is nonembeddable over  $\mathbb{D}$ . If  $(y, z)$  was pairwise-connected in  $\mu$  and  $2^{r-1} \geq \min\{|\Gamma|, |\Phi|\}$ , then the support of  $(y, z)$  in  $\mu^+$  is full.*

*Proof.* Define  $(x, \dots, x) = \bar{x} \in \Sigma^+$  for all  $x \in \Sigma$ . Note that either  $T^+(\bar{x}, y, z) = T(x, y, z)$ , or else  $|T^+(\bar{x}, y, z)| < 1$ , where we are trivially done by Lemma 4.2. So  $T^+$  is nonembeddable. The second claim on the support of  $(y, z)$  in  $\mu^+$  is clear.  $\square$

Performing the path trick maintains pairwise-connectivity.

**Lemma 4.5.** *If  $\mu$  is pairwise-connected, then  $\mu^+$  is pairwise-connected.*

*Proof.*  $(y, z)$ -pairwise-connectivity is clear, so we focus on  $(x, y)$ -pairwise-connectivity by symmetry. We first show that any two symbols  $y, y' \in \Gamma$  are connected. This is because  $\mu$  is pairwise-connected, and  $((x, x, \dots, x), y) \in \text{supp}(\mu^+)$  for any  $(x, y) \in \text{supp}(\mu)$ . Now, every  $\vec{x} \in \Sigma^+$  is adjacent to some  $y \in \Gamma$ , which completes the proof.  $\square$

We now show that performing the path trick can only decrease the value polynomially.

**Lemma 4.6.** *If there are functions  $F : \Sigma^n \rightarrow \mathbb{D}$ ,  $G : \Gamma^n \rightarrow \mathbb{D}$ ,  $H : \Phi^n \rightarrow \mathbb{D}$  with*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \geq \varepsilon,$$

*then for  $\mu^+, T^+$  as defined above, there are functions  $F^+ : (\Sigma^{2^r-1})^n \rightarrow \mathbb{D}$  and  $\tilde{H} : \Phi^n \rightarrow \mathbb{D}$  with*

$$\left| \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [F^+(x)G(y)\tilde{H}(z) \prod_{i=1}^n T^+(x_i, y_i, z_i)] \right| \geq \varepsilon^{2^r}.$$

*Proof.* We start by performing a simple reduction on  $H$ . Define

$$\tilde{H}(z) := \mathbb{E}_{(x,y,z') \sim \mu^{\otimes n}} [\overline{F(x)G(y)T(x,y,z)} | z' = z].$$

Then note that

$$\begin{aligned} \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z)T(x,y,z)] \right| &= \left| \mathbb{E}_{z \sim \mu_z} [H(z)\overline{\tilde{H}(z)}] \right| \leq \|H\|_2 \|\tilde{H}\|_2 \leq \|\tilde{H}\|_2 \\ &= \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)\tilde{H}(z)T(x,y,z)]^{1/2}. \end{aligned}$$

Hence  $\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)\tilde{H}(z)T(x,y,z)] \right| \geq \varepsilon^2$ . The remainder of the proof involves repeated application of Cauchy-Schwarz.

We show the following by induction on  $r$ :

$$\left| \mathbb{E}_{\substack{x_1, x'_1, \dots, x_{2^{r-1}}, x'_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}+1} \\ z_1, \dots, z_{2^{r-1}}} \left[ \prod_{i=1}^{2^{r-1}} F(x_i) \overline{F(x'_i)} \cdot G(y_1) \overline{G(y_{2^{r-1}+1})} \prod_{i=1}^{2^{r-1}} T(x_i, y_i, z_i) \overline{T(x'_i, y_{i+1}, z_i)} \right] \right| \geq \varepsilon^{2^r}. \quad (3)$$

Let us establish the base case  $r = 1$ .

$$\begin{aligned} \varepsilon^2 &\leq \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)\tilde{H}(z)T(x,y,z)] \right| \\ &= \left| \mathbb{E}_{\substack{x_1, y_1, x'_1, y_2 \\ z_1}} [F(x_1) \overline{F(x'_1)} G(y_1) \overline{G(y_2)} T(x_1, y_1, z_1) \overline{T(x'_1, y_2, z_1)}] \right|. \end{aligned}$$

The inductive step will follow by squaring this, and using Cauchy-Schwarz on the variable other than  $y_{2^{r-1}+1}$ . Precisely,

$$\begin{aligned} \varepsilon^{2^{r+1}} &\leq \left| \mathbb{E}_{\substack{x_1, x'_1, \dots, x_{2^{r-1}}, x'_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}+1} \\ z_1, \dots, z_{2^{r-1}}} \left[ \prod_{i=1}^{2^{r-1}} F(x_i) \overline{F(x'_i)} \cdot G(y_1) \overline{G(y_{2^{r-1}+1})} \prod_{i=1}^{2^{r-1}} T(x_i, y_i, z_i) \overline{T(x'_i, y_{i+1}, z_i)} \right] \right|^2 \\ &\leq \mathbb{E}_{y_{2^{r-1}+1}} \left| \mathbb{E}_{\substack{x_1, x'_1, \dots, x_{2^{r-1}}, x'_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}} \\ z_1, \dots, z_{2^{r-1}}} \left[ \prod_{i=1}^{2^{r-1}} F(x_i) \overline{F(x'_i)} \cdot G(y_1) \prod_{i=1}^{2^{r-1}} T(x_i, y_i, z_i) \overline{T(x'_i, y_{i+1}, z_i)} \right] \right|^2. \end{aligned}$$

Note that this last expression exactly equals (3) for  $r + 1$ , completing the induction. Finally, the definition of  $\tilde{H}$  gives that the LHS expression in (3) can be expressed as:

$$\left| \mathbb{E}_{\substack{x_1, x'_1, \dots, x_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}} \\ z_1, \dots, z_{2^{r-1}}} \left[ \prod_{i=1}^{2^{r-1}} F(x_i) \prod_{i=1}^{2^{r-1}-1} \overline{F(x'_i)} \cdot G(y_1) \tilde{H}(z_{2^{r-1}}) \prod_{i=1}^{2^{r-1}} T(x_i, y_i, z_i) \prod_{i=1}^{2^{r-1}-1} \overline{T(x'_i, y_{i+1}, z_i)} \right] \right|.$$

So we can set  $F^+(\vec{x}) = \prod_{i=1}^{2^{r-1}} F(x_i) \prod_{i=1}^{2^{r-1}-1} \overline{F(x'_i)}$ . This is exactly what we wanted to show, once we note that only the  $T$  part of the previous expression depends on  $y_2, \dots, y_{2^{r-1}}, z_1, \dots, z_{2^{r-1}-1}$ .  $\square$

## 4.2 Merging Symbols

Let us go back to the setting of a distribution  $\mu$  over  $\Sigma \times \Gamma \times \Phi$  (you can think of it as if we have already applied a path trick to  $x$ ). Consider a graph  $G$  with vertices in  $\Sigma$ , where  $x, x' \in \Sigma$  are connected if and only if there are  $y \in \Gamma, z \in \Phi$  such that  $(x, y, z), (x', y, z) \in \text{supp}(\mu)$ . Give each connected component in  $G$  a single representative vertex, and let  $\text{rep}(x)$  denote the representative of the component of  $x \in \Sigma$ . Let  $\Sigma^- = \{\text{rep}(x) : x \in \Sigma\}$ . The first goal is to reduce to the case where  $(\text{rep}(x), y, z) \in \text{supp}(\mu)$ .

**Lemma 4.7** (Merging symbols). *If  $F : \Sigma^n \rightarrow \mathbb{D}$ ,  $G : \Gamma^n \rightarrow \mathbb{D}$ ,  $H : \Phi^n \rightarrow \mathbb{D}$  satisfy*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \geq \varepsilon,$$

*then there is a distribution  $\mu'$  over  $\Sigma \times \Gamma \times \Phi$  with  $\{(\text{rep}(x), y, z) : (x, y, z) \in \text{supp}(\mu)\} \subseteq \text{supp}(\mu')$ , nonembeddable  $\tilde{T} : \text{supp}(\mu') \rightarrow \mathbb{D}$ , and  $\tilde{F} : \Sigma^n \rightarrow \mathbb{D}$  with*

$$\left| \mathbb{E}_{(x,y,z) \sim (\mu')^{\otimes n}} [\tilde{F}(x)G(y)H(z) \prod_{i=1}^n \tilde{T}(x_i, y_i, z_i)] \right| \geq \varepsilon^{2M},$$

where  $M = |\Sigma|$ .

*Proof.* Define the matrix  $\mathbf{A} \in \mathbb{C}^{\Sigma \times (\Gamma \times \Phi)}$  by  $\mathbf{A}_{(x,(y,z))} = \mu(x, y, z)T(x, y, z)$ . Let  $\mathbf{D}_x = \text{diag}(\mu_x)$  and  $\mathbf{D}_{yz} = \text{diag}(\mu_{y,z})$ . Define the vector  $R(y, z) = G(y)H(z)$ .

In the calculations below, to simplify notation we by abuse of notation let  $\mathbf{A}, \mathbf{D}_x, \mathbf{D}_{yz}$  denote  $\mathbf{A}^{\otimes n}, \mathbf{D}_x^{\otimes n}, \mathbf{D}_{yz}^{\otimes n}$  respectively. By the hypothesis and Cauchy-Schwarz we get:

$$\varepsilon \leq |\overline{F}^* \mathbf{A} R| \leq \|\mathbf{D}_x^{1/2} F\|_2 \|\mathbf{D}_x^{-1/2} \mathbf{A} R\|_2 \leq \|\mathbf{D}_x^{-1/2} \mathbf{A} R\|_2,$$

so expanding gives

$$(\mathbf{D}_{yz}^{1/2} R)^* (\mathbf{D}_{yz}^{-1/2} \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A} \mathbf{D}_{yz}^{-1/2}) \mathbf{D}_{yz}^{1/2} R \geq \varepsilon^2.$$

Using  $\|\mathbf{D}_{yz}^{1/2} R\|_2 \leq 1$  and repeated Cauchy-Schwarz gives us

$$(\mathbf{D}_{yz}^{1/2} R)^* (\mathbf{D}_{yz}^{-1/2} \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A} \mathbf{D}_{yz}^{-1/2})^M \mathbf{D}_{yz}^{1/2} R \geq \varepsilon^{2M}.$$

Thus for  $\tilde{F} = \mathbf{D}_x^{-1} \mathbf{A} R$  we know that

$$\tilde{F}^* \mathbf{A} (\mathbf{D}_{yz}^{-1} \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A})^{M-1} R \geq \varepsilon^{2M}. \quad (4)$$

Now we interpret this to extract the desired  $\tilde{T}$  and  $\mu'$ . Consider a random walk between  $\Sigma$  and  $(\Gamma \times \Phi)$  as follows, where each step is sampled proportional to  $\mu$ . The total number of steps is  $2M-1$ , starting from  $(y_1, z_1) \sim \mu_{y,z}$ . It goes through vertices  $(y_1, z_1), x_1, (y_2, z_2), \dots, (y_M, z_M), x_M$ . Then we can set  $\mu'(x, y, z) = \Pr[x = x_M, y = y_1, z = z_1]$  and

$$\tilde{T}(x, y, z) = \mathbb{E} \left[ \prod_{i=1}^M T(x_i, y_i, z_i) \prod_{i=1}^{M-1} \overline{T(x_i, y_{i+1}, z_{i+1})} \mid x = x_M, y = y_1, z = z_1 \right].$$

Then (4) exactly tells us that  $\mu', \tilde{T}$  satisfy the conclusion of the lemma. To check that  $\tilde{T}$  is nonembeddable note that  $\tilde{T}(x, y, z) = T(x, y, z)$  for all  $(x, y, z) \in \text{supp}(\mu)$ , or  $|\tilde{T}(x, y, z)| < 1$ , where we are done by [Lemma 4.2](#).  $\square$

Now, by the random restriction trick ([Lemma 2.3](#)), we can ensure that  $\text{supp}(\mu^-)$  contains exactly  $(\text{rep}(x), y, z)$ .

**Lemma 4.8.** *Under the hypotheses of [Lemma 4.7](#), there is a distribution  $\mu^-$  with  $\text{supp}(\mu^-) = \{(\text{rep}(x), y, z) : (x, y, z) \in \text{supp}(\mu)\}$ ,  $n' \geq cn$ , and functions  $\tilde{F} : (\Sigma^-)^{n'} \rightarrow \mathbb{D}, \tilde{G} : \Gamma^{n'} \rightarrow \mathbb{D}, \tilde{H} : \Phi^{n'} \rightarrow \mathbb{D}$  such that*

$$\left| \mathbb{E}_{(x,y,z) \sim (\mu^-)^{\otimes n'}} [\tilde{F}(x) \tilde{G}(y) \tilde{H}(z) \prod_{i=1}^{n'} T(x_i, y_i, z_i)] \right| \geq \varepsilon^{2M}/2.$$

We have reduced the size of the alphabet and support of the distribution so we may worry that  $T : \text{supp}(\mu^-) \rightarrow \{-1, 1\}$  may become embeddable into  $\mathbb{D}$ . We show that if this is the case, then we can show that the game has exponentially decreasing value. This can be viewed as an instance of two player games.

**Lemma 4.9.** *If  $T : \text{supp}(\mu^-) \rightarrow \mathbb{D}$  is embeddable over  $\mathbb{D}$  then for some  $c := c(T, \mu) > 0$ ,*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} \left[ F(x) G(y) H(z) \prod_{i=1}^n T(x_i, y_i, z_i) \right] \right| \leq 2^{-cn}.$$



*Proof.* We first show that if  $T : \text{supp}(\mu^-) \rightarrow \mathbb{D}$  is embeddable, then there are no functions  $A : \Sigma \rightarrow \mathbb{D}$  and  $B : (\Gamma \times \Phi) \rightarrow \mathbb{D}$  for which  $A(x)B(y, z) = T(x, y, z)$  for all  $(x, y, z) \in \text{supp}(\mu)$ . Assume such  $A, B$  existed. We will show how to create an embedding on  $\text{supp}(\mu)$  from one on  $\text{supp}(\mu^-)$ . Take an embedding  $A'(x)B'(y)C'(z) = T(x, y, z)$  for all  $(x, y, z) \in \text{supp}(\mu^-)$ . Then we have

$$T(x, y, z) = A(x)B(y, z) = \frac{A(x)}{A(\text{rep}(x))} T(\text{rep}(x), y, z) = \frac{A(x)A'(\text{rep}(x))}{A(\text{rep}(x))} B'(y)C'(z),$$

hence  $T$  admits an embedding on  $\mu$ , a contradiction.

So we may assume that  $T : \Sigma \times (\Gamma \times \Phi) \rightarrow \mathbb{D}$  is nonembeddable. Critically, we have combined  $\Gamma, \Phi$  so that this is two dimensional now. Thus, the conclusion follows from the two-dimensional case, in [Lemma 2.2](#).  $\square$

### 4.3 The Relaxed Base Case

In this section, we reduce to  $\mu$  which satisfy a certain *relaxed base case*. This is to handle the *Horn SAT obstruction*, as introduced in [\[BKM23a\]](#).

**Definition 4.10** (Relaxed base case). We say that a distribution  $\mu$  on  $\Sigma \times \Gamma \times \Phi$  and  $\Sigma' \subseteq \Sigma$  satisfy the relaxed base case if:

- $\mu_{y,z}$  is a product distribution.
- $(y, z)$  uniquely determines  $x$  in  $\text{supp}(\mu)$ .
- For a nonembeddable tensor  $T$ , functions  $G : \Gamma \rightarrow \mathbb{C}$ ,  $H : \Phi \rightarrow \mathbb{C}$ , and  $F : \Sigma \rightarrow \mathbb{C}$  with

$$I_{\text{eff}}[F(x)] = \mathbb{E}_{x, x' \sim \mu_x} [1_{x, x' \in \Sigma'} |F(x) - F(x')|^2] \geq \tau \|F\|_2^2,$$

we have for a constant  $c := c(T, \mu)$  and  $M = \max\{|\Sigma|, |\Gamma|, |\Phi|\}$ ,

$$\left| \mathbb{E}_{(x, y, z) \sim \mu} [F(x)G(y)H(z)T(x, y, z)] \right| \leq (1 - c\tau^{100M}) \|F\|_2 \|G\|_2 \|H\|_2.$$

The following lemma gives a reduction to the relaxed base case.

**Lemma 4.11** (Reduction to relaxed base case). *If there are functions  $F : \Sigma^n \rightarrow \mathbb{D}$ ,  $G : \Gamma^n \rightarrow \mathbb{D}$ ,  $H : \Phi^n \rightarrow \mathbb{D}$  with*

$$\left| \mathbb{E}_{(x, y, z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \geq \varepsilon,$$

*then we can find  $\Sigma^+, \Gamma^+, \Phi$  and  $\mu^+$  on  $\Sigma^+ \times \Gamma^+ \times \Phi$  and  $\Sigma' \subseteq \Sigma^+$  satisfying the relaxed base case, and  $n' \geq cn$ , functions  $\tilde{F} : (\Sigma^+)^{n'} \rightarrow \mathbb{D}$ ,  $\tilde{G} : (\Gamma^+)^{n'} \rightarrow \mathbb{D}$ ,  $\tilde{H} : \Phi^{n'} \rightarrow \mathbb{D}$ , and nonembeddable  $\tilde{T} : \text{supp}(\mu^+) \rightarrow \mathbb{D}$  with*

$$\left| \mathbb{E}_{(x, y, z) \sim (\mu^+)^{\otimes n'}} [\tilde{F}(x)\tilde{G}(y)\tilde{H}(z) \prod_{i=1}^{n'} \tilde{T}(x_i, y_i, z_i)] \right| \geq \varepsilon^{O_M(1)}. \quad (5)$$

*Proof.* We construct  $\mu^+$  as follows starting from  $\mu$ .

1. Apply the path trick ([Lemma 4.6](#)) to  $y$ . The result is a distribution  $\mu^{(1)}$  on  $\Sigma \times \Gamma^+ \times \Phi$  such that  $\mu_{x,z}^{(1)}$  has full support.

2. Apply the path trick (Lemma 4.6) to  $x$ . The result is a distribution  $\mu^{(2)}$  on  $\Sigma^+ \times \Gamma^+ \times \Phi$  such that  $\mu_{y,z}^{(2)}$  is full. We also will require that for all  $z$  and  $\vec{x} = (x, \dots, x) \in \Sigma^+$ , there is a  $\vec{y} \in \Gamma^+$  with  $(\vec{x}, y, z) \in \text{supp}(\mu^{(2)})$ . This follows because if  $(x, \vec{y}, z) \in \text{supp}(\mu^{(1)})$ , then  $(\vec{x}, \vec{y}, z) \in \text{supp}(\mu^{(2)})$ .
3. Merge symbols in  $\Sigma^+$  using Lemmas 4.7 and 4.8. Relabel the alphabet as  $\Sigma^+$  still. Call this distribution  $\mu^{(3)}$ .
4. Apply a random restriction so that  $(y, z)$  is uniform. This is our distribution  $\mu'$ .
5. Set  $\Sigma' = \{\text{rep}((x, \dots, x)) : x \in \Sigma\}$ .

The existence of  $\tilde{F}, \tilde{G}, \tilde{H}$ , and nonembeddable  $\tilde{T}$  satisfying (5) follows from combining Lemmas 4.6 to 4.9. It suffices to show that  $\mu'$  satisfies the relaxed base case. The first two properties follow by construction. We proceed to showing the third property.

In fact we only use the weaker property that  $\mathbb{E}_{x \sim \mu_x} [1_{x \in \Sigma'} |F(x)|^2] \geq \tau \|F\|_2^2$ . Assume that  $\|F\|_2 = \|G\|_2 = \|H\|_2 = 1$ . For simplicity of notation, we use  $\lesssim, \gtrsim$  in this proof to suppress constants depending on  $M, \mu$ .

Let us multiply  $F(x)$  pointwise by the same unit complex number so that

$$\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)T(x,y,z)]$$

is a positive real number. We may assume this throughout. Then we have the equality

$$|\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)T(x,y,z)]| = 1 - \frac{1}{2} \mathbb{E}_{(x,y,z) \sim \mu} \left[ |\overline{F(x)} - G(y)H(z)T(x,y,z)|^2 \right].$$

Assume for contradiction that  $|\overline{F(x)} - G(y)H(z)T(x,y,z)|^2 \leq c^2 \tau^{100M}$  for all  $(x,y,z) \in \text{supp}(\mu)$ , for some constant  $c$ . In particular,  $\|F(x) - |G(y)||H(z)|\| \leq c\tau^{50M}$ .

Now, there is  $x \in \Sigma'$  with  $|F(x)|^2 \gtrsim \tau$ . For all  $z \in \Phi$  there is  $y \in \Gamma^+$  with  $(x,y,z) \in \text{supp}(\mu)$ . Using that  $|G(y)| \lesssim 1$  for all  $y$  (as  $\|G\|_2 = 1$ ), we deduce that

$$|H(z)| \geq \frac{|F(x)| - c\tau^{50M}}{|G(y)|} \gtrsim \sqrt{\tau}.$$

We next establish that  $|G(y)|, |F(x)| \gtrsim \tau^{M/2}$  for all  $x \in \Sigma^+, y \in \Gamma^+$ . Indeed, for any  $(x,y)$  such that there is a  $z \in \Phi$  with  $(x,y,z) \in \text{supp}(\mu)$ , then  $|F(x)| \geq |G(y)||H(z)| - c\tau^{50M} \gtrsim |G(y)|\sqrt{\tau} - c\tau^{50M}$ , and

$$|G(y)| \geq \frac{|F(x)| - c\tau^{50M}}{|H(z)|} \gtrsim |F(x)| - c\tau^{50M},$$

as  $|H(z)| \lesssim 1$ . This implies that  $|G(y)|, |F(x)| \gtrsim \tau^{M/2}$  for all  $x \in \Sigma^+, y \in \Gamma^+$  because the  $x, y$  coordinates are pairwise connected, so there is a path from the  $x \in \Sigma'$  with  $|F(x)| \gtrsim \tau^{1/2}$  to any  $x \in \Sigma^+$  of length at most  $M$ .

Now, because  $T$  is nonembeddable, there is  $(x,y,z) \in \text{supp}(\mu)$  with

$$\left| T(x,y,z) - \frac{\overline{F(x)}}{|F(x)|} \cdot \frac{\overline{G(y)}}{|G(y)|} \cdot \frac{\overline{H(z)}}{|H(z)|} \right| \gtrsim 1.$$

From this, it is direct that

$$|T(x,y,z)G(y)H(z) - \overline{F(x)}| \gtrsim \min\{|G(y)H(z)|, |F(x)|\} \gtrsim \tau^M.$$

□

## 4.4 Reducing to High-Degree Functions

In this section, we will deduce [Theorem 2](#) from the following bound on functions with high effective degrees, which we show in [Section 5](#).

**Theorem 4.** *If  $\mu$  satisfies the relaxed base case ([Definition 4.10](#)), then there are constants  $c := c(T, \mu), C := C(T, \mu) > 0$  such that for all  $G : \Gamma^n \rightarrow \mathbb{C}, H : \Phi^n \rightarrow \mathbb{C}$ , and  $F : \Sigma^n \rightarrow \mathbb{C}$  that is effectively homogeneous with effective degree  $d$ ,*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \leq 2^{-c \cdot d(d/n)^C} \|F\|_2 \|G\|_2 \|H\|_2.$$

To deduce [Theorem 2](#) from [Theorem 4](#), the idea is to decompose  $F = \mathcal{T}_{I-\rho}^{\text{eff}} F + (1 - \mathcal{T}_{I-\rho}^{\text{eff}})F$ , where  $I$  is the identity operator. We can handle the former term because it causes the distribution  $\mu$  to mix over  $\Sigma'$ , at which point the underlying game becomes essentially connected. The second term has only high-degree terms, and we apply [Theorem 4](#). To achieve the desired quantitative dependence on  $n$  in [Theorem 2](#), we will use a combination of noise operators instead of a single operator to more effectively eliminate low-degree terms.

Our combination of noise operators comes from the following lemma.

**Lemma 4.12** (Polynomial approximation). *For any  $d \in \mathbb{Z}_{\geq 0}$  and  $\varepsilon \in (0, 1/2)$ , there are  $\rho_0, \dots, \rho_d \in [0, 1 - \varepsilon]$  and  $c_0, \dots, c_d \in \mathbb{R}$  such that:*

1.  $\sum_{j=0}^d c_j \rho_j^k = 1$  for  $k = 0, 1, \dots, d$ ,
2.  $0 \leq \sum_{j=0}^d c_j \rho_j^k \leq 1$  for  $k \geq d + 1$ .
3.  $\sum_{j=0}^d |c_j| \leq 2^{O(d\sqrt{\varepsilon})}$ .

*Proof.* The proof uses Chebyshev polynomials and other tricks, and is deferred to [Appendix A](#).  $\square$

We also prove that applying a single noise operator  $\mathcal{T}_{1-\xi}^{\text{eff}}$  to  $F$  leads to a game whose value is exponentially decaying.

**Lemma 4.13.** *For  $F : \Sigma^n \rightarrow \mathbb{D}, G : \Gamma^n \rightarrow \mathbb{D}, H : \Phi^n \rightarrow \mathbb{D}$ ,  $\mu$  constructed as in [Lemma 4.11](#), and any  $\xi > 0$ , there is a constant  $c := c(T, \mu) > 0$  such that*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [\mathcal{T}_{1-\xi}^{\text{eff}} F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \leq 2^{-c\xi^n}.$$

*Proof.* Recall that  $(y, z)$  determines  $x$  in the support of  $\mu$ . Thus we can (by abuse of notation) write  $T(y, z) = T(x, y, z)$ . Let  $\mu'$  be the distribution defined as follows: sample  $(x, y, z) \sim \mu$ , then  $x' \sim N_{1-\xi}^{\text{eff}}(x)$ , and return  $(x', y, z)$ . Then by definition,

$$\begin{aligned} & \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [\mathcal{T}_{1-\xi}^{\text{eff}} F(x)G(y)H(z) \prod_{i=1}^n T(y_i, z_i)] \right| \\ &= \left| \mathbb{E}_{(x,y,z) \sim (\mu')^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(y_i, z_i)] \right|. \end{aligned}$$

For sufficiently small constant  $c := c(T, \mu)$ , we can write  $\mu' = c\xi\bar{\mu} + (1 - c\xi)\nu$  for the distribution  $\bar{\mu}$  which is sampled as follows: sampled  $(x, y, z) \sim \mu$  conditioned on  $x \in \Sigma'$ , and then resample  $x \in \Sigma'$

independently. Taking a random restriction (Lemma 2.3) gives that it suffices to show that for all  $F : \Sigma^{n'} \rightarrow \mathbb{D}, G : \Gamma^{n'} \rightarrow \mathbb{D}, H : \Phi^{n'} \rightarrow \mathbb{D}$ , we have

$$\mathbb{E}_{(x,y,z) \sim (\bar{\mu})^{\otimes n'}} [F(x)G(y)H(z) \prod_{i=1}^{n'} T(y_i, z_i)] \leq 2^{-cn'}.$$

Note that  $\bar{\mu}$  is connected, so it suffices to show that  $T$  is nonembeddable over  $\text{supp}(\bar{\mu})$ . This follows from the fact that Lemmas 4.7 and 4.9 only use values  $T(x, y, z)$  for  $x \in \Sigma'$  to show nonembeddability of  $T$ .  $\square$

We can now show our main analytic result, Theorem 2.

*Proof of Theorem 2.* By combining all our reductions in Lemmas 4.6 to 4.8 and 4.11 (which either increase  $\varepsilon$  by polynomial factors, or decrease  $n$  by a constant), we assume that  $\mu$  satisfies the relaxed base case. Let  $I$  be the identity operator on  $L_2(\Sigma^n, \mu_x^{\otimes n})$ . For parameters  $d \in \mathbb{Z}_{\geq 0}$  (chosen later) and  $\varepsilon = 1/4$ , let  $\{c_j\}, \{\rho_j\}$  be as in Lemma 4.12, and define  $\mathcal{T} = \sum_{j=0}^d c_j \mathcal{T}_{\rho_j}^{\text{eff}}$ . We now write

$$\begin{aligned} & \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \\ & \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [\mathcal{T}F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] + \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [(I - \mathcal{T})F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)]. \end{aligned}$$

We handle the first term using Lemma 4.13. Indeed,

$$\begin{aligned} & \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [\mathcal{T}F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \\ & \leq \sum_{j=0}^d |c_j| \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [\mathcal{T}_{\rho_j}^{\text{eff}} F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \leq 2^{O(d\sqrt{\varepsilon})} 2^{-c\varepsilon n}, \end{aligned}$$

by Lemma 4.13 and that  $\rho_j \leq 1 - \varepsilon$  for all  $j$  by Lemma 4.12. We use Theorem 4 to handle the second term. Indeed,

$$\begin{aligned} & \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [(I - \mathcal{T})F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \\ & = \left| \sum_{k=0}^n (1 - \sum_{j=0}^d c_j \rho_j^k) \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F^{\text{eff}=k}(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \\ & \leq \sum_{k=d+1}^n \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F^{\text{eff}=k}(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \leq \sum_{k=d+1}^n 2^{-ck(k/n)^C} \leq n \cdot 2^{-d(d/n)^C}, \end{aligned}$$

where the first inequality uses Items 1 and 2 of Lemma 4.12, and the second inequality uses Lemma 5.1 and  $\|F^{\text{eff}=k}\|_2 \leq \|F\|_2 \leq 1$  for all  $k$ . To conclude, set  $d = c'n$  for sufficiently small constant  $c'$ , so that the total is at most  $2^{O(d\sqrt{\varepsilon})} 2^{-c\varepsilon n} + n \cdot 2^{-d(d/n)^C} \leq 2^{-c'n}$ .  $\square$

## 5 Main Induction: No Complex Embeddings

The goal of this section is to show [Theorem 4](#) by induction on  $n$ .

**Theorem 4.** *If  $\mu$  satisfies the relaxed base case ([Definition 4.10](#)), then there are constants  $c := c(T, \mu), C := C(T, \mu) > 0$  such that for all  $G : \Gamma^n \rightarrow \mathbb{C}, H : \Phi^n \rightarrow \mathbb{C}$ , and  $F : \Sigma^n \rightarrow \mathbb{C}$  that is effectively homogeneous with effective degree  $d$ ,*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right| \leq 2^{-c \cdot d(d/n)^C} \|F\|_2 \|G\|_2 \|H\|_2.$$

Let  $\mathcal{P}^{(n,d)}$  be the class of all functions  $F : \Sigma^n \rightarrow \mathbb{C}$  such that all monomials in  $F$  have effective degree at least  $d$ . Define

$$\gamma_{n,d} := \sup_{\substack{\|F\|_2 \leq 1, F \in \mathcal{P}^{(n,d)} \\ \|G\|_2 \leq 1, \|H\|_2 \leq 1}} \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] \right|.$$

[Theorem 4](#) follows from the following lemma.

**Lemma 5.1.** *If  $\mu$  satisfies the relaxed base case, then  $\gamma_{n,d} \leq (1 - c(d/n)^{O_M(1)}) \gamma_{n-1,d-1}$ .*

*Proof of [Theorem 4](#).* Iterating [Lemma 5.1](#) a total of  $d/2$  times gives:

$$\gamma_{n,d} \leq (1 - c(d/n)^{O_M(1)})^{d/2} \leq \exp(-c \cdot d(d/n)^{O_M(1)}). \quad \square$$

The proof of [Lemma 5.1](#) proceeds by performing an SVD on the functions  $F, G, H$ .

**Lemma 5.2 (SVD).** *For a function  $F : \Sigma^n \rightarrow \mathbb{C}$  with  $\|F\|_2 = 1$ , for  $M = |\Sigma|$ , there are functions  $F_1, \dots, F_M : \Sigma^{n-1} \rightarrow \mathbb{C}, F'_1, \dots, F'_M : \Sigma \rightarrow \mathbb{C}$ , and nonnegative real numbers  $\kappa_1, \dots, \kappa_M$  satisfying:*

1.  $F_1, \dots, F_M$  form an orthonormal basis.
2.  $F'_1, \dots, F'_M$  form an orthonormal basis.
3.  $\sum_{r=1}^M \kappa_r^2 = 1$ .
4.  $F(x) = \sum_{r=1}^M \kappa_r F_r(x_1, \dots, x_{n-1}) F'_r(x_n)$  for all  $x \in \Sigma^n$ .

Note that each  $F_r \in \mathcal{P}^{(n-1,d-1)}$  because each  $F'_r$  has degree 1.

For notational simplicity, let  $I = \{1, \dots, n-1\}$  and  $J = \{n\}$ . We write  $x_I = (x_1, \dots, x_{n-1})$  and  $x_J = x_n$ . Towards showing [Lemma 5.1](#), because  $F$  has all monomials of effective degree at least  $d$ , we know that its effective influence satisfies  $I_{\text{eff}}[F] \geq d$ . Hence there is an index  $j \in [n]$  with  $I_{j, \text{eff}}[F] \geq d/n$ . Without loss of generality, we let  $j = n$ . Let  $\tau = d/n$ .

This actually implies that the average effective variance of the  $F'_r$  is large.

**Lemma 5.3.** *We have:  $\sum_{r \in [M]} \kappa_r^2 I_{\text{eff}}[F'_r] = I_{n, \text{eff}}[F] \geq \tau$ .*

*Proof.* Indeed, if  $(x_n, y_n)$  are sampled by the distribution used to define  $I_{\text{eff}}$ , we know

$$\begin{aligned}
I_{n,\text{eff}}[F] &= \mathbb{E}_{x_I, (x_n, y_n)} [|F(x_I, x_n) - F(x_I, y_n)|^2] = \mathbb{E}_{x_I, (x_n, y_n)} \left[ \left| \sum_{r \in [M]} \kappa_r F_r(x_I) (F'_r(x_n) - F'_r(y_n)) \right|^2 \right] \\
&= \mathbb{E}_{x_I, (x_n, y_n)} \left[ \sum_{r_1, r_2 \in [M]} \kappa_{r_1} \kappa_{r_2} F_{r_1}(x_I) \overline{F_{r_2}(x_I)} (F'_{r_1}(x_n) - F'_{r_1}(y_n)) \overline{(F'_{r_2}(x_n) - F'_{r_2}(y_n))} \right] \\
&= \mathbb{E}_{(x_n, y_n)} \left[ \sum_{r \in [M]} \kappa_r^2 |F'_r(x_n) - F'_r(y_n)|^2 \right] = \sum_{r \in [M]} \kappa_r^2 I_{\text{eff}}[F'_r]
\end{aligned}$$

□

Perform SVDs on  $F, G, H$  using [Lemma 5.2](#). We write

$$F(x) = \sum_{r=1}^M \kappa_r F_r(x_I) F'_r(x_J), G(y) = \sum_{s=1}^M \lambda_s G_s(y_I) G'_s(y_J), H(z) = \sum_{t=1}^M \mu_t H_t(x_I) H'_t(x_J).$$

Define the quantities:

$$\begin{aligned}
\widehat{F}_r(s, t) &:= \mathbb{E}_{(x, y, z) \sim \mu^{\otimes I}} [F_r(x) G_s(y) H_t(z) \prod_{i \in I} T(x_i, y_i, z_i)], \\
\widehat{F}'_r(s, t) &:= \mathbb{E}_{(x, y, z) \sim \mu} [F'_r(x) G'_s(y) H'_t(z) T(x, y, z)].
\end{aligned}$$

Then by definition we know that

$$\mathbb{E}_{(x, y, z) \sim \mu^{\otimes n}} [F(x) G(y) H(z) \prod_{i=1}^n T(x_i, y_i, z_i)] = \sum_{r, s, t \in [M]} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t).$$

## 5.1 Finding a Singular Value Gap

Let  $\delta < \Delta$  satisfy that none of the  $\kappa_r, \lambda_s, \mu_t$  lie in  $[\delta, \Delta]$ , and  $\delta \leq \left(\frac{\Delta}{CM\tau}\right)^{1000M}$ . We can find such  $\delta, \Delta$  with  $\delta \geq \tau^{O_M(1)}$ . Let  $R = \{r : \kappa_r \geq \Delta\}, S = \{s : \lambda_s \geq \Delta\}, T = \{t : \mu_t \geq \Delta\}$ . Note that because  $|\widehat{F}_r(s, t)| \leq \gamma_{n-1, d-1}, |\widehat{F}'_r(s, t)| \leq 1$  we know

$$\left| \sum_{r, s, t \in [M]} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t) \right| \leq \left| \sum_{r \in R, s \in S, t \in T} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t) \right| + \delta M^3 \gamma_{n-1, d-1}. \quad (6)$$

## 5.2 Cauchy-Schwarz and Simple Observations

For the remainder of the section, we assume that [Lemma 5.1](#) is false for the choice of  $F, G, H$ . By the Cauchy-Schwarz inequality we get:

$$\left| \sum_{r \in R, s \in S, t \in T} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t) \right| \leq \sqrt{\sum_{s \in S, t \in T} \lambda_s^2 \mu_t^2 \sum_{r \in R} |\widehat{F}_r(s, t)|^2} \sqrt{\sum_{r \in R} \kappa_r^2 \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2}. \quad (7)$$

We have the following simple observations:

**Lemma 5.4.** For all  $s, t$  we have:  $\sum_{r \in [M]} |\widehat{F}_r(s, t)|^2 \leq \gamma_{n-1, d-1}^2$ .

*Proof.* Write

$$\widetilde{F}_{s,t} = \frac{\sum_{r \in [M]} \overline{\widehat{F}_r(s, t)} F_r}{\sqrt{\sum_{r \in [M]} |\widehat{F}_r(s, t)|^2}}.$$

Clearly  $\|\widetilde{F}_{s,t}\|_2 = 1$ . Thus, we get

$$\sum_{r \in [M]} |\widehat{F}_r(s, t)|^2 = \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes I}} \left[ \widetilde{F}_{s,t}(x) G_r(y) H_s(y) \prod_{i \in I} T(x_i, y_i, z_i) \right] \right|^2 \leq \gamma_{n-1, d-1}^2.$$

□

Analogously, we have:

**Lemma 5.5.** For all  $s, t$  we have:  $\sum_{r \in [M]} |\widehat{F}'_r(s, t)|^2 \leq 1$ .

*Proof.* The proof is identical to that of [Lemma 5.4](#).

□

**Lemma 5.6.** For all  $r$  we have:  $\sum_{s,t \in [M]} |\widehat{F}'_r(s, t)|^2 \leq 1$ .

*Proof.* This is one of the few places in the proof we use that  $x$  is determined by  $(y, z)$  under the relaxed based case ([Definition 4.10](#)). To clarify this throughout this proof, we will write  $x$  as  $x(y, z)$ . Define the functions  $U_{s,t} : \Gamma \times \Phi \rightarrow \mathbb{C}$  as  $U_{s,t}(y, z) = G'_s(y) H'_t(z) T(x(y, z), y, z)$ . These functions are orthogonal over  $L_2(\Gamma \times \Phi, \mu_{y,z})$ . Indeed for  $(s_1, t_1)$  and  $(s_2, t_2)$ ,

$$\begin{aligned} \mathbb{E}_{(y,z) \sim \mu_{y,z}} [U_{s_1, t_1}(y, z) \overline{U_{s_2, t_2}(y, z)}] &= \mathbb{E}_{(y,z) \sim \mu_{y,z}} [G'_{s_1}(y) \overline{G'_{s_2}(y)} H'_{t_1}(z) \overline{H'_{t_2}(z)} |T(x(y, z), y, z)|^2] \\ &= \mathbb{E}_{(y,z) \sim \mu_{y,z}} [G'_{s_1}(y) \overline{G'_{s_2}(y)} H'_{t_1}(z) \overline{H'_{t_2}(z)}] = \mathbf{1}_{(s_1, t_1) = (s_2, t_2)}, \end{aligned}$$

where we have used that  $\mu_{y,z}$  is a product distribution,  $|T(x, y, z)| = 1$  for all  $(x, y, z) \in \text{supp}(\mu)$ , and orthogonality of the families  $G'_s$  and  $H'_t$ . The lemma follows because  $\widehat{F}'_r(s, t) = \langle F'_r, \overline{U_{s,t}} \rangle$ . □

Note now that these lemmas show:

$$\sum_{s \in S, t \in T} \lambda_s^2 \mu_t^2 \sum_{r \in R} |\widehat{F}_r(s, t)|^2 \leq \gamma_{n-1, d-1}^2 \quad \text{and} \quad \sum_{r \in R} \kappa_r^2 \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \leq 1.$$

Thus, if [Lemma 5.1](#) fails, we must have near-equality of [Lemmas 5.4](#) and [5.6](#).

**Lemma 5.7.** For all  $s \in S, t \in T$  we have:  $\sum_{r \in R} |\widehat{F}_r(s, t)|^2 \geq (1 - \frac{4\delta M^3}{\Delta^4}) \gamma_{n-1, d-1}^2$ .

*Proof.* If the contrary holds, then

$$\sum_{s \in S, t \in T} \lambda_s^2 \mu_t^2 \sum_{r \in R} |\widehat{F}_r(s, t)|^2 \leq (1 - 4\delta M^3) \gamma_{n-1, d-1}^2,$$

because  $\lambda_s, \mu_t \geq \Delta$ . Combining this with (6) and (7) would show [Lemma 5.1](#). □

**Lemma 5.8.** For all  $r \in R$  we have:  $\sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \geq 1 - \frac{4\delta M^3}{\Delta^2}$ .

*Proof.* If the contrary holds, then

$$\sum_{r \in R} \kappa_r^2 \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \leq 1 - 4\delta M^3,$$

because  $\kappa_r \geq \Delta$ . Combining this with (6) and (7) would show [Lemma 5.1](#).  $\square$

The remainder of the section is split into two cases depending on whether  $|R| < |S||T|$  or  $|R| \geq |S||T|$ . The former is handled by showing that we can perturb  $F_r, G_s, H_t$  to reach a contradiction. The latter is handled by using the one-dimensional bound guaranteed by the relaxed base case.

### 5.3 Handling the Case $|R| < |S||T|$

In this case we will show how to build functions  $\widetilde{F}, \widetilde{G}, \widetilde{H}$  that violate the inductive hypothesis. Define the  $|R|$  dimensional vectors  $V^{(s,t)}$ , defined as  $V_r^{(s,t)} = \widehat{F}'_r(s, t)$ . Because  $|R| < |S||T|$ , we can find some  $(s_1, t_1)$  and  $(s_2, t_2)$  such that  $|\langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle|$  is non-negligible.

**Lemma 5.9.** *For  $q > d$ , and vectors  $v_1, \dots, v_q \in \mathbb{C}^d$ , there is  $i \neq j$  such that*

$$|\langle v_i, v_j \rangle| \geq d^{-1} \|v_i\|_2 \|v_j\|_2.$$

*Proof.* By scaling we may assume that  $\|v_i\|_2 = 1$  for all  $i$ . Let  $\mathbf{V} \in \mathbb{C}^{q \times d}$  be the matrix whose  $i$ -th row is  $v_i$ . Then  $\mathbf{V}\mathbf{V}^*$  has entries  $\langle v_i, v_j \rangle$ , and has rank at most  $d$ . Hence

$$\sum_{i,j} |\langle v_i, v_j \rangle|^2 = \text{Tr}((\mathbf{V}\mathbf{V}^*)^2) \geq d \cdot (q/d)^2 = q^2/d.$$

Thus there is an  $i \neq j$  with

$$|\langle v_i, v_j \rangle|^2 \geq \frac{q^2/d - q}{q(q-1)} = \frac{q-d}{d(q-1)} \geq \frac{1}{d^2}$$

as long as  $q \geq d+1$ .  $\square$

Let  $(s_1, t_1) \neq (s_2, t_2)$  (we do allow for  $s_1 = s_2$  or  $t_1 = t_2$ ) satisfy

$$|\langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle| \geq |R|^{-1} \|V^{(s_1, t_1)}\|_2 \|V^{(s_2, t_2)}\|_2 \geq \frac{1}{2M} \gamma_{n-1, d-1}^2, \quad (8)$$

where we have used [Lemma 5.4](#).

Now we define the perturbed functions  $\widetilde{G}, \widetilde{H}$ . Let  $\alpha_s$  for  $s \in S$  and  $\beta_t$  for  $t \in T$  be unit complex variables. Let  $A_s$  for  $s \in S$  and  $B_t$  for  $t \in T$  be nonnegative real numbers satisfying  $\sum_{s \in S} A_s^2 = 1$  and  $\sum_{t \in T} B_t^2 = 1$ . Define

$$\widetilde{G} = \sum_{s \in S} \alpha_s A_s G_s \quad \text{and} \quad \widetilde{H} = \sum_{t \in T} \beta_t B_t H_t.$$

Note that  $\|\widetilde{G}\|_2 = \|\widetilde{H}\|_2 = 1$ . Define the coefficients

$$\widetilde{F}_r = \mathbb{E}_{(x,y,z) \sim \mu^I} [F_r(x) \widetilde{G}(y) \widetilde{H}(z) \prod_{i \in I} T(x_i, y_i, z_i)],$$



and the function  $\tilde{F} = \frac{\sum_{r \in R} \widetilde{F}_r F_r}{\sqrt{\sum_{r \in R} |\widetilde{F}_r|^2}}$ . Then  $\|\tilde{F}\|_2 = 1$  and thus

$$\sum_{r \in R} |\widetilde{F}_r|^2 = \left| \mathbb{E}_{(x,y,z) \sim \mu^I} [\tilde{F}(x) \tilde{G}(y) \tilde{H}(z) \prod_{i \in I} T(x_i, y_i, z_i)] \right|^2 \leq \gamma_{n-1, d-1}^2, \quad (9)$$

for any choice of  $\alpha_s, \beta_t, A_s, B_t$  by the inductive hypothesis. Expanding the leftmost term gives

$$\begin{aligned} \sum_{r \in R} |\widetilde{F}_r|^2 &= \sum_{r, s, t, s', t'} \alpha_s A_s \overline{\alpha_{s'}} A_{s'} \beta_t B_t \overline{\beta_{t'}} B_{t'} \widehat{F}_r(s, t) \overline{\widehat{F}_r(s', t')} \\ &= \sum_{s, t, s', t'} \alpha_s A_s \overline{\alpha_{s'}} A_{s'} \beta_t B_t \overline{\beta_{t'}} B_{t'} \langle V^{(s, t)}, V^{(s', t')} \rangle. \end{aligned}$$

Let us take the expectation of the previous expression over  $\alpha_s$  for  $s \notin \{s_1, s_2\}$  and  $\beta_t$  for  $t \notin \{t_1, t_2\}$  being independent, uniform unit complex variables. This gives

$$\begin{aligned} &\mathbb{E}_{\substack{\alpha_s, s \notin \{s_1, s_2\} \\ \beta_t, t \notin \{t_1, t_2\}}} \left[ \sum_{s, t, s', t'} \alpha_s \overline{\alpha_{s'}} \beta_t \overline{\beta_{t'}} \langle V^{(s, t)}, V^{(s', t')} \rangle \right] \\ &= \sum_{s \in S, t \in T} A_s^2 B_t^2 \|V^{(s, t)}\|_2^2 \end{aligned} \quad (10)$$

$$+ \sum_{\substack{s \in S \\ t, t' \in \{t_1, t_2\}, t \neq t'}} A_s^2 \beta_t B_t \overline{\beta_{t'}} B_{t'} \langle V^{(s, t)}, V^{(s, t')} \rangle \quad (11)$$

$$+ \sum_{\substack{s, s' \in \{s_1, s_2\}, s \neq s' \\ t \in T}} \alpha_s A_s \overline{\alpha_{s'}} A_{s'} B_t^2 \langle V^{(s, t)}, V^{(s', t)} \rangle \quad (12)$$

$$+ \sum_{\substack{s, s' \in \{s_1, s_2\}, s \neq s' \\ t, t' \in \{t_1, t_2\}, t \neq t'}} \alpha_s A_s \overline{\alpha_{s'}} A_{s'} \beta_t B_t \overline{\beta_{t'}} B_{t'} \langle V^{(s, t)}, V^{(s', t')} \rangle. \quad (13)$$

By [Lemma 5.7](#) we know that the expression in (10) is at least  $(1 - \frac{4\delta M^3}{\Delta^4}) \gamma_{n-1, d-1}^2$ . The rest of the argument is split into subcases depending on whether  $s_1 = s_2$ ,  $t_1 = t_2$ , or neither.

**Case 1: We have  $|\langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle| > \frac{4\delta M^3}{\Delta^4} \gamma_{n-1, d-1}^2$  for some  $s_1 = s_2$  or  $t_1 = t_2$ :** By symmetry, we only consider the case  $s_1 = s_2$ . Let  $s = s_1$ . In this case, the terms in (12) and (13) vanish. We set  $A_s = 1$  and all other  $A_{s'} = 0$ . set  $B_{t_1} = 1/\sqrt{2}$  and  $B_{t_2} = 1/\sqrt{2}$ . The the expression in (11) simplifies to

$$\frac{1}{2} \beta_{t_1} \overline{\beta_{t_2}} \langle V^{(s, t_1)}, V^{(s, t_2)} \rangle + \frac{1}{2} \beta_{t_2} \overline{\beta_{t_1}} \langle V^{(s, t_2)}, V^{(s, t_1)} \rangle = \text{Re} \left( \beta_{t_1} \overline{\beta_{t_2}} \langle V^{(s, t_1)}, V^{(s, t_2)} \rangle \right).$$

Thus, we should choose  $\beta_{t_1} \overline{\beta_{t_2}}$  to make the final expression a positive real number. Then its value is  $|\langle V^{(s, t_1)}, V^{(s, t_2)} \rangle| > \frac{4\delta M^3}{\Delta^4} \gamma_{n-1, d-1}^2$  by assumption. So the total sum is greater than

$$\left(1 - \frac{4\delta M^3}{\Delta^4}\right) \gamma_{n-1, d-1}^2 + \frac{4\delta M^3}{\Delta^4} \gamma_{n-1, d-1}^2 = \gamma_{n-1, d-1}^2,$$

a contradiction.

**Case 1 does not hold and**  $|\langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle| \geq \frac{1}{2M} \gamma_{n-1, d-1}^2$  **for some**  $s_1 \neq s_2$  **and**  $t_1 \neq t_2$ :  
Set  $A_{s_1} = A_{s_2} = B_{t_1} = B_{t_2} = 1/\sqrt{2}$ . We first bound the expressions in (11) and (12). Because **Case 1** does not hold, the expression in (11) can be lower bounded by

$$- \max_{s \in \{s_1, s_2\}} |\langle V^{(s, t_1)}, V^{(s, t_2)} \rangle| \geq -\frac{4\delta M^3}{\Delta^4} \gamma_{n-1, d-1}^2,$$

and an identical bound follows for (12). The expression in (11) can be simplified to

$$\begin{aligned} & \frac{1}{4} \left( \alpha_{s_1} \overline{\alpha_{s_2}} \beta_{t_1} \overline{\beta_{t_2}} \langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle + \alpha_{s_2} \overline{\alpha_{s_1}} \beta_{t_2} \overline{\beta_{t_1}} \langle V^{(s_2, t_2)}, V^{(s_1, t_1)} \rangle \right. \\ & \left. + \alpha_{s_2} \overline{\alpha_{s_1}} \beta_{t_1} \overline{\beta_{t_2}} \langle V^{(s_2, t_1)}, V^{(s_1, t_2)} \rangle + \alpha_{s_1} \overline{\alpha_{s_2}} \beta_{t_2} \overline{\beta_{t_1}} \langle V^{(s_1, t_2)}, V^{(s_2, t_1)} \rangle \right) \\ & = \frac{1}{2} \operatorname{Re} \left( \alpha_{s_1} \overline{\alpha_{s_2}} \beta_{t_1} \overline{\beta_{t_2}} \langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle + \alpha_{s_2} \overline{\alpha_{s_1}} \beta_{t_1} \overline{\beta_{t_2}} \langle V^{(s_2, t_1)}, V^{(s_1, t_2)} \rangle \right). \end{aligned} \quad (14)$$

Define  $\alpha = \alpha_{s_1} \overline{\alpha_{s_2}}$  and  $\beta = \beta_{t_1} \overline{\beta_{t_2}}$ . Pick  $\alpha, \beta$  uniformly at random so that  $\alpha\beta \langle V^{(s_1, t_1)}, V^{(s_2, t_2)} \rangle$  is a positive real number. Say  $\alpha\beta = \omega$ . Then  $\mathbb{E}[\alpha^{-1}\beta] = \mathbb{E}[\omega^{-1}\beta^2] = 0$ , so the second term in (14) vanishes. Thus in expectation, the total value of equations (10) to (13) is at least

$$\left(1 - \frac{4\delta M^3}{\Delta^4}\right) \gamma_{n-1, d-1}^2 - \frac{4\delta M^3}{\Delta^4} \gamma_{n-1, d-1}^2 - \frac{4\delta M^3}{\Delta^4} \gamma_{n-1, d-1}^2 + \frac{1}{2M} \gamma_{n-1, d-1}^2 > \gamma_{n-1, d-1}^2,$$

a contradiction.

#### 5.4 Handling the Case $|R| \geq |S||T|$

In this case we use the relaxed base case (Definition 4.10). We start with the following observation.

**Lemma 5.10.** *If  $|R| \geq |S||T|$  then for all  $s \in S, t \in T$  we have  $\sum_{r \in R} |\widehat{F}'_r(s, t)|^2 \geq 1 - \frac{4\delta M^5}{\Delta^2}$ .*

*Proof.* By Lemma 5.8 we know that

$$\sum_{r \in R, s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \geq |R| \left(1 - \frac{4\delta M^3}{\Delta^2}\right) \geq |S||T| \left(1 - \frac{4\delta M^3}{\Delta^2}\right).$$

Because  $\sum_{r \in R} |\widehat{F}'_r(s, t)|^2 \leq 1$  for all  $s \in S, t \in T$  by Lemma 5.5, the conclusion follows.  $\square$

For  $s \in S, t \in T$ , define the functions

$$\widetilde{F}_{st} = \frac{\sum_{r \in R} \overline{\widehat{F}'_r(s, t)} F'_r}{\sqrt{\sum_{r \in R} |\widehat{F}'_r(s, t)|^2}}.$$

Note that  $\|\widetilde{F}_{st}\|_2 = 1$  and

$$\left( \mathbb{E}_{(x, y, z) \sim \mu} \left[ \widetilde{F}_{st}(x) G'_s(y) H'_t(z) T(x, y, z) \right] \right)^2 = \sum_{r \in R} |\widehat{F}'_r(s, t)|^2 \geq 1 - \frac{4\delta M^5}{\Delta^2}, \quad (15)$$

by [Lemma 5.10](#). Now we leverage the relaxed base case property of  $\mu$  to get that

$$I_{\text{eff}}[\widetilde{F}_{st}] \leq \left( \frac{4\delta M^5}{c\Delta^2} \right)^{\frac{1}{100M}} := \delta',$$

for all  $s \in S$  and  $t \in T$ . The remainder of the proof converts this into an effective influence bound on the  $F'_r$  functions themselves to contradict [Lemma 5.3](#).

In the following, treat all functions to be on the space  $L_2(\Gamma \times \Phi, \mu_{y,z})$ , where we recall that  $(y, z)$  is a product distribution and determines  $x$  uniquely in the relaxed base case. Note that for all  $s \in S, t \in T$  that

$$\|\widetilde{F}_{st} - G'_s H'_t T\|_2 = \|\widetilde{F}_{st}\|_2^2 + \|G'_s H'_t T\|_2^2 - 2\mathbb{E}_{(x,y,z)} \left[ \widetilde{F}_{st}(x) G'_s(y) H'_t(z) T(x, y, z) \right] \leq \frac{8\delta M^5}{\Delta^2}, \quad (16)$$

where we have used (15) and that the expectation is real. For all  $r \in R$  we have

$$\left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) \widetilde{F}_{st}} \right\|_2 \leq \left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) G'_s H'_t T} \right\|_2 + \sum_{s \in S, t \in T} \widehat{F}'_r(s, t) \left\| G'_s H'_t T - \widetilde{F}_{st} \right\|_2.$$

By (16) the second term can be bounded by

$$\sum_{s \in S, t \in T} \widehat{F}'_r(s, t) \left\| G'_s H'_t T - \widetilde{F}_{st} \right\|_2 \leq M^2 \cdot \frac{8\delta M^5}{\Delta^2} = \frac{8\delta M^7}{\Delta^2}.$$

The first term can be simplified to get

$$\left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) G'_s H'_t T} \right\|_2 = \left( \|\overline{F'_r}\|_2^2 - \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \right)^{1/2} \leq \sqrt{\frac{4\delta M^3}{\Delta^2}},$$

where we have used that the functions  $G'_s H'_t T$  are orthonormal (see the proof of [Lemma 5.6](#)), and [Lemma 5.8](#). By the triangle inequality again, we get

$$\begin{aligned} I_{\text{eff}}[F'_r] &\leq I_{\text{eff}} \left[ \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) \widetilde{F}_{st}} \right] + \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)| I_{\text{eff}}[\widetilde{F}_{st}] \\ &\leq \left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) \widetilde{F}_{st}} \right\|_2 + \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)| I_{\text{eff}}[\widetilde{F}_{st}] \\ &\leq \sqrt{\frac{4\delta M^3}{\Delta^2}} + \frac{8\delta M^7}{\Delta^2} + M^2 \delta' \leq 2M^2 \delta', \end{aligned}$$

by our choice of  $\delta = \left( \frac{\Delta}{cM\tau} \right)^{1000M}$ . This gives that

$$\sum_{r \in [M]} \kappa_r^2 I_{\text{eff}}[F'_r] \leq M\delta^2 + \sum_{r \in R} \kappa_r^2 I_{\text{eff}}[F'_r] \leq M\delta^2 + 2M^2 \delta' < \tau,$$

by the choice of  $\delta$  again, because  $\delta' = \left( \frac{4\delta M^5}{c\Delta^2} \right)^{\frac{1}{100M}}$ . This contradicts [Lemma 5.3](#).

## 6 Complex Embeddings: Reductions

The system  $f(x) + g(y) + h(z) = t(x, y, z) \pmod{m}$  has a solution for all  $(x, y, z) \in \text{supp}(\mu)$  iff it has a solution for all prime power factors  $p^k$  of  $m$  by the Chinese remainder theorem. So we assume  $m = p^k$  for the remainder of the paper. If  $T$  is nonembeddable over  $\mathbb{D}$ , then we are done by [Corollary 4.1](#). So we assume that  $T$  is embeddable over  $\mathbb{D}$ , and hence there is an integer  $N$  and functions  $a : \Sigma \rightarrow \mathbb{Z}_{p^k N}, b : \Gamma \rightarrow \mathbb{Z}_{p^k N}, c : \Phi \rightarrow \mathbb{Z}_{p^k N}$  with

$$a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{p^k N}.$$

We assume that  $N$  is minimal. Thus the equation  $f(x) + g(y) + h(z) = t(x, y, z) \pmod{p^k}$  can be rewritten as  $\tilde{f}(x) + \tilde{g}(y) + \tilde{h}(z) = 0 \pmod{p^k N}$  where  $\tilde{f}(x) = N \cdot f(x) - a(x)$ .

This motivates us to apply the techniques of [\[BKM24\]](#) to reduce to the case where the distribution  $\mu$  is over all  $(x, y, z) \in \mathcal{A}^3$  with  $x + y + z = 0$ , where  $\mathcal{A}$  is a finite abelian group (the *master group*) that governs all abelian embeddings of  $\mu$ .

### 6.1 Master Group and Embedding

We start by introducing the master embedding, which essentially captures all the abelian embeddings of  $\mu$ . The master embedding maps into a finite abelian group as long as  $\mu$  has no  $\mathbb{Z}$ -embeddings.

**Definition 6.1** (Master Embedding). Let  $\mu$  be a distribution on  $\Sigma \times \Gamma \times \Phi$ , and let  $M = \max\{|\Sigma|, |\Gamma|, |\Phi|\}$ . Let  $r = O_M(1)$  be sufficiently large. Let  $(\sigma_1, \gamma_1, \phi_1), \dots, (\sigma_s, \gamma_s, \phi_s)$  be all embeddings of  $\mu$  into cyclic groups  $\mathcal{A}_1, \dots, \mathcal{A}_s$  of size at most  $r$ . Then a master embedding of  $\mu$  into  $\prod_{i \in [s]} \mathcal{A}_i$  is given by

$$\sigma_{\mathbf{m}}(x) = (\sigma_1(x), \dots, \sigma_s(x)), \quad \gamma_{\mathbf{m}}(y) = (\gamma_1(y), \dots, \gamma_s(y)), \quad \phi_{\mathbf{m}}(z) = (\phi_1(z), \dots, \phi_s(z)).$$

As argued in [\[BKM24\]](#), for sufficiently large  $r = O_M(1)$ , the master embedding captures all embeddings that are not equal up to simple transformations.

**Lemma 6.2.** *If  $\mu$  has no  $\mathbb{Z}$ -embeddings, then  $\sigma(x) = \sigma(x')$  for all embeddings  $\sigma, \gamma, \phi$  of  $\mu$  if and only if  $\sigma_{\mathbf{m}}(x) = \sigma_{\mathbf{m}}(x')$ , and similar for the  $y, z$  variables.*

We also do a reduction to the embedding functions  $a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{p^k N}$  to ensure that  $(a, b, c)$  are constant on the master embedding. Specifically, define  $\tilde{a}(x) = a(x) \pmod{N}$ , and similar for  $\tilde{b}, \tilde{c}$ , and let  $\tilde{t}(x, y, z) = (\tilde{a}(x) + \tilde{b}(y) + \tilde{c}(z))/N$ . We claim that the game with target values  $\tilde{t}$  is equivalent to the game with target value  $t$ . Indeed, note that if  $f(x) + g(y) + h(z) = t(x, y, z)$  then for

$$\tilde{f}(x) = f(x) + \frac{\tilde{a}(x) - a(x)}{N}, \quad \tilde{g}(y) = g(y) + \frac{\tilde{b}(y) - b(y)}{N}, \quad \tilde{h}(z) = h(z) + \frac{\tilde{c}(z) - c(z)}{N}$$

we have  $\tilde{f}(x) + \tilde{g}(y) + \tilde{h}(z) = \tilde{t}(x, y, z) \pmod{p^k}$ .

Thus, from now on, we make the following assumption.

**Assumption 6.3.** *We write  $N \cdot t(x, y, z) = a(x) + b(y) + c(z) \pmod{p^k N}$  where  $0 \leq a(x), b(y), c(z) < N$  for all  $x, y, z$ .*

This way, we know that  $a, b, c$  is an embedding, hence it is captured by the master embedding.

**Lemma 6.4.** *Under [Assumption 6.3](#),  $(a, b, c)$  is an embedding. Thus, if  $\sigma_{\mathbf{m}}(x) = \sigma_{\mathbf{m}}(x')$ , we have  $a(x) = a(x')$ .*

## 6.2 Modest Noise Operator and Modest Influences

The goal of this section is to define the *modest noise operator* and *modest influences*.

Let  $\mu$  be a distribution on  $\Sigma \times \Gamma \times \Phi$ , and consider a set  $\Sigma' \subseteq \Sigma$ . Let  $\mathcal{A}$  be the master group, and for each  $a \in \mathcal{A}$ , let  $\Sigma'_a = \{x \in \Sigma' : \sigma_m(x) = a\}$ . Create an orthonormal basis of  $L^2(\Sigma, \mu_x)$  called  $B = B' \cup \cup_{a \in \mathcal{A}} B_a$  as follows.  $B_a$  is an orthonormal basis for functions supported on  $\Sigma'_a$  that are orthogonal to constant.  $B'$  is an orthonormal basis for functions that are constant on each  $\Sigma'_a$ .

Any function  $F \in L_2(\Sigma^n, \mu_x^{\otimes n})$  can be written as  $F = \sum_{\chi \in B^{\otimes n}} \widehat{F}(\chi) \chi$ . For a character  $\chi$ , define its *modest degree* as  $\text{moddeg}(\chi) = \{i : \chi_i \in \cup_{a \in \mathcal{A}} B_a\}$ . This allows us to define the modest degree decomposition

$$F^{\text{modest}=i} = \sum_{\chi \in B^{\otimes n} : \text{moddeg}(\chi)=i} \widehat{F}(\chi) \chi.$$

Now define the modest influence as

$$I_{\text{modest}}[F] = \sum_{i \in [n]} I_{i, \text{modest}}[F] = \mathbb{E}_{(x,y) \in \Sigma^n} [|F(x) - F(y)|^2],$$

where  $(x, y)$  are sampled as follows. Sample  $x \in \Sigma^n$  according to  $\mu_x^{\otimes n}$ . If  $x_i \notin \Sigma'$  set  $y_i = x_i$ . Otherwise, set  $y_i$  to be from  $\mu_x$  conditioned on  $y_i \in \Sigma'_{\sigma_m(x_i)}$ .

Analogous to the case of standard influences we know that

$$I_{i, \text{modest}}[F] = \sum_{\chi \in B^{\otimes n}, \chi_i \in B \setminus B'} |\widehat{F}(\chi)|^2 \quad \text{and} \quad I_{\text{modest}}[F] = \sum_{\chi \in B^{\otimes n}} \text{moddeg}(\chi) |\widehat{F}(\chi)|^2.$$

We finally define the modest noise operator  $\mathcal{T}_{1-\rho}^{\text{modest}}$ . Define  $\mathcal{T}_{1-\rho}^{\text{modest}} F(x) = \mathbb{E}_{y \sim N_{1-\rho}^{\text{modest}}(x)} [F(y)]$ , where  $N_{1-\rho}^{\text{modest}}(x)$  is the following distribution. For each  $i \in [n]$ , with probability  $1 - \rho$ , set  $y_i = x_i$ . Otherwise with probability  $\rho$ , do the following. If  $x_i \in \Sigma \setminus \Sigma'$ , set  $y_i = x_i$ . Otherwise, sample  $y_i$  from  $\mu_x$  conditioned on  $y_i \in B'_{\sigma_m(x_i)}$ . Analogous to the standard noise operator, we have

$$\mathcal{T}_{1-\rho}^{\text{modest}} F = \sum_{0 \leq i \leq n} (1 - \rho)^i F^{\text{modest}=i}.$$

## 6.3 Saturating the Master Embedding via Path Tricks

The goal of this section is to perform multiple path tricks as in [Section 4.1](#). This is with two goals in mind. The first is to enlarge the sets  $\{\sigma_m(x) : x \in \Sigma\}$ ,  $\{\gamma_m(y) : y \in \Gamma\}$ ,  $\{\phi_m(z) : z \in \Phi\}$  until each of them forms a group. In fact, it is easy to argue that the three groups must be identical.

The secondary goal will be to work towards establishing a variation of the relaxed base case.

Throughout this section,  $\mu, t$  will denote a distribution and target value, and  $f : \Sigma^n \rightarrow \mathbb{Z}_m^n, g : \Gamma^n \rightarrow \mathbb{Z}_m^n, h : \Phi^n \rightarrow \mathbb{Z}_m^n$  will be strategies. Let the winning probability of  $f, g, h$  be represented as

$$\text{val}_{\mu, t}(f, g, h) = \Pr_{(x,y,z) \sim \mu^{\otimes n}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \pmod{m} \quad \forall i = 1, 2, \dots, n].$$

**Lemma 6.5.** *Let  $\mu, t$  represent a game of value less than 1, and let  $\text{val}_{\mu, t}(f, g, h) \geq \varepsilon$ . If  $\mu^+$  represents the result of applying a path trick to  $\mu$  with  $r$  steps ([Definition 4.3](#)), then there is  $t^+$  such that  $\mu^+, t^+$  has value less than 1, and there are  $\tilde{f}, \tilde{g}, \tilde{h}$  with  $\text{val}_{\mu^+, t^+}(\tilde{f}, \tilde{g}, \tilde{h}) \geq \varepsilon^{2^r}$ .*

*Proof.* Without loss of generality, let  $\mu^+$  be the result of an  $x$  path trick. In this case, we set

$$t^+(\vec{x}, y_1, z_{2^{r-1}-1}) = \sum_{i=1}^{2^{r-1}} t(x_i, y_i, z_i) - \sum_{i=1}^{2^{r-1}-1} t(x'_i, y_{i+1}, z_i).$$

This definition actually does not depend on  $y_2, \dots, y_{2^{r-1}}, z_1, \dots, z_{2^{r-1}-2}$ . Indeed, because  $a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{p^k N}$  we have

$$N \cdot t^+(\vec{x}, y_1, z_{2^{r-1}-1}) = \sum_{i=1}^{2^{r-1}} a(x_i) - \sum_{i=1}^{2^{r-1}-1} a(x'_i) + b(y_1) + c(z_{2^{r-1}-1}) \pmod{p^k N},$$

so  $t^+$  is uniquely determined modulo  $p^k$ .

Now,  $\mu^+, t^+$  has value less than 1 because for all  $(x, y, z) \in \text{supp}(\mu)$ , we have  $t^+((x, \dots, x), y, z) = t(x, y, z)$ . We will show the following by induction.

$$\Pr_{\substack{x_1, x'_1, \dots, x_{2^{r-1}}, x'_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}+1} \\ z_1, \dots, z_{2^{r-1}}}} \left[ \sum_{i=1}^{2^{r-1}} f(x_i) - f(x'_i) + g(y_1) - g(y_{2^{r-1}+1}) = \sum_{i=1}^{2^{r-1}} t(x_i, y_i, z_i) - t(x'_i, y_{i+1}, z_i) \right] \geq \varepsilon^{2^r}. \quad (17)$$

Let us establish the case  $r = 1$ . Because  $\text{val}_{\mu, t}(f, g, h) \geq \varepsilon$ , we get

$$\begin{aligned} \varepsilon^2 &\leq \mathbb{E}_z \left[ \Pr_{(x_1, y_1, z_1) \sim \mu} [f(x_1) + g(y_1) - t(x_1, y_1, z_1) = h(z_1) | z_1 = z] \right]^2 \\ &\leq \mathbb{E}_{z'} \left[ \Pr_{(x_1, y_1, z_1) \sim \mu} [f(x_1) + g(y_1) - t(x_1, y_1, z_1) = h(z_1) | z_1 = z] \right]^2 \\ &= \mathbb{E}_z \Pr_{\substack{(x_1, y_1, z_1) \sim \mu \\ (x'_1, y_2, z_1) \sim \mu}} [f(x_1) + g(y_1) - t(x_1, y_1, z_1) = f(x_2) + g(y_2) - t(x_2, y_2, z_1) = h(z_1) | z_1 = z] \\ &\leq \mathbb{E}_z \Pr_{\substack{(x_1, y_1, z_1) \sim \mu \\ (x'_1, y_2, z_1) \sim \mu}} [f(x_1) - f(x_2) + g(y_1) - g(y_2) = t(x_1, y_1, z_1) - t(x'_1, y_2, z_1) | z_1 = z] \quad (18) \end{aligned}$$

$$= \Pr_{\substack{x_1, x'_1 \\ y_1, y_2 \\ z_1}} [f(x_1) - f(x'_1) + g(y_1) - g(y_2) = t(x_1, y_1, z_1) - t(x'_1, y_2, z_1)]. \quad (19)$$

We complete the induction with a similar approach. By induction

$$\begin{aligned} \varepsilon^{2^{r+1}} &\leq \mathbb{E}_{y_{2^{r-1}+1}} \left[ \Pr_{\substack{x_1, x'_1, \dots, x_{2^{r-1}}, x'_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}} \\ z_1, \dots, z_{2^{r-1}}}} \left[ \sum_{i=1}^{2^{r-1}} f(x_i) - f(x'_i) + g(y_1) - g(y_{2^{r-1}+1}) = \sum_{i=1}^{2^{r-1}} t(x_i, y_i, z_i) - t(x'_i, y_{i+1}, z_i) \right] \right]^2 \\ &\leq \mathbb{E}_{y_{2^{r-1}+1}} \left[ \Pr_{\substack{x_1, x'_1, \dots, x_{2^{r-1}}, x'_{2^{r-1}} \\ y_1, \dots, y_{2^{r-1}} \\ z_1, \dots, z_{2^{r-1}}}} \left[ \sum_{i=1}^{2^{r-1}} f(x_i) - f(x'_i) + g(y_1) - g(y_{2^{r-1}+1}) = \sum_{i=1}^{2^{r-1}} t(x_i, y_i, z_i) - t(x'_i, y_{i+1}, z_i) \right] \right]^2. \end{aligned}$$

Expanding the last expression and removing the  $g(y_{2^{r-1}+1})$  like in (18), (19) completes the induction. To conclude, we discuss the choices of  $\tilde{f}, \tilde{g}, \tilde{h}$ . We set  $\tilde{f}(\vec{x}) = \sum_{i=1}^{2^{r-1}} f(x_i) - \sum_{i=1}^{2^{r-1}-1} f(x'_i)$  and

$\tilde{g}(y) = y$ . Finally, for each  $z \in \Phi$  we randomly choose  $\tilde{h}(z)$  as follows: generate  $(x, y, z')$  from  $\mu$  conditioned on  $z' = z$ , and set  $\tilde{h}(z) = t(x, y, z') - f(x) - g(y)$ . Then (17) exactly tells us that  $\mathbb{E}_{\tilde{h}}[\text{val}_{\mu^+, t^+}(\tilde{f}, \tilde{g}, \tilde{h})] \geq \varepsilon^{2^r}$  as desired.  $\square$

We study how embeddings evolve under the path trick.

**Lemma 6.6.** *Let  $\mu^+$  be the result of applying an  $x$  path trick to  $\mu$ . If  $\sigma^+, \gamma^+, \phi^+$  is an embedding of  $\mu^+$ , then there must be an embedding  $\sigma, \gamma, \phi$  of  $\mu$  such that  $\gamma^+(y) = b(y)$  for all  $y \in \Gamma$ ,  $\phi^+(z) = \phi(z)$  for all  $z \in \Phi$ , and  $\sigma^+(\vec{x}) = \sum_{i=1}^{2^{r-1}} \sigma(x_i) - \sum_{i=1}^{2^{r-1}-1} \sigma(x'_i)$ .*

*Proof.* Recall that  $((x, \dots, x), y, z) \in \text{supp}(\mu^+)$ . Thus, there are embeddings  $\sigma, \gamma, \phi$  of  $\mu$  with  $\gamma^+(y) = \gamma(y)$  and  $\phi^+(z) = \phi(z)$ . If  $(\vec{x}, y_1, z_{2^{r-1}-1}) \in \text{supp}(\mu^+)$ , then

$$\sigma^+(\vec{x}) = -\gamma(y_1) - \phi(z_{2^{r-1}-1}) = \sum_{i=1}^{2^{r-1}} \sigma(x_i) - \sum_{i=1}^{2^{r-1}-1} \sigma(x'_i).$$

$\square$

Finally, we discuss how to augment the image of the master embedding.

**Lemma 6.7.** *Consider the following sequence of steps to a distribution  $\mu$ .*

1. *Perform a  $z$  path trick with  $r > \log_2 M$ .*
2. *Perform a  $y$  path trick with  $r > \log_2 M$ .*
3. *Perform a  $x$  path trick with  $r = 2$ .*

*Then if  $\mu$  had master embeddings  $\sigma_m, \gamma_m, \phi_m$ , then the new distribution  $\mu^+$  has master embeddings  $\sigma_m^+, \gamma_m^+, \phi_m^+$  generated according to Lemma 6.6. In particular,  $\text{Image}(\gamma_m) \subseteq \text{Image}(\gamma_m^+)$ ,  $\text{Image}(\phi_m) \subseteq \text{Image}(\phi_m^+)$ , and for all  $x_1, x_2, x_3 \in \Sigma$  we have  $\sigma_m(x_1) - \sigma_m(x_2) + \sigma_m(x_3) \in \text{Image}(\sigma_m^+)$ .*

*Proof.* Everything before the final sentence follows from Lemma 6.6. Let  $\mu', \mu''$  after step 1, 2 respectively. In  $\mu'$ ,  $(x, y)$  has full support. In  $\mu''$ ,  $(x, z)$  has full support, and for all  $x \in \Sigma, y \in \Gamma$  there is some  $z$  with  $(x, (y, \dots, y), z) \in \text{supp}(\mu'')$ . Now, we find a path containing any  $x_1, x_2, x_3 \in \Sigma$ . Pick a  $y \in \Gamma$ . Let  $z_1, z_2$  be such that  $(x_1, (y, \dots, y), z_1) \in \text{supp}(\mu'')$  and  $(x_2, (y, \dots, y), z_2) \in \text{supp}(\mu'')$ . Because  $(x, z)$  has full support in  $\mu''$ , there is  $\vec{y}$  with  $(x_3, \vec{y}, z_2) \in \text{supp}(\mu'')$ , as desired.  $\square$

By shifting, assume that  $0 \in \text{Image}(\sigma_m), \text{Image}(\gamma_m), \text{Image}(\phi_m)$ .

**Lemma 6.8** (Saturated embedding). *Let  $\mu, t$  represent a game of value less than 1, with strategies  $f, g, h$  with  $\text{val}_{\mu, t}(f, g, h) \geq \varepsilon$ . Then there is a game  $\mu^+, t^+$  of value less than 1, and strategies  $\tilde{f}, \tilde{g}, \tilde{h}$  with  $\text{val}_{\mu^+, t^+}(\tilde{f}, \tilde{g}, \tilde{h}) \geq \varepsilon^{O_M(1)}$ . Additionally:*

1. *The master embeddings  $\sigma_m, \gamma_m, \phi_m$  of  $\mu^+$  satisfy for some finite group  $\mathcal{A}$ :*

$$\text{Image}(\sigma_m) = \text{Image}(\gamma_m) = \text{Image}(\phi_m) = \mathcal{A}.$$

2.  *$(y, z)$  has full support in  $\mu^+$ .*

*Proof.* Repeating Lemma 6.7 multiple times, and its symmetric  $y, z$  versions, we eventually reach a distribution  $\mu^+$  where  $\text{Image}(\sigma_m), \text{Image}(\gamma_m), \text{Image}(\phi_m)$  are subgroups. These must be the same subgroup because (say)  $(x, y)$  has full support. Then for any  $a \in \text{Image}(\sigma_m), b \in \text{Image}(\gamma_m)$ , we know that  $-a - b \in \text{Image}(\phi_m)$ . Let  $\mathcal{A} = \text{Image}(\sigma_m)$ . Item 2 can be guaranteed by applying one more  $x$  path trick. The bound  $\text{val}_{\mu^+, t^+}(\tilde{f}, \tilde{g}, \tilde{h}) \geq \varepsilon^{O_M(1)}$  follows since we apply Lemma 6.5 a finite number of times.  $\square$

## 6.4 Merging Symbols

We apply symbol merging in this section as in [Section 4.2](#).

**Lemma 6.9.** *Let  $(\mu, t)$  be a game of value less than 1, with strategies  $f, g, h$  with  $\text{val}_{\mu, t}(f, g, h) \geq \varepsilon$ . Then there is a distribution  $\mu'$  with  $\{(\text{rep}(x), y, z) : (x, y, z) \in \text{supp}(\mu)\} \subseteq \text{supp}(\mu')$ ,  $t'$  on  $\text{supp}(\mu')$  with game value less than 1, and strategy  $\tilde{f}$  such that  $\text{val}_{\mu', t'}(\tilde{f}, g, h) \geq \varepsilon^{2M}$ .*

*Proof.* To avoid repeating arguments, we will show this by combining [Lemmas 2.1](#) and [4.7](#). We use the same notation as the proof of [Lemma 4.7](#). In particular,  $\mu'$  is still given by the random walk between  $\Sigma$  and  $(\Gamma \times \Phi)$ , and we let  $N \cdot t'(x, y, z) = a(x) + b(y) + c(z) \pmod{p^k N}$ . Clearly,  $(\mu', t')$  has game value less than 1 because  $t'(x, y, z) = t(x, y, z)$  for  $(x, y, z) \in \text{supp}(\mu)$ . Finally, we set  $\tilde{f}(x)$  as follows: pick  $(y, z)$  as a random neighbor of  $x$  proportional to mass in  $\mu$ , and set  $\tilde{f}(x) = t(x, y, z) - g(y) - h(z)$ . Define  $\tilde{F}_S, G_S, H_S$  as in [Lemma 2.1](#), and  $R_S(y, z) = G_S(y)H_S(z)$ .

$$\begin{aligned} \mathbb{E}_{\tilde{f}}[\text{val}_{\mu', t'}(\tilde{f}, g, h)] &= \mathbb{E}_{S \in \mathbb{Z}_m^n} \mathbb{E}_{\tilde{f}}[\tilde{F}_S^* \mathbf{A} (\mathbf{D}_{yz}^{-1} \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A})^{M-1} R_S] \\ &= \mathbb{E}_{S \in \mathbb{Z}_m^n} \left[ (\mathbf{D}_{yz}^{1/2} R_S)^* (\mathbf{D}_{yz}^{-1/2} \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A} \mathbf{D}_{yz}^{-1/2})^M \mathbf{D}_{yz}^{1/2} R_S \right] \\ &\geq \mathbb{E}_{S \in \mathbb{Z}_m^n} \left[ (R_S^* \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A} R_S)^M \right] \geq \mathbb{E}_{S \in \mathbb{Z}_m^n} \left[ R_S^* \mathbf{A}^* \mathbf{D}_x^{-1} \mathbf{A} R_S \right]^M \\ &\geq \mathbb{E}_{S \in \mathbb{Z}_m^n} \left[ |\overline{F}_S^* \mathbf{A} R_S|^2 \right]^M \geq \mathbb{E}_{S \in \mathbb{Z}_m^n} \left[ \overline{F}_S^* \mathbf{A} R_S \right]^{2M} \\ &= \text{val}_{\mu, t}(f, g, h)^{2M} \geq \varepsilon^{2M}, \end{aligned}$$

where we have applied Jensen's inequality and [Lemma 2.1](#) several times.  $\square$

Now we show that we can reduce to when the support is exactly  $(\text{rep}(x), y, z)$ . This is essentially a repeat of the arguments in [Lemmas 4.8](#) and [4.9](#).

**Lemma 6.10.** *Let  $(\mu, t)$  be a game of value less than 1, with strategies  $f, g, h$  with  $\text{val}_{\mu, t}(f, g, h) \geq \varepsilon$ . Then there is a distribution  $\mu^-$  with support  $\{(\text{rep}(x), y, z) : (x, y, z) \in \text{supp}(\mu)\}$ ,  $(y, z)$  is a product distribution,  $t^-$  on  $\text{supp}(\mu^-)$  with game value less than 1, and strategies  $\tilde{f}, \tilde{g}, \tilde{h}$  on  $n' \geq cn$  variables with  $\text{val}_{\mu^-, t^-}(\tilde{f}, \tilde{g}, \tilde{h}) \geq \varepsilon^{2M}/2$ .*

*Proof.* We may assume that  $(\mu, t)$  are as constructed in [Lemma 6.9](#). Then  $\mu^-, t^-, \tilde{f}, \tilde{g}, \tilde{h}$  are formed by taking a random restriction. It suffices to show that  $(\mu^-, t^-)$  has game value less than 1. We show that otherwise, the two player game on  $\Sigma \times (\Gamma \times \Phi)$  with target value  $t(x, y, z)$  was unwinnable, so we can apply a two-player parallel repetition theorem to reach a contradiction [[Raz98](#)]. Indeed, assume there are strategies  $r(x) + s(y, z) = t(x, y, z) \pmod{m}$ , and  $f(x) + g(y) + h(z) = t^-(x, y, z)$  for  $(x, y, z) \in \text{supp}(\mu^-)$ . Then clearly

$$\begin{aligned} t(x, y, z) &= r(x) + s(y, z) = (r(x) - r(\text{rep}(x)) + r(\text{rep}(x)) + s(y, z)) \\ &= (r(x) - r(\text{rep}(x)) + t^-(x, y, z) = [f(x) + r(x) - r(\text{rep}(x))] + g(y) + h(z), \end{aligned}$$

so  $t$  admits a strategy with value 1, a contradiction.  $\square$

## 6.5 Establishing the Master Relaxed Base Case

The goal of this section is to establish an analogue of the relaxed base case in this setting.

**Definition 6.11** (Master relaxed base case). We say that a distribution  $\mu$  on  $\Sigma \times \Gamma \times \Phi$  and  $\Sigma' \subseteq \Sigma$  satisfies the *master relaxed base case* if:



- $\mu_{y,z}$  is a product distribution.
- $(y, z)$  uniquely determines  $x$  in  $\text{supp}(\mu)$ .
- For functions  $G : \Gamma \rightarrow \mathbb{C}$ ,  $H : \Phi \rightarrow \mathbb{C}$ , and  $F : \Sigma \rightarrow \mathbb{C}$  with

$$I_{\text{modest}}[F] = \mathbb{E}_{x,x' \sim \mu_x} [1_{x,x' \in \Sigma'} 1_{\sigma_m(x) = \sigma_m(x')} |F(x) - F(x')|^2] \geq \tau \|F\|_2^2,$$

we have for a constant  $c := c(\mu)$  and  $M = \max\{|\Sigma|, |\Gamma|, |\Phi|\}$  that

$$\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)] \leq (1 - c\tau^{100M}) \|F\|_2 \|G\|_2 \|H\|_2.$$

As in [Lemma 4.11](#), we show that applying two path tricks, merging, and taking a random restriction results in a distribution  $\mu$  that satisfies the master relaxed base case. In the next section, it will be important to remember that we got to a distribution satisfying the relaxed base case in this manner, because we require that the original distribution is contained inside it.

**Lemma 6.12.** *Consider a distribution  $\mu$  on  $\Sigma \times \Gamma \times \Phi$ . Perform the following sequence of operations to  $\mu$ .*

1. Apply the path trick ([Lemma 6.5](#)) to  $\mu$  on coordinate  $y$  to reach a distribution  $\mu^{(1)}$  on  $\Sigma \times \Gamma^+ \times \Phi$ .
2. Apply the path trick ([Lemma 6.5](#)) to  $\mu^{(1)}$  on coordinate  $x$  to reach a distribution  $\mu^{(2)}$  on  $\Sigma^+ \times \Gamma^+ \times \Phi$ .
3. Merge symbols in  $\Sigma^+$  using ([Lemma 6.10](#)) to reach a distribution  $\mu^{(3)}$ .
4. Random restrict  $\mu^{(3)}$  so that  $(y, z)$  is uniform. Call this distribution  $\mu^+$ .
5. Set  $\Sigma' = \{\text{rep}((x, \dots, x)) : x \in \Sigma\}$ .

Then  $\mu^+$  satisfies the master relaxed base case with the set  $\Sigma'$ .

*Proof.* The first two items of [Definition 6.11](#) are satisfied by construction. We focus on the last item.

Relabel  $\Sigma^+, \Gamma^+$  as  $\Sigma, \Gamma$  for simplicity. Start with the observation that  $\mathbb{E}_{x \sim \mu_x} [1_{x \in \Sigma'} |F(x)|^2] \geq \tau \|F\|_2^2$ . Assume that  $\|F\|_2 = \|G\|_2 = \|H\|_2 = 1$ . For simplicity of notation, we use  $\lesssim, \gtrsim$  in this proof to suppress constants depending on  $M, \mu$ .

Let us multiply  $F(x)$  pointwise by the same unit complex number so that

$$\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)]$$

is a positive real number. We may assume this throughout. Then we have the equality

$$|\mathbb{E}_{(x,y,z) \sim \mu} [F(x)G(y)H(z)]| = 1 - \frac{1}{2} \mathbb{E}_{(x,y,z) \sim \mu} \left[ |\overline{F(x)} - G(y)H(z)|^2 \right].$$

Assume for contradiction that  $|\overline{F(x)} - G(y)H(z)|^2 \leq c\tau^{100M}$  for all  $(x, y, z) \in \text{supp}(\mu)$ . In particular,  $\|F(x) - |G(y)||H(z)\| \leq c\tau^{50M}$ .

As in the proof of [Lemma 4.11](#), we can establish that  $|F(x)|, |G(y)|, |H(z)| \gtrsim \tau^{M/2}$  for all  $x \in \Sigma, y \in \Gamma, z \in \Phi$ . Now, we actually establish that  $|F(x)|$  must be nearly constant by using that  $\mu$  has no  $\mathbb{Z}$ -embeddings. Recall that  $\|F(x) - |G(y)||H(z)\| \lesssim \tau^{50M}$ . Using that  $|F(x)| \gtrsim \tau^{M/2}$

gives us  $|1 - |F(x)|^{-1}|G(y)||H(z)|| \lesssim \tau^{49M}$ . Define  $f'(x) = -\log |F(x)|$ ,  $g'(y) = \log |G(y)|$ ,  $h'(z) = \log |H(z)|$ . Then we get

$$|f'(x) + g'(y) + h'(z)| \lesssim \tau^{49M} \quad \text{for all } (x, y, z) \in \text{supp}(\mu),$$

i.e.,  $f', g', h'$  are approximately a  $\mathbb{Z}$ -embedding. By the Dirichlet approximation theorem, for  $N = (C\tau)^{9M}$ , there is some  $q \leq N$  and integers  $\sigma(x), \gamma(y), \phi(z)$  with

$$\left| f'(x) - \frac{\sigma(x)}{q} \right| \leq \frac{1}{qN^{1/(3M)}}, \quad \left| g'(y) - \frac{\gamma(y)}{q} \right| \leq \frac{1}{qN^{1/(3M)}}, \quad \left| h'(z) - \frac{\phi(z)}{q} \right| \leq \frac{1}{qN^{1/(3M)}}.$$

Thus

$$|\sigma(x) + \gamma(y) + \phi(z)| \leq q|f'(x) + g'(y) + h'(z)| + \frac{3}{N^{1/(3M)}} \lesssim N\tau^{49M} + 3(C\tau)^{-3} < 1.$$

Hence  $\sigma, \gamma, \phi$  form a  $\mathbb{Z}$ -embedding, so they are all constant. Thus

$$|\log |F(x)| - \log |F'(x)|| \leq \frac{1}{N^{1/(3M)}} = (C\tau)^{-3},$$

so  $|1 - |F(x)|/|F(x')|| \lesssim (C\tau)^{-3}$  for all  $x, x'$ .

To conclude, we will establish that the arguments of  $F(x)$  are nearly constant. Recall that  $|\overline{F(x)} - G(y)H(z)| \lesssim \tau^{50M}$  and  $|F(x)|, |G(y)|, |H(z)| \gtrsim \tau^{M/2}$ , so

$$\tau^{49M} \gtrsim \left| \frac{\overline{F(x)}}{|F(x)|} - \frac{G(y)}{|G(y)|} \frac{H(z)}{|H(z)|} \right| = \left| 1 - \frac{F(x)}{|F(x)|} \frac{G(y)}{|G(y)|} \frac{H(z)}{|H(z)|} \right|.$$

Let  $f'(x) = \arg(F(x))/(2\pi)$ ,  $g'(y) = \arg(G(y))/(2\pi)$ ,  $h'(z) = \arg(H(z))/(2\pi)$ . A direct calculation gives that  $\|f'(x) + g'(y) + h'(z)\|_{\mathbb{R}/\mathbb{Z}} \lesssim \tau^{49M}$ , i.e.  $f', g', h'$  nearly form an embedding. Applying Dirichlet approximation again gives

$$\left| f'(x) - \frac{\sigma(x)}{q} \right| \leq \frac{1}{qN^{1/(3M)}}, \quad \left| g'(y) - \frac{\gamma(y)}{q} \right| \leq \frac{1}{qN^{1/(3M)}}, \quad \left| h'(z) - \frac{\phi(z)}{q} \right| \leq \frac{1}{qN^{1/(3M)}}.$$

This gives that

$$\left\| \frac{1}{q}(\sigma(x) + \gamma(y) + \phi(z)) \right\|_{\mathbb{R}/\mathbb{Z}} \leq \|f'(x) + g'(y) + h'(z)\|_{\mathbb{R}/\mathbb{Z}} + \frac{3}{qN^{1/(3M)}} < 1/q.$$

Thus,  $\sigma(x) + \gamma(y) + \phi(z) \equiv 0 \pmod{q}$ , i.e. forms an embedding. Thus, for any  $x, x'$  with  $\sigma_m(x) = \sigma_m(x')$  we have  $\sigma(x) = \sigma(x')$  by the definition of the master embedding. Thus, for any  $x, x'$  with  $\sigma_m(x) = \sigma_m(x')$  we have that  $|f'(x) - f'(x')| \leq (C\tau)^{-3}$ . Combining this with  $|1 - |F(x)|/|F(x')|| \lesssim (C\tau)^{-3}$  for all  $x, x'$  gives  $|F(x) - F(x')| \lesssim \tau^{-3}$  for any  $\sigma_m(x) = \sigma_m(x')$ , contradicting the hypothesis.  $\square$

## 6.6 Applying the Modest Noise Operator

In [Section 7](#) we will show the following bound on functions with high modest degree.

**Theorem 5.** *If  $\mu$  satisfies the master relaxed base case ([Definition 6.11](#)), then for any  $G : \Gamma^n \rightarrow \mathbb{C}$ ,  $H : \Phi^n \rightarrow \mathbb{C}$ , and  $F : \Sigma^n \rightarrow \mathbb{C}$  that with all terms having modest degree at least  $d$ :*

$$|\mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z)]| \leq 2^{-c \cdot d(d/n)^C} \|F\|_2 \|G\|_2 \|H\|_2.$$

We will apply this bound to reduce to studying functions that are constant on the master embedding. The idea is to apply the modest noise operator, apply [Theorem 5](#) to remove the high-degree terms, and then take a random restriction onto a distribution that distributes mass equally over inputs with the same master embedding.

**Lemma 6.13.** *Let  $\mu$  be a saturated distribution on  $\Sigma \times \Gamma \times \Phi$ , and let  $\nu$  be a distribution on  $\mathcal{A} \times \Gamma \times \Phi$  sampled as follows: sample  $(y, z) \sim \mu_{y,z}$ , and return  $(-\sigma_{\mathbf{m}}(y) - \sigma_{\mathbf{m}}(z), y, z)$ .*

*For any constant  $c' > 0$ , for all sufficiently small  $c$  (in terms of  $c'$ ), and  $\text{val}_{t,\mu}(f, g, h) \geq \varepsilon$  for  $\varepsilon \geq 2^{-cn}$ , there are strategies  $\tilde{f}, \tilde{g}, \tilde{h}$  on  $n/8$  coordinates with  $\text{val}_{\nu,t}(\tilde{f}, \tilde{g}, \tilde{h}) \gtrsim \varepsilon^{O_M(1)}/2^{c'n}$ .*

*Proof.* Perform path tricks and merging as in [Lemma 6.12](#) to transform  $\mu$  into  $\mu^+$  which satisfies the master relaxed base case. For simplicity, relabel  $f, g, h$  as the resulting strategies. Note that this reduces  $n$  by a constant, due to random restrictions, and  $\text{val}_{\mu^+,t}(f, g, h) \geq \varepsilon'$  for  $\varepsilon' = \varepsilon^{O_M(1)}$ .

Let  $d \in \mathbb{Z}_{>0}$  be chosen later,  $\hat{\varepsilon} = 1/4$ , and  $\{c_i\}, \{\rho_i\}$  as in [Lemma 4.12](#) for  $\varepsilon := \hat{\varepsilon}$ . Let  $\mathcal{T} = \sum_{j=0}^d c_j \mathcal{T}_{\rho_j}^{\text{modest}}$ . Let  $F_S, G_S, H_S$  be as defined in [Lemma 2.1](#). Let  $a, b, c$  and  $N$  satisfy  $a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{p^k N}$ , and define

$$A(x) := \exp\left(\frac{2\pi i \cdot a(x)}{p^k N}\right), \quad B(y) := \exp\left(\frac{2\pi i \cdot b(y)}{p^k N}\right), \quad \text{and} \quad C(z) := \exp\left(\frac{2\pi i \cdot c(z)}{p^k N}\right).$$

For  $S \subseteq \mathbb{Z}_m^n$ , define  $A_S(x) = \prod_{i \in [n]} A(x)^{S_i}$ , so that  $T_S(x, y, z) = A_S(x)B_S(y)C_S(z)$  for all  $(x, y, z) \in \text{supp}(\mu)$ . By [Lemma 2.1](#),

$$\begin{aligned} \text{val}_{\mu^+,t}(f, g, h) &= \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [F_S(x)G_S(y)H_S(z)T_S(x, y, z)] \\ &= \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [(I - \mathcal{T})(A_S F_S)(x)(B_S G_S)(y)(C_S H_S)(z)] \\ &\quad + \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [\mathcal{T}(A_S F_S)(x)(B_S G_S)(y)(C_S H_S)(z)]. \end{aligned}$$

We will bound the first term using [Theorem 5](#). Indeed,

$$\begin{aligned} &\left| \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [(I - \mathcal{T})(A_S F_S)(x)(B_S G_S)(y)(C_S H_S)(z)] \right| \\ &= \left| \sum_{k=0}^n (1 - \sum_{j=0}^d c_j \rho_j^k) \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [(A_S F_S)^{\text{modest}=k}(x)(B_S G_S)(y)(C_S H_S)(z)] \right| \\ &\leq \sum_{k=d+1}^n 2^{-ck(k/n)^C} \leq n \cdot 2^{-cd(d/n)^C}, \end{aligned}$$

where we have used [Lemma 4.12](#) and [Lemma 7.1](#). For the choice  $d = c'n$  for  $c' \asymp c^{1/(C+1)}$ , we know that  $2^{-\Omega(d(d/n)^C)} \leq \varepsilon'/2$ , because  $\varepsilon' \geq 2^{-cn}$ . Thus [Lemma 2.1](#) and triangle inequality give

$$\begin{aligned} &\sum_{j=0}^d |c_j| \left| \mathbb{E}_{S \sim \mathbb{Z}_m^n} \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [\mathcal{T}_{\rho_j}^{\text{modest}}(A_S F_S)(x)(B_S G_S)(y)(C_S H_S)(z)] \right| \\ &\geq \text{val}_{\mu^+,t}(f, g, h) - \varepsilon'/2 \geq \varepsilon'/2. \end{aligned}$$

Because  $\sum_{j=1}^k |c_j| \leq 2^{O(d)}$  by [Lemma 4.12](#), there is some  $j$  with

$$\left| \mathbb{E}_{S \sim \mathbb{Z}_m^n} \mathbb{E}_{(x,y,z) \sim (\mu^+)^{\otimes n}} [\mathcal{T}_{\rho_j}^{\text{modest}}(A_S F_S)(x)(B_S G_S)(y)(C_S H_S)(z)] \right| \geq \varepsilon'/2^{O(d)} \gtrsim \varepsilon^{O_M(1)}/2^{O(c'n)}.$$

Define the distribution  $\mu'$  as follows: take a sample  $(x, y, z) \sim \mu^+$ , and then resample  $x' \sim N_{1-\rho_j}^{\text{modest}}(x)$ . Then the above expression equals

$$\left| \mathbb{E}_{S \sim \mathbb{Z}_m^n} \mathbb{E}_{(x,y,z) \sim (\mu')^{\otimes n}} [(A_S F_S)(x)(B_S G_S)(y)(C_S H_S)(z)] \right| = \text{val}_{t, \mu'}(f, g, h),$$

by Lemma 2.1 again. Thus  $\text{val}_{t, \mu'}(f, g, h) \gtrsim \varepsilon^{O_M(1)} / 2^{O(c'n)}$ . Formally, we need to define the tensor  $t$  on the support of  $\mu'$ . However, by Lemma 6.4 we know that for any  $x$ ,  $\sigma_m$  is constant on the support of  $N_{1-\rho_j}^{\text{modest}}(x)$ , so  $t$  just inherits its value.

To conclude the proof, we will take a random restriction of  $\mu'$ . Note that all atoms have in  $\mu'$  value at least  $1 - \rho_j \geq \hat{\varepsilon}$  for all  $j = 0, 1, \dots, d$ , by Lemma 4.12. By construction,  $\mu^+$  contained all points of the form  $(\text{rep}((x, \dots, x)), (y, \dots, y), z)$  for  $(x, y, z) \in \text{supp}(\mu)$ . Thus we can take a random restriction onto a distribution  $\nu$  supported on  $(x', y, z)$  for all  $(y, z) \in \text{supp}(\mu)$  and  $\sigma_m(x') = \sigma_m(\text{rep}(x)) = -\sigma_m(y) - \sigma_m(z)$ . The random restriction reduces  $n$  to at least  $n' = \hat{\varepsilon}n/2 = n/8$  coordinates, and only affects the winning probability by an exponentially small amount (see Lemma 2.3).  $\square$

Apply Lemma 6.13 to reduce to a distribution  $\mu$  on  $\mathcal{A} \times \Gamma \times \Phi$ . Because  $(y, z)$  is already a product distribution in the master relaxed base case, we can take another random restriction so make  $\gamma_m(y), \phi_m(z)$  uniform over  $\mathcal{A}$ . Thus, we have reduced to the distribution on  $\mathcal{A}^3$  which is uniform over all  $(x, y, z) \in \mathcal{A}^3$  with  $x + y + z = 0$ . Theorem 3, whose proof is based on additive combinatorics (as in [BKM23c]), shows that in this case,  $\varepsilon \leq 2^{-cn}$ . This completes the proof of Theorem 1, after we show Theorem 5 and Theorem 3.

## 7 Main Induction: Complex Embeddings

This section is dedicated to proving Theorem 5. The proof is nearly identical to that of Theorem 4.

**Theorem 5.** *If  $\mu$  satisfies the master relaxed base case (Definition 6.11), then for any  $G : \Gamma^n \rightarrow \mathbb{C}$ ,  $H : \Phi^n \rightarrow \mathbb{C}$ , and  $F : \Sigma^n \rightarrow \mathbb{C}$  that with all terms having modest degree at least  $d$ :*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z)] \right| \leq 2^{-c \cdot d(d/n)^C} \|F\|_2 \|G\|_2 \|H\|_2.$$

Define  $\mathcal{Q}^{(n,d)}$  to be the set of functions  $F : \Sigma^n \rightarrow \mathbb{C}$  all of whose terms have modest degree at least  $d$ . Define

$$\gamma_{n,d} := \sup_{\substack{\|F\|_2 \leq 1, F \in \mathcal{Q}^{(n,d)} \\ \|G\|_2 \leq 1, \|H\|_2 \leq 1}} \left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [F(x)G(y)H(z)] \right|.$$

Then Theorem 5 follows from the following lemma.

**Lemma 7.1.** *If  $\mu$  satisfies the master relaxed base case, then  $\gamma_{n,d} \leq (1 - c(d/n)^{O_M(1)}) \gamma_{n-1, d-1}$ .*

*Proof of Theorem 5.* Apply Lemma 7.1 a total of  $d/2$  times, giving

$$\gamma_{n,d} \leq (1 - c(d/n)^{O_M(1)})^{d/2} \leq \exp(-c \cdot d(d/n)^{O_M(1)}). \quad \square$$

The remainder of the section is devoted to establishing Lemma 7.1. Because  $F \in \mathcal{Q}^{(n,d)}$  we know that  $I_{\text{modest}}[F] \geq d$ , so  $\tau := I_{j, \text{modest}}[F] \geq d/n$  for some  $j$ . Without loss of generality, let  $j = n$ , and perform an SVD (Lemma 5.2) to write  $F(x) = \sum_{r=1}^M \kappa_r F_r(x_I) F_r'(x_n)$ .

Analogous to Lemma 5.3, the sum of modest influences of the  $F_r'$  are large.

**Lemma 7.2.** *We have:  $\sum_{r \in [M]} \kappa_r^2 I_{\text{modest}}[F'_r] = I_{n, \text{modest}}[F] \geq \tau$ .*

*Proof.* Indeed, if  $(x_n, y_n)$  are sampled by the distribution used to define  $I_{\text{modest}}$ , we know

$$\begin{aligned}
I_{n, \text{modest}}[F] &= \mathbb{E}_{x_I, (x_n, y_n)} [|F(x_I, x_n) - F(x_I, y_n)|^2] \\
&= \mathbb{E}_{x_I, (x_n, y_n)} \left[ \left| \sum_{r \in [M]} \kappa_r F_r(x_I) (F'_r(x_n) - F'_r(y_n)) \right|^2 \right] \\
&= \mathbb{E}_{x_I, (x_n, y_n)} \left[ \sum_{r_1, r_2 \in [M]} \kappa_{r_1} \kappa_{r_2} F_{r_1}(x_I) \overline{F_{r_2}(x_I)} (F'_{r_1}(x_n) - F'_{r_1}(y_n)) \overline{(F'_{r_2}(x_n) - F'_{r_2}(y_n))} \right] \\
&= \mathbb{E}_{(x_n, y_n)} \left[ \sum_{r \in [M]} \kappa_r^2 |F'_r(x_n) - F'_r(y_n)|^2 \right] = \sum_{r \in [M]} \kappa_r^2 I_{\text{modest}}[F'_r]
\end{aligned}$$

□

Given the SVDs

$$F(x) = \sum_{r \in [M]} \kappa_r F_r(x_I) F'_r(x_J), \quad G(y) = \sum_{s \in [M]} \lambda_s G_s(x_I) G'_s(x_J), \quad H(z) = \sum_{t \in [M]} \mu_t H_t(z_I) H'_t(z_J),$$

define the quantities

$$\begin{aligned}
\widehat{F}_r(s, t) &:= \mathbb{E}_{(x, y, z) \sim \mu \otimes I} [F_r(x) G_s(y) H_t(z)] \\
\widehat{F}'_r(s, t) &:= \mathbb{E}_{(x, y, z) \sim \mu} [F'_r(x) G'_s(y) H'_t(z)].
\end{aligned}$$

Then by definition we have that

$$\mathbb{E}_{(x, y, z) \sim \mu \otimes n} [F(x) G(y) H(z)] = \sum_{r, s, t \in [M]} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t).$$

## 7.1 Finding a Singular Value Gap

Let  $\delta < \Delta$  satisfy that none of the  $\kappa_r, \lambda_s, \mu_t$  lie in  $[\delta, \Delta]$ , and  $\delta \leq \left(\frac{\Delta}{CM\tau}\right)^{1000M}$ . We can find such  $\delta, \Delta$  with  $\delta \geq \tau^{O_M(1)}$ . Let  $R = \{r : \kappa_r \geq \Delta\}, S = \{s : \lambda_s \geq \Delta\}, T = \{t : \mu_t \geq \Delta\}$ . Note that because  $|\widehat{F}_r(s, t)| \leq \gamma_{n-1, d-1}, |\widehat{F}'_r(s, t)| \leq 1$  we know

$$\left| \sum_{r, s, t \in [M]} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t) \right| \leq \left| \sum_{r \in R, s \in S, t \in T} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t) \right| + \delta M^3 \gamma_{n-1, d-1}. \quad (20)$$

## 7.2 Cauchy-Schwarz and Simple Observations

For the remainder of the section, we assume that [Lemma 7.1](#) is false for the choice of  $F, G, H$ . By the Cauchy-Schwarz inequality we get:

$$\left| \sum_{r \in R, s \in S, t \in T} \kappa_r \lambda_s \mu_t \widehat{F}_r(s, t) \widehat{F}'_r(s, t) \right| \leq \sqrt{\sum_{s \in S, t \in T} \lambda_s^2 \mu_t^2 \sum_{r \in R} |\widehat{F}_r(s, t)|^2} \sqrt{\sum_{r \in R} \kappa_r^2 \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2}. \quad (21)$$

The following simple observations are analogous to their counterparts [Lemmas 5.4 to 5.8](#) in [Section 5](#).

**Lemma 7.3.** For all  $s, t$  we have:  $\sum_{r \in [M]} |\widehat{F}_r(s, t)|^2 \leq \gamma_{n-1, d-1}^2$ .

**Lemma 7.4.** For all  $s, t$  we have:  $\sum_{r \in [M]} |\widehat{F}'_r(s, t)|^2 \leq 1$ .

**Lemma 7.5.** For all  $r$  we have:  $\sum_{s, t \in [M]} |\widehat{F}'_r(s, t)|^2 \leq 1$ .

**Lemma 7.6.** For all  $s \in S, t \in T$  we have:  $\sum_{r \in R} |\widehat{F}_r(s, t)|^2 \geq (1 - \frac{4\delta M^3}{\Delta^4}) \gamma_{n-1, d-1}^2$ .

**Lemma 7.7.** For all  $r \in R$  we have:  $\sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \geq 1 - \frac{4\delta M^3}{\Delta^2}$ .

The case  $|R| < |S||T|$  is handled identically to [Section 5.3](#) (imagine that the tensor  $T \equiv 1$ ). So in the remainder of the section, we focus on the case  $|R| \geq |S||T|$ .

### 7.3 Handling the Case $|R| \geq |S||T|$

In this case we use the master relaxed base case ([Definition 4.10](#)). We use the analogue of [Lemma 5.10](#), whose proof is identical.

**Lemma 7.8.** If  $|R| \geq |S||T|$  then for all  $s \in S, t \in T$  we have  $\sum_{r \in R} |\widehat{F}'_r(s, t)|^2 \geq 1 - \frac{4\delta M^5}{\Delta^2}$ .

For  $s \in S, t \in T$ , define the functions

$$\widetilde{F}_{st} = \frac{\sum_{r \in R} \overline{\widehat{F}'_r(s, t)} F'_r}{\sqrt{\sum_{r \in R} |\widehat{F}'_r(s, t)|^2}}.$$

Note that  $\|\widetilde{F}_{st}\|_2 = 1$  and

$$\left( \mathbb{E}_{(x, y, z) \sim \mu} \left[ \widetilde{F}_{st}(x) G'_s(y) H'_t(z) \right] \right)^2 = \sum_{r \in R} |\widehat{F}'_r(s, t)|^2 \geq 1 - \frac{4\delta M^5}{\Delta^2}, \quad (22)$$

by [Lemma 7.8](#). Now we leverage the master relaxed base case property of  $\mu$  to get that

$$I_{\text{modest}}[\widetilde{F}_{st}] \leq \left( \frac{4\delta M^5}{c\Delta^2} \right)^{\frac{1}{100M}} := \delta',$$

for all  $s \in S$  and  $t \in T$ . The remainder of the proof converts this into an effective influence bound on the  $F'_r$  functions themselves to contradict [Lemma 7.2](#).

In the following, treat all functions to be on the space  $L_2(\Gamma \times \Phi, \mu_{y, z})$ , where we recall that  $(y, z)$  is a product distribution and determines  $x$  uniquely in the relaxed base case. Note that for all  $s \in S, t \in T$  that

$$\|\widetilde{F}_{st} - G'_s H'_t\|_2 = \|\widetilde{F}_{st}\|_2^2 + \|G'_s H'_t\|_2^2 - 2\mathbb{E}_{(x, y, z)} \left[ \widetilde{F}_{st}(x) G'_s(y) H'_t(z) \right] \leq \frac{8\delta M^5}{\Delta^2}, \quad (23)$$

where we have used (22) and that the expectation is real. For all  $r \in R$  we have

$$\left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t)} \widetilde{F}_{st} \right\|_2 \leq \left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t)} G'_s H'_t \right\|_2 + \sum_{s \in S, t \in T} \widehat{F}'_r(s, t) \left\| G'_s H'_t - \widetilde{F}_{st} \right\|_2.$$

By (23) the second term can be bounded by

$$\sum_{s \in S, t \in T} \widehat{F}'_r(s, t) \left\| G'_s H'_t - \widetilde{F}_{st} \right\|_2 \leq M^2 \cdot \frac{8\delta M^5}{\Delta^2} = \frac{8\delta M^7}{\Delta^2}.$$

The first term can be simplified to get

$$\left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) G'_s H'_t} \right\|_2 = \left( \|F'_r\|_2^2 - \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)|^2 \right)^{1/2} \leq \sqrt{\frac{4\delta M^3}{\Delta^2}},$$

where we have used that the functions  $G'_s H'_t$  are orthonormal (because  $\mu_{y, z}$  is a product distribution), and [Lemma 7.7](#). By the triangle inequality again, we get

$$\begin{aligned} I_{\text{modest}}[F'_r] &\leq I_{\text{modest}} \left[ \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) \widetilde{F}_{st}} \right] + \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)| I_{\text{modest}}[\widetilde{F}_{st}] \\ &\leq \left\| \overline{F'_r} - \sum_{s \in S, t \in T} \overline{\widehat{F}'_r(s, t) \widetilde{F}_{st}} \right\|_2 + \sum_{s \in S, t \in T} |\widehat{F}'_r(s, t)| I_{\text{modest}}[\widetilde{F}_{st}] \\ &\leq \sqrt{\frac{4\delta M^3}{\Delta^2}} + \frac{8\delta M^7}{\Delta^2} + M^2 \delta' \leq 2M^2 \delta', \end{aligned}$$

by our choice of  $\delta = \left(\frac{\Delta}{CM\tau}\right)^{1000M}$ . This gives that

$$\sum_{r \in [M]} \kappa_r^2 I_{\text{modest}}[F'_r] \leq M\delta^2 + \sum_{r \in R} \kappa_r^2 I_{\text{modest}}[F'_r] \leq M\delta^2 + 2M^2 \delta' < \tau,$$

by the choice of  $\delta$  again, because  $\delta' = \left(\frac{4\delta M^5}{c\Delta^2}\right)^{\frac{1}{100M}}$ . This contradicts [Lemma 7.2](#).

## 8 Additive Combinatorics

This section is dedicated to proving [Theorem 3](#).

**Theorem 3.** *Let  $\mathcal{A}$  be a finite abelian group and  $t : \mathcal{A}^3 \rightarrow \mathbb{Z}_{p^k}$ . Let  $N > 1$  be minimal so that there are functions  $a, b, c : \mathcal{A} \rightarrow \mathbb{Z}_{p^k N}$  with  $a(x) + b(y) + c(z) = N \cdot t(x, y, z) \pmod{p^k N}$  for all  $x + y + z = 0$ . Then there is a constant  $c = c(t, \mathcal{A}) > 0$  such that for any  $f, g, h : \mathcal{A}^n \rightarrow \mathbb{Z}_{p^k}$ ,*

$$\Pr_{\substack{(x, y, z) \in (\mathcal{A}^n)^3 \\ x + y + z = 0}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \pmod{p^k}, i = 1, 2, \dots, n] \leq 2^{-cn}.$$

We show [Theorem 3](#) in the remainder of this section. By minimality, we may assume that  $N$  is a power of  $p$ . Let  $N = p^j$ . Assume that

$$\Pr_{\substack{(x, y, z) \in (\mathcal{A}^n)^3 \\ x + y + z = 0}} [f(x)_i + g(y)_i + h(z)_i = t(x_i, y_i, z_i) \pmod{p^k}, i = 1, 2, \dots, n] \geq \varepsilon,$$

where  $\varepsilon > 2^{-cn}$ , for contradiction. Then writing  $\tilde{f}(x) = N \cdot f(x) - a(x)$ ,  $\tilde{g}(y) = N \cdot g(y) - b(y)$ , and  $\tilde{h}(z) = N \cdot h(z) - c(z)$ , where  $\tilde{f}, \tilde{g}, \tilde{h} : \mathcal{A}^n \rightarrow \mathbb{Z}_{p^k N}^n$ , we get that

$$\Pr_{\substack{(x, y, z) \in (\mathcal{A}^n)^3 \\ x + y + z = 0}} [\tilde{f}(x) + \tilde{g}(y) + \tilde{h}(z) = 0 \pmod{p^k N}, i = 1, 2, \dots, n] \geq \varepsilon.$$

We can follow the approach [[BKM23c](#), Section 2] to conclude the following.

**Lemma 8.1.** *There is a subset  $\mathcal{A}' \subseteq \mathcal{A}^n$  with  $|\mathcal{A}'| \geq \Omega(\varepsilon^{O(p^k N)}) |\mathcal{A}|^n$  such that  $\tilde{f}$  is an order- $p^k N$  Freiman homomorphism on  $\mathcal{A}'$ .*

At this point, write  $\mathcal{A} = \prod_{i=1}^L \mathbb{Z}_{p_i}^{k_i}$ . We have the map  $a : \mathcal{A} \rightarrow \mathbb{Z}_{p^k N}$ . Consider its restriction modulo  $p$ , which we write as  $\bar{a} : \mathcal{A} \rightarrow \mathbb{Z}_p$  defined as  $\bar{a}(x) = a(x) \pmod{p}$ . We show that  $\bar{a}$  must be the shift of a homomorphism.

**Lemma 8.2.** *For sufficiently small constants  $c$ , if  $\varepsilon \geq 2^{-cn}$  then  $\bar{a} : \mathcal{A} \rightarrow \mathbb{Z}_p$  must take the form  $\bar{a}(x) = c + \sum_{i=1}^L c_i x_i \pmod{p}$ , where  $x = (x_1, \dots, x_L) \in \mathcal{A}$ .*

*Proof.* Because  $\tilde{f}$  is an order 4 Freiman homomorphism we know for all  $u, v, w, x \in (\mathcal{A}')^4$  with  $u + v = w + x$  we have  $\tilde{f}(u) + \tilde{f}(v) = \tilde{f}(w) + \tilde{f}(x) \pmod{p^k N}$ . Recall that  $\tilde{f}(x) = N \cdot f(x) - a(x) = -\bar{a}(x) \pmod{p}$ . So if  $u + v = w + x$  then  $\bar{a}(u) + \bar{a}(v) = \bar{a}(w) + \bar{a}(x) \pmod{p}$ . The number of such 4-tuples is at least  $|\mathcal{A}'|^4 / |\mathcal{A}|^n \geq \Omega(\varepsilon^{O(p^k N)}) |\mathcal{A}|^{3n}$ , so  $\bar{a}$  satisfies:

$$\Pr_{(u,v,w,x) \in (\mathcal{A}^n)^4: u+v=w+x} [\bar{a}(u_i) + \bar{a}(v_i) = \bar{a}(w_i) + \bar{a}(x_i) \pmod{p} \quad \forall i \in [n]] \geq \Omega(\varepsilon^{O(p^k N)}).$$

If  $\bar{a}(u) + \bar{a}(v) \neq \bar{a}(w) + \bar{a}(x)$  for some  $(u, v, w, x) \in \mathcal{A}^4$  with  $u + v = w + x$  then the above probability is at most  $(1 - 1/|\mathcal{A}|^3)^n \ll \varepsilon^{O(p^k N)}$ . So  $\bar{a}(u) + \bar{a}(v) = \bar{a}(w) + \bar{a}(x)$  for all  $u + v = w + x$ . This implies the conclusion.  $\square$

Now we show that many of the coefficients  $c_i = 0$  in fact.

**Lemma 8.3.** *Under the hypotheses of Lemma 8.2 we have that  $c_i = 0$  if  $p_i \neq p$ , or  $p_i = p$  and  $k_i < k + j$ .*

*Proof.* If  $c_i \neq 0$  for some  $p_i \neq p$ , then  $\bar{a}$  is not the shift of a homomorphism from  $\mathcal{A} \rightarrow \mathbb{Z}_p$ .

Now consider  $i$  with  $p_i = p$  and  $k_i < k + j$ . Split  $\mathcal{A}^n$  into  $|\mathcal{A}|^n / p^n$  groups as follows. We first describe how to split  $(\mathbb{Z}_{p^{k_i}})^n$  into  $(p^{k_i})^n / p^n$  groups. This is done by splitting  $\mathbb{Z}_{p^{k_i}}$  into  $p^{k_i-1}$  groups of size  $p$ :

$$\{0, 1, \dots, p-1\} \cup \{p, p+1, \dots, 2p-1\}, \dots, \{p^{k_i} - p, \dots, p^{k_i} - 1\},$$

and then taking the direct product  $n$  times. The point is that for any  $x, y \in (\mathbb{Z}_{p^{k_i}})^n$  in the same group,  $x \not\equiv y \pmod{p}$ , as  $n$ -dimensional vectors. Now we split  $\mathcal{A}^n$  as follows:  $(x_1, \dots, x_L)$  and  $(y_1, \dots, y_L)$  are in the same group if and only if  $x_{i'} = y_{i'}$  for any  $i' \neq i$ , and  $(x_i, y_i)$  are in the same group of  $(\mathbb{Z}_{p^{k_i}})^n$  as described.

Because  $|\mathcal{A}'| > |\mathcal{A}|^n / p^n$ , the Pigeonhole Principle tells us that we can find  $x, y \in \mathcal{A}'$  in the same group. By construction,  $p^{k_i} x = p^{k_i} y$ . Because  $\tilde{f}$  was an order  $p^k N > p^{k_i}$  Freiman homomorphism, we deduce that  $p^{k_i} \tilde{f}(x) = p^{k_i} \tilde{f}(y) \pmod{p^{k+j}}$ , so  $\tilde{f}(x) = \tilde{f}(y) \pmod{p}$  as  $k_i < k + j$ . Thus,  $\bar{a}(x) = \bar{a}(y)$ . Because  $x_{i'} = y_{i'}$  for all  $i' \neq i$  by construction, using Lemma 8.2 gives us  $c_i(x_i - y_i) = 0 \pmod{p}$ . By construction  $x_i \not\equiv y_i \pmod{p}$  as vectors, so  $c_i = 0$ .  $\square$

By symmetry, we can apply both Lemmas 8.2 and 8.3 to  $b, c$  to write:

$$\bar{b}(y) = d + \sum_{i=1}^L d_i y_i \pmod{p} \quad \text{and} \quad \bar{c}(z) = e + \sum_{i=1}^L e_i z_i \pmod{p}.$$

We will use this to contradict minimality of  $N$ . To start, we first show a relationship between the coefficients  $c_i, d_i, e_i$ .

**Lemma 8.4** (Matching coefficients). *If  $\bar{a}(x) = c + \sum_{i=1}^L c_i x_i$ ,  $\bar{b}(y) = d + \sum_{i=1}^L d_i y_i$ , and  $\bar{c}(z) = e + \sum_{i=1}^L e_i z_i$  then  $c + d + e = 0 \pmod{p}$  and  $c_i = d_i = e_i$  (as elements of  $\mathbb{Z}_p$ ).*



*Proof.* Recall that  $\bar{a}(x) + \bar{b}(y) + \bar{c}(z) = N \cdot t(x, y, z) = 0 \pmod{p}$  for all  $x + y + z = 0$  in  $\mathcal{A}$ . Thus  $c + d + e = 0 \pmod{p}$  and  $c_i = d_i = e_i$ .  $\square$

Finally we reach a contradiction by showing that  $N$  was not minimal.

**Lemma 8.5.** *Assume  $c + d + e = 0$  over  $\mathbb{Z}$ . Define  $\tilde{a}(x) = a(x) - c - \sum_{i=1}^L c_i x_i$ ,  $\tilde{b}(y) = b(y) - d - \sum_{i=1}^L c_i y_i$ ,  $\tilde{c}(z) = c(z) - e - \sum_{i=1}^L c_i z_i$ . Then  $\tilde{a}(x) + \tilde{b}(y) + \tilde{c}(z) = a(x) + b(y) + c(z) \pmod{p^k N}$  and  $\tilde{a}(x), \tilde{b}(y), \tilde{c}(z) = 0 \pmod{p}$  for any  $x, y, z \in \mathcal{A}$ .*

*Proof.* The final claim from Lemma 8.2. For all  $(x, y, z) \in \mathcal{A}^3$  with  $x + y + z = 0$  we know that

$$(a(x) + b(y) + c(z)) - (\tilde{a}(x) + \tilde{b}(y) + \tilde{c}(z)) = (c + d + e) + \sum_{i=1}^L c_i(x_i + y_i + z_i).$$

If  $p_i = p$  and  $k_i < k + j$  then  $c_i = 0$ . Otherwise,  $x_i + y_i + z_i = 0 \pmod{p^k N}$ , as  $N = p^j$ . In both cases,  $\sum_{i=1}^L c_i(x_i + y_i + z_i) = 0 \pmod{p^k N}$  as desired.  $\square$

This contradicts the minimality of  $N$ , because we can shift  $c, d, e$  by multiples of  $p$  to ensure that  $c + d + e = 0$  (recall that  $c + d + e = 0 \pmod{p}$  by Lemma 8.4), and then divide everything by  $p$  to reduce  $N$  to  $N/p$ .

## References

- [ABSS97] Sanjeev Arora, László Babai, Jacques Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *J. Comput. Syst. Sci.*, 54(2):317–331, 1997.
- [AK09] Noga Alon and Bo’az Klartag. Economical toric spines via Cheeger’s inequality. *J. Topol. Anal.*, 1(2):101–111, 2009.
- [ALM<sup>+</sup>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.
- [AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.
- [BBCR13] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM J. Comput.*, 42(3):1327–1363, 2013.
- [BBLV13] Jop Briët, Harry Buhrman, Troy Lee, and Thomas Vidick. Multipartite entanglement in XOR games. *Quantum Inf. Comput.*, 13(3-4):334–360, 2013.
- [BG15] Mark Braverman and Ankit Garg. Small value parallel repetition for general games. In Rocco A. Servedio and Ronitt Rubinfeld, editors, *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 335–340. ACM, 2015.
- [BGKW88] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In Janos Simon, editor, *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 113–131. ACM, 1988.
- [BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, pcps, and nonapproximability-towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998.

- [BKM22] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable  $k$ -csp: I. In Stefano Leonardi and Anupam Gupta, editors, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 976–988. ACM, 2022.
- [BKM23a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable  $k$ -csp: II. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 632–642. ACM, 2023.
- [BKM23b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable  $k$ -csp: III. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 643–655. ACM, 2023.
- [BKM23c] Mark Braverman, Subhash Khot, and Dor Minzer. Parallel repetition for the GHZ game: Exponential decay. In *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2023, Santa Cruz, CA, USA, November 6-9, 2023*, pages 1337–1341. IEEE, 2023.
- [BKM24] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable  $k$ -csp: IV. In Bojan Mohar, Igor Shinkar, and Ryan O’Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1423–1434. ACM, 2024.
- [BM21] Mark Braverman and Dor Minzer. Optimal tiling of the euclidean space using permutation-symmetric bodies. In Valentine Kabanets, editor, *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPICs*, pages 5:1–5:48. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.
- [BRWY13] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 746–755. IEEE Computer Society, 2013.
- [CHTW04] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 236–249. IEEE Computer Society, 2004.
- [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.
- [DHVY17] Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expanding games. In *8th Innovations in Theoretical Computer Science Conference*, volume 67 of *LIPICs. Leibniz Int. Proc. Inform.*, pages Art. No. 37, 16. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2017.
- [DRS05] Irit Dinur, Oded Regev, and Clifford D. Smyth. The hardness of 3-uniform hypergraph coloring. *Comb.*, 25(5):519–535, 2005.
- [DS14] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 624–633. ACM, 2014.
- [Fei91] Uriel Feige. On the success probability of the two provers in one-round proof systems. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference, Chicago, Illinois, USA, June 30 - July 3, 1991*, pages 116–123. IEEE Computer Society, 1991.
- [Fei98] Uriel Feige. A threshold of  $\ln n$  for approximating set cover. *J. ACM*, 45(4):634–652, 1998.
- [FGL<sup>+</sup>96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996.

- [FKO07] Uriel Feige, Guy Kindler, and Ryan O’Donnell. Understanding parallel repetition requires understanding foams. In *22nd Annual IEEE Conference on Computational Complexity (CCC 2007), 13-16 June 2007, San Diego, California, USA*, pages 179–192. IEEE Computer Society, 2007.
- [FRS94] Lance Fortnow, John Rompel, and Michael Sipser. On the power of multi-prover interactive protocols. *Theor. Comput. Sci.*, 134(2):545–557, 1994.
- [FV02] Uriel Feige and Oleg Verbitsky. Error reduction by parallel repetition—a negative result. *Combinatorica*, 22(4):461–478, 2002.
- [GHM<sup>+</sup>21] Uma Girish, Justin Holmgren, Kunal Mittal, Ran Raz, and Wei Zhan. Parallel repetition for the GHZ game: A simpler proof. *arXiv preprint arXiv:2107.06156*, 2021. Available at <https://arxiv.org/pdf/2107.06156.pdf>.
- [GHM<sup>+</sup>22] Uma Girish, Justin Holmgren, Kunal Mittal, Ran Raz, and Wei Zhan. Parallel repetition for all 3-player games over binary alphabet. In Stefano Leonardi and Anupam Gupta, editors, *STOC ’22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 998–1009. ACM, 2022.
- [GHS02] Venkatesan Guruswami, Johan Hastad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM J. Comput.*, 31(6):1663–1686, 2002.
- [GMRZ22] Uma Girish, Kunal Mittal, Ran Raz, and Wei Zhan. Polynomial bounds on parallel repetition for all 3-player games with binary inputs. In Amit Chakrabarti and Chaitanya Swamy, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2022, September 19-21, 2022, University of Illinois, Urbana-Champaign, USA (Virtual Conference)*, volume 245 of *LIPICs*, pages 6:1–6:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
- [Gow01] William T Gowers. A new proof of Szemerédi’s theorem. *Geometric & Functional Analysis GAF*, 11(3):465–588, 2001.
- [Has01] Johan Hastad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.
- [Hol09] Thomas Holenstein. Parallel repetition: Simplification and the no-signaling case. *Theory Comput.*, 5(1):141–172, 2009.
- [HR20] Justin Holmgren and Ran Raz. A parallel repetition theorem for the ghz game. *arXiv preprint arXiv:2008.05059*, 2020. Available at <https://arxiv.org/pdf/2008.05059.pdf>.
- [Kho02a] Subhash Khot. Hardness results for approximate hypergraph coloring. In John H. Reif, editor, *Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada*, pages 351–359. ACM, 2002.
- [Kho02b] Subhash Khot. Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *43rd Symposium on Foundations of Computer Science (FOCS 2002), 16-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 23–32. IEEE Computer Society, 2002.
- [KORW08] Guy Kindler, Ryan O’Donnell, Anup Rao, and Avi Wigderson. Spherical cubes and rounding in high dimensions. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 189–198. IEEE Computer Society, 2008.
- [Pol12] D. H. J. Polymath. A new proof of the density Hales-Jewett theorem. *Ann. of Math. (2)*, 175(3):1283–1327, 2012.
- [PRW97] Itzhak Parnafes, Ran Raz, and Avi Wigderson. Direct product results and the GCD problem, in old and new communication models. In Frank Thomson Leighton and Peter W. Shor, editors, *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, El Paso, Texas, USA, May 4-6, 1997*, pages 363–372. ACM, 1997.
- [Rao11] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011.

- [Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.
- [Ver96] Oleg Verbitsky. Towards the parallel repetition conjecture. *Theoret. Comput. Sci.*, 157(2):277–282, 1996.

## A Proof of Lemma 4.12

*Proof.* Fix  $\rho_0, \dots, \rho_d \in [0, 1 - \varepsilon]$  to be chosen later, and let  $c_j = \frac{\prod_{i \neq j} (1 - \rho_i)}{\prod_{i \neq j} (\rho_j - \rho_i)}$  for  $j = 0, \dots, d$ . We will show that [Items 1 and 2 of Lemma 4.12](#) are satisfied for any choice of  $\rho_j$ .

For [Item 1](#) (where  $k \leq d$ ) consider the polynomial  $p(x) = x^k$ . By the Lagrange interpolation formula, we know that

$$p(x) = \sum_{j=0}^d \frac{\prod_{i \neq j} (x - \rho_i)}{\prod_{i \neq j} (\rho_j - \rho_i)} p(\rho_j) = \sum_{j=0}^d \frac{\prod_{i \neq j} (x - \rho_i)}{\prod_{i \neq j} (\rho_j - \rho_i)} \rho_j^k.$$

Taking  $x = 1$  gives us  $1 = p(1) = \sum_{j=0}^d c_j \rho_j^k$ , as desired.

Define  $A_k = \sum_{j=0}^d c_j \rho_j^k$ . We will show that  $A_k \geq A_{k+1}$  for all  $k \geq 0$ . Because  $\lim_{k \rightarrow \infty} A_k = 0$  evidently, and  $A_d = 1$  by the above,  $0 \leq A_k \leq 1$  for all  $k \geq d + 1$ , as desired. Note that

$$A_k - A_{k+1} = \sum_{j=0}^d \frac{\prod_{i \neq j} (1 - \rho_i)}{\prod_{i \neq j} (\rho_j - \rho_i)} (\rho_j^k - \rho_j^{k+1}) = \prod_{j=0}^d (1 - \rho_j) \cdot \sum_{j=0}^d \frac{\rho_j^k}{\prod_{i \neq j} (\rho_j - \rho_i)}.$$

The positivity of this expression follows from  $\prod_{j=0}^d (1 - \rho_j) > 0$ , as  $\rho_0, \dots, \rho_d < 1$ , and the identity

$$\sum_{j=0}^d \frac{\rho_j^k}{\prod_{i \neq j} (\rho_j - \rho_i)} = \sum_{\substack{a_0 + \dots + a_d = k - d \\ a_0, \dots, a_d \in \mathbb{Z}_{\geq 0}}} \rho_0^{a_0} \rho_1^{a_1} \dots \rho_d^{a_d} \geq 0. \quad (24)$$

We establish this identity below, in [Lemma A.1](#).

We proceed to [Item 3](#). Let

$$T_d(x) := \frac{1}{2} \left( (x + \sqrt{x^2 - 1})^d + (x - \sqrt{x^2 - 1})^d \right) \quad (25)$$

denote the  $d$ -th Chebyshev polynomial, and let  $\tilde{T}(x) = T_d(\frac{2}{1-\varepsilon}x - 1)$ . Let  $\rho_j = \frac{1-\varepsilon}{2}(\cos(j\pi/d) + 1)$ . Note that  $0 \leq \rho_j \leq 1 - \varepsilon$ , and  $\tilde{T}(\rho_j) = T_d(\cos(j\pi/d)) = \cos(j\pi) = (-1)^j$ . Let  $a_0, \dots, a_d$  be the coefficients of  $\tilde{T}$ , i.e.,  $\tilde{T}(x) = \sum_{k=0}^d a_k x^k$ . We have that

$$\tilde{T}(1) = \sum_{k=0}^d a_k = \sum_{k=0}^d a_k \sum_{j=0}^d c_j \rho_j^k = \sum_{j=0}^d c_j \sum_{k=0}^d a_k \rho_j^k = \sum_{j=0}^d c_j \tilde{T}(\rho_j) = \sum_{j=0}^d c_j (-1)^j = \sum_{j=0}^d |c_j|,$$

where the final equality uses the definition  $c_j = \frac{\prod_{i \neq j} (1 - \rho_i)}{\prod_{i \neq j} (\rho_j - \rho_i)}$  to deduce that  $c_j (-1)^j > 0$  for all  $j$  (recall that  $\rho_0 > \rho_1 > \dots > \rho_d$ ). To conclude, note that

$$\tilde{T}(1) = T \left( \frac{1 + \varepsilon}{1 - \varepsilon} \right) \leq T(1 + 4\varepsilon) \leq \exp(O(d\sqrt{\varepsilon})),$$

where the final inequality comes from taking  $x = 1 + 4\varepsilon$  in [\(25\)](#).  $\square$

**Lemma A.1.** (24) holds for any distinct real numbers  $\rho_0, \dots, \rho_d$ , and integer  $k \geq d$ .

*Proof.* Using the Vandermonde determinant, one can compute that

$$\sum_{j=0}^d \frac{\rho_j^k}{\prod_{i \neq j} (\rho_j - \rho_i)} = \det \begin{bmatrix} \rho_0^k & \rho_1^k & \cdots & \rho_d^k \\ \rho_0^{d-1} & \rho_1^{d-1} & \cdots & \rho_d^{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} / \det \begin{bmatrix} \rho_0^d & \rho_1^d & \cdots & \rho_d^d \\ \rho_0^{d-1} & \rho_1^{d-1} & \cdots & \rho_d^{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

This is the Schur polynomial corresponding to the partition  $\lambda := (0, 0, \dots, k - d)$ . The Giambelli formula says that this the Schur polynomial can be alternatively expressed as

$$\sum_T \rho_0^{t_0} \cdots \rho_d^{t_d},$$

where  $T$  ranges over all semistandard Young Tableaux for partition  $\lambda$ , and  $t_i$  is the number of occurrences of  $i$  in  $T$ . Evidently, this is equivalent to the RHS of (24), for  $\lambda$  with a single piece of size  $k - d$ , and labels between  $0, \dots, d$ .  $\square$