



On Approximability of Satisfiable k -CSPs: VII

Amey Bhangale* Subhash Khot† Yang P. Liu‡ Dor Minzer§

Abstract

Let $\Sigma_1, \dots, \Sigma_k$ be finite alphabets, and let μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ in which the probability of each atom is at least α . We prove that if μ does not admit Abelian embeddings, and $f_i : \Sigma_i \rightarrow \mathbb{C}$ are 1-bounded functions (for $i = 1, \dots, k$) such that

$$\left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} [f_1(x_1) \dots f_k(x_k)] \right| \geq \varepsilon,$$

then there exists $L : \Sigma_1^n \rightarrow \mathbb{C}$ of degree at most d and $\|L\|_2 \leq 1$ such that $|\langle f_1, L \rangle| \geq \delta$, where d and $\delta > 0$ depend only on k, α and ε . This answers the analytic question posed by Bhangale, Khot, and Minzer (STOC 2022). We also prove several extensions of this result that are useful in subsequent applications.

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

†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.

1 Introduction

This paper continues the investigation of the approximability of constraints satisfaction problems [BKM22, BKM23a, BKM23b, BKM24a, BKM24b, BKLM24a]. While previous papers dealt with 3-ary predicates, the focus of the current paper is on k -ary predicates for $k > 3$. Our primary contribution is a set of new analytical inequalities for a general family of k -ary distributions, extending results of Mossel [Mos10] about the class of connected distributions.

1.1 Constraint Satisfaction Problems

The Decision Version

Constraint satisfaction problems (CSPs in short) are some of the most fundamental problems in computer science. For a finite alphabet Σ and a predicate $P : \Sigma^k \rightarrow \{0, 1\}$, an instance Ψ of the problem P -CSP consists of a set of variables x_1, x_2, \dots, x_n and a collection of constraints C_1, C_2, \dots, C_m , where each constraint is of the form $P(x_{i_1}, x_{i_2}, \dots, x_{i_k}) = 1$. The goal is to decide, given an instance Ψ , whether there is an assignment $A : \{x_1, \dots, x_n\} \rightarrow \Sigma$ satisfying all of the constraints or not. For a collection of predicates $\mathcal{P} \subseteq \{P : \Sigma^k \rightarrow \{0, 1\}\}$, the problem \mathcal{P} -CSP is defined analogously, where now each constraint C_i takes the form $P(x_{i_1}, x_{i_2}, \dots, x_{i_k}) = 1$ for some $P \in \mathcal{P}$.¹

A systematic study of the complexity of solving CSPs began by Schaefer [Sch78], who demonstrated that for any \mathcal{P} with $|\Sigma| = 2$, the problem \mathcal{P} -CSP is either in **P** or else is **NP**-complete. The well-known Dichotomy Conjecture by Feder and Vardi [FV98], suggested that this assertion holds for all finite alphabets Σ . This conjecture was recently confirmed by Bulatov and independently by Zhuk [Bul17, Zhu20], following a long line of work using the methods of abstract algebra.

The Optimization Version

In the maximization version of \mathcal{P} -CSP, called Max- \mathcal{P} -CSP, the task is to find an assignment to the variables that satisfies the maximum fraction of the constraints. An α -approximation algorithm is a polynomial-time algorithm which always returns an assignment satisfying at least $\alpha \cdot \text{OPT}$ fraction of the constraints, where OPT is the value of the optimum assignment. The focus of the current work is on approximability of fully satisfiable instances, meaning the case that the algorithm is guaranteed that $\text{OPT} = 1$.

The PCP Theorem [FGL+96, ALM+98, AS98] implies that it is **NP**-hard to approximate many CSPs within some constant factor $\alpha < 1$. Håstad [Hås01] showed that for specific CSPs, one can obtain hardness results for much better approximation ratios. For instance, he showed that for any $\varepsilon > 0$, 3-SAT cannot be approximated within factor $\frac{7}{8} + \varepsilon$ in polynomial time, unless **P** = **NP**.² His result [Hås01] is in fact stronger: even if the formula is promised to be *fully satisfiable*, it is still **NP**-hard to find an assignment satisfying more than $(\frac{7}{8} + \varepsilon)$ -fraction of the clauses. In another notable result, Håstad showed given 3-LIN instance³ promised to be $(1 - \varepsilon)$ -satisfiable, it is **NP**-hard to find an assignment satisfying more than $(\frac{1}{2} + \varepsilon)$ -fraction of the constraints. Unlike 3-SAT, under the stronger promise that the instance is fully satisfiable, one can efficiently find an assignment that satisfies all constraints. Thus, knowing that a CSP instance is fully satisfiable sometimes allows for better approximation algorithms.

¹Some well known problems, such as 3-SAT for example, involve taking negations of variables. To put such problems in the framework we discuss, one takes a collection \mathcal{P} of 8 predicates, corresponding to the predicate $P(x, y, z) = x \vee y \vee z$ applied on all negation patterns on 3 variables.

²Note that a random assignment satisfies $\frac{7}{8}$ -fraction of the clauses in expectation, so Håstad's result in fact asserts that the best approximation algorithm for 3-SAT is (essentially) obtained by randomly sampling an assignment.

³This CSP is over \mathbb{F}_2 , with constraints of the type $x_{i_1} \oplus x_{i_2} \oplus x_{i_3} = 1$ or 0.

By now, the theory of approximability of *almost satisfiable* CSPs is pretty well understood. Indeed, assuming the Unique-Games Conjecture [Kho02], we know that the best possible polynomial time algorithms for any CSP (in the almost satisfiable regime) is given by a semi-definite programming algorithm (SDP in short). The first such result is from [KKMO07, MOO05], where it is proved that assuming the Unique-Games Conjecture, the Goemans and Williamson [GW95] approximation algorithm for the Max-Cut problem is essentially the best possible. Raghavendra [Rag08] generalized this result to *all* CSPs. To do so, he formulated a natural SDP relaxation of any given CSP instance, and showed that a generic randomized rounding procedure achieves the best possible approximation ratio. Towards this end, he showed that any (c, s) integrality gap instance for this relaxation can be transformed to a hardness result: assuming the Unique-Games Conjecture, for all $\varepsilon > 0$, given an instance promised to be $(c - \varepsilon)$ -satisfiable, it **NP**-hard to find an assignment that satisfies at least $(s + \varepsilon)$ -fraction of the constraints.

In comparison, the theory of approximability of satisfiable CSPs is much less developed. Besides results in a few special cases, such as the result of Håstad [Hås01] discussed above, not much is known, even assuming conjectures in the spirit of the Unique-Games Conjecture. One of the primary goals of the recent line of research [BKM22, BKM23a, BKM23b, BKM24a, BKM24b, BKLM24a] is to further develop this theory for general classes of predicates. Ultimately, the hope is to establish analogous results to Raghavendra’s Theorem [Rag08] (but for the case of satisfiable instances) under conjectures such as the Rich 2-to-1 Games Conjecture [BKM21] (see also [BKLM22]). The current paper is a continuation of this line of research.

1.2 Towards Satisfiable CSPs

The most important building-block in Raghavendra’s result (as well as in many other PCPs) is the construction of a suitable dictatorship test. A function $f : \Sigma^n \rightarrow \Sigma$ is called a dictatorship function if $f(x)$ depends only on one of x ’s coordinates. A dictatorship test is a randomized procedure which queries f at a few (correlated) locations randomly and, based on these, decides if f is a dictatorship function or far from any dictator function.

There are three important properties of the test which are useful in getting hardness of approximation results for CSPs. The first one is the *completeness parameter* c , which is the probability that the test accepts a dictatorship function. The second property is the *soundness parameter* s , which is the maximum probability the test accepts a far-from-dictatorship function. The third property is the decision predicate that the test uses. If the test uses predicates from the class \mathcal{P} , has completeness c and soundness s , then it can be used to prove a UG-hardness result for Max- \mathcal{P} -CSP with completeness $(c - \varepsilon)$ and soundness $(s + \varepsilon)$, for any constant $\varepsilon > 0$.

Raghavendra proved his result by showing how to transform a (c, s) integrality gap instance for Max- \mathcal{P} -CSP to a dictatorship test with completeness $(c - \varepsilon)$, soundness $(s + \varepsilon)$ and predicates \mathcal{P} , and then combined it with a reduction from Unique-Games. His technique inherently cannot be used to prove hardness results for satisfiable instances. For once, using the Unique-Games Conjecture, one inherently loses perfect completeness, as Unique-Games are easy on satisfiable instances. However, this may not be a very fundamental loss, as there is a feasible variant of the Unique-Games Conjecture, known as the Rich 2-to-1 Games Conjecture [BKM21], that has perfect completeness. It is plausible that assuming the Rich 2-to-1 Games Conjecture, one could transform dictatorship tests into hardness results without any loss in the completeness. Therefore, it seems that the more fundamental place where perfect completeness is lost is in the design of the dictatorship test.

The paper [BKM22] initiated a systematic study of completely characterizing the precise approximability of every k -ary CSP on satisfiable instances. They considered natural dictatorship

tests for a large class of predicates similar to the ones given by Raghavendra, except that they do not lose the perfect completeness which we now describe. The starting point is an instance Ψ of P -CSP, whose SDP value is 1, and the integral value (i.e., maximum fraction of the constraints that can be satisfied by an assignment) is s . The SDP solution consists of vectors as well as local distribution for each constraint. Since the SDP value is 1, all these local distributions are supported on the satisfying assignments to P . Let μ_i be the local distribution corresponding to the i^{th} constraint of the instance. The test is as follows (here $\varepsilon > 0$ is a small constant independent of n):

Given $f : \Sigma^n \rightarrow \Sigma$,

1. Select a constraint C_i from Ψ with probability proportional to its weight.
2. Construct a $k \times n$ matrix by sampling each column of the matrix independently according to μ_i .
3. Let x_j be the j^{th} row of the matrix.
4. Check if $P(f(x_1), f(x_2), \dots, f(x_k)) = 1$.

If f is a dictatorship function, then the test accepts with probability 1. This follows because for every i , the distribution μ_i is supported on the satisfying assignments to P and therefore every column of the matrix is from $P^{-1}(1)$. What about the acceptance probability of the test when f is far from dictator functions? When a small loss in the completeness parameter is allowed, one can modify the distributions μ_i slightly and ensure that they are fully supported on Σ^k , in which case one can appeal to results of Mossel [Mos10]. When no loss in completeness is allowed, one has to work with the distribution μ_i as is, and this task becomes significantly harder.

The paper [BKM22] relates the soundness analysis of the test to the following k -wise correlation of functions $f_i : \Sigma_i^n \rightarrow [-1, 1]$ with respect to the distribution μ supported on $P^{-1}(1)$. Specifically, the following general question was posed in [BKM22].

Question 1. (Informal) Find the necessary and sufficient condition on a distribution μ over $\Sigma_1 \times \dots \times \Sigma_k$, such that

$$\left| \mathbb{E}_{(x_1, x_2, \dots, x_k) \sim \mu^{\otimes n}} [f_1(x_1) f_2(x_2) \cdots f_k(x_k)] \right| \rightarrow 0 \quad \text{as } d \rightarrow \infty, \quad (1)$$

for all functions $f_i : \Sigma_i^n \rightarrow [-1, 1]$ such that at least one function (essentially) has degree at least d .

Mossel [Mos10] showed that if the distribution μ is *connected*, then (1) as above holds. Here, a distribution μ is called connectedness if looking at the graph whose vertices consist of $\text{supp}(\mu)$, and two vertices are adjacent if they differ on exactly one coordinate, is connected. The connectedness condition, however, is not necessary, as noted implicitly in [BK21]. Let G be a non-Abelian group with no dimension one representation, and take $P : G^3 \rightarrow \{0, 1\}$ as $P^{-1}(1) = \{(x, y, z) \mid x \cdot y \cdot z = 1_G\}$, along with the distribution μ that is uniform on $P^{-1}(1)$. The distribution μ is not connected, but yet [BK21] prove that (1) still holds.

1.3 Abelian Embeddings

Motivated by Question 1, the work [BKM22] suggested the notion of *Abelian embeddings* of distributions, defined as follows.

Definition 1.1. We say that a distribution μ on $\Sigma_1 \times \cdots \times \Sigma_k$ admits an Abelian embedding if there exist an Abelian group G and mappings $\sigma_i : \Sigma_i \rightarrow G$, $1 \leq i \leq k$ not all constant such that for all $(a_1, a_2, \dots, a_k) \in \text{supp}(\mu)$, it holds that $\sum_{i=1}^k \sigma_i(a_i) = 0_G$.

First, we note that for (1) to hold, it is necessary for μ to not admit any Abelian embedding. Indeed, suppose that μ admits an Abelian embedding, and fix G and $\sigma_1, \dots, \sigma_k$ as in Definition 1.1. Fix some non-trivial character $\chi \in \widehat{G}$ and define $f_i(x_i) = \prod_{j=1}^n \chi(\sigma_i((x_i)_j))$. Note that for each $(x_1, x_2, \dots, x_k) \in \text{supp}(\mu^{\otimes n})$, the product $f_1(x_1)f_2(x_2) \cdots f_k(x_k)$ is equal to

$$\prod_{i=1}^k \prod_{j=1}^n \chi(\sigma_i((x_i)_j)) = \prod_{j=1}^n \prod_{i=1}^k \chi(\sigma_i((x_i)_j)) = \prod_{j=1}^n \chi\left(\sum_{i=1}^k \sigma_i((x_i)_j)\right) = \prod_{j=1}^n \chi(0_G) = 1.$$

In particular, the expectation of $f_1(x_1)f_2(x_2) \cdots f_k(x_k)$ over $\mu^{\otimes n}$ is bounded away from 0 in absolute value. Yet, for large n , for each i such that σ_i is not constant the corresponding f_i is a (essentially) high-degree function.⁴ Motivated by these examples and a long-term application to approximability of CSPs on satisfiable instances, the authors of [BKM22] hypothesized that having no Abelian embeddings is also a sufficient condition. More precisely:

Conjecture 1.2. (Informal): *For a distribution μ on Σ^k , Conclusion (1) holds if and only if μ admits no Abelian embedding.*

Conjecture 1.2 was proved for $k = 3$ in [BKM23a]. Follow-up works [BKM23b, BKM24a, BKM24b, BKLM24a] further studied the case of $k = 3$ and distributions μ that do admit Abelian embeddings, and gave a characterization of functions f_1, f_2, f_3 that may exhibit such a correlation (provided that the distribution μ is pairwise-connected, which is a much milder requirement).

1.4 Our Contributions

1.4.1 Main Result

Our first result is a confirmation of Conjecture 1.2 for all k .

Theorem 1 (Main theorem). *Let k be a positive integer and let μ be a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$ that does not admit any Abelian embedding, and in which the probability of each atom is at least α . Then, for every $\varepsilon > 0$ there is $\delta := \delta(\alpha, \varepsilon) > 0$ such that for sufficiently large n , if 1-bounded functions $f_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then $\text{Stab}_{1-\delta}(f_i) \geq \delta$ for all $i = 1, 2, \dots, k$. Quantitatively, we can take

$$\delta := \exp(-\exp(\dots \exp(\varepsilon^{-O_\alpha(1)})))$$

where the number of exponentials is at most $k^{O(k)}$.

Remark. *The condition that $\text{Stab}_{1-\delta}(f_i) \leq \delta$ (see Definition 2.2) serves as a convenient proxy for the condition that the function f_i is essentially of high degree. Indeed, if $\text{Stab}_{1-\delta}(f_i) \leq \delta$ then the Fourier mass of f_i on degrees less than $\frac{1}{\delta}$ is at most $O(\delta)$. Conversely, if the Fourier mass on degrees less than $\frac{1}{\delta} \log(\frac{2}{\delta})$ is at most $\frac{\delta}{2}$, then $\text{Stab}_{1-\delta}(f_i) \leq \delta$.*

⁴The functions here are complex-valued with absolute value 1; one can appropriately take their real or imaginary parts if one insists on having real-valued functions.

Implications to CSP approximability: combining [Theorem 1](#) with the transformation from [\[BKM22\]](#) converting integrality gap instances to dictatorship tests, we have the following result:

Theorem 2. *Let $P: \Sigma^k \rightarrow \{0, 1\}$ be a predicate that satisfies the following conditions: (1) P admits no Abelian embedding, (2a) there exists an instance of Max- P -CSP that has a $(1, s)$ -integrality gap for the basic SDP relaxation, (2b) on every constraint, the local distribution in the SDP solution admits no Abelian embedding. Then, for every $\varepsilon > 0$, there is a dictatorship test for P -CSP that has perfect completeness and soundness $s + \varepsilon$.*

In other words, [Theorem 2](#) gives some evidence that for predicates with no Abelian embedding, the best possible approximation algorithm is still the natural SDP relaxation with an appropriate rounding scheme.

1.4.2 Local Inverse Theorems under Milder Assumptions

The proof of [Theorem 2](#) requires a structural result for functions f_1, \dots, f_k exhibiting k -wise correlation under a milder assumption on μ . Namely, instead of requiring that μ does not admit an Abelian embedding, we require it to be *pairwise-connected* as well as that its marginal on the first $k - 1$ coordinates, denoted by μ_{-k} , admits no Abelian embedding.

Definition 1.3. A distribution μ over $\Sigma_1 \times \dots \times \Sigma_k$ is pairwise-connected if for all $i \neq j$ the bipartite graph $(\Sigma_i \cup \Sigma_j, \text{supp}(\mu_{i,j}))$ is connected.

In this setting we have the following result:

Lemma 1.4. *Let k be a positive integer, let μ be a pairwise-connected distribution over $\Sigma_1 \times \dots \times \Sigma_k$ in which the probability of each atom is at least α , and assume that μ_{-k} admits no Abelian embeddings. Then, for every $\varepsilon > 0$ there is $\delta := \delta(\alpha, \varepsilon) > 0$ such that for all sufficiently large n , if 1-bounded functions $f_i: \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then a random restriction of f_1 correlates to a product function, i.e.,

$$\Pr_{I \sim_{1-\delta}[n], z \sim \mu_1^I} \left[\exists \{P_i: \Sigma_1 \rightarrow \mathbb{C}, \|P_i\|_\infty \leq 1\}_{i \in \bar{I}} \quad \text{with} \quad \left| \mathbb{E}_{x \sim \mu_1^{\bar{I}}} \left[(f_1)_{I \rightarrow z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geq \delta \right] \geq \delta.$$

Quantitatively, $\delta = \exp(-\exp(\dots \exp(\varepsilon^{-O_\alpha(1)})))$ where the number of exponentials is $k^{O(k)}$.

If we strengthen the assumption on μ_{-k} , and assume that it is connected, then we can improve upon the quantitative aspect of [Lemma 1.4](#), and get the following result:

Lemma 1.5. *Let k be a positive integer, let μ be a pairwise-connected distribution over $\Sigma_1 \times \dots \times \Sigma_k$ in which the probability of each atom is at least α , and assume that μ_{-k} is connected. Then, for every $\varepsilon > 0$ there is a constant $\delta := \delta(\alpha, \varepsilon) > 0$ such that for all sufficiently large n , if 1-bounded functions $f_i: \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then a random restriction of f_1 correlates to a product function, i.e.,

$$\Pr_{I \sim_{1-\delta} [n], z \sim \mu_1^I} \left[\exists \{P_i : \Sigma_1 \rightarrow \mathbb{C}, \|P_i\|_\infty \leq 1\}_{i \in \bar{I}} \text{ with } \left| \mathbb{E}_{x \sim \mu_1^{\bar{I}}} \left[(f_1)_{I \rightarrow z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geq \delta \right] \geq \delta.$$

Quantitatively, we can take $\delta = \exp(-\varepsilon^{-O_\alpha(1)})$.

1.4.3 A Global Inverse Theorem

Finally, we state and prove a global inverse theorem in the setting of [Lemma 1.4](#). By a global inverse theorem, we mean a result that makes an assertion regarding the global structure of the functions f_i 's, as opposed to only a local one after random restrictions. Indeed, combining [Lemma 1.4](#) and the restriction inverse theorem, namely [[BKLM24a](#), Theorem 9], we get the following conclusion.

Theorem 3. *Let k be a positive integer, let μ be a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$ in which the probability of each atom is at least α , and assume that μ_{-k} admits no Abelian embeddings. Then, for every $\varepsilon > 0$ there are $\delta := \delta(\alpha, \varepsilon) > 0$ and $d = d(\alpha, \varepsilon) \in \mathbb{N}$ such that for all sufficiently large n , if 1-bounded functions $f_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then there exists $L : \Sigma_1^n \rightarrow \mathbb{C}$ of degree at most d and $\|L\|_2 \leq 1$, as well as a product function $P : \Sigma_1^n \rightarrow \mathbb{C}$ of the form $P(x) = \prod_{i=1}^n P_i(x_i)$ where $|P_i(x_i)| = 1$ for all i and x , such that $|\langle f, L \cdot P \rangle| \geq \delta$.

1.5 Subsequent and Future Works

Subsequent works: in [[BKLM24b](#)], we use [Lemma 1.5](#) with $k = 4$ to give the first reasonable bounds for the density Hales-Jewett theorem on $\{0, 1, 2\}^n$. The density Hales-Jewett problem refers to the problem of determining the maximum density of a subset $S \subseteq \{0, 1, \dots, k-1\}^n$ that avoids all *combinatorial lines* of length k . Here, we say that a k -tuple $x^{(1)}, \dots, x^{(k)} \in S$ forms a combinatorial line of length k if not all $x^{(j)}$ are equal, and for each coordinate $i = 1, 2, \dots, n$,

$$(x_i^{(1)}, \dots, x_i^{(k)}) \in \{(0, \dots, 0), (1, \dots, 1), \dots, (k-1, \dots, k-1), (0, 1, 2, \dots, k-1)\}.$$

The density Hales-Jewett theorem, proved by Furstenberg and Katznelson [[FK89](#), [FK91](#)], asserts that for every $k \in \mathbb{N}$ and $\delta > 0$, for sufficiently large $n \geq n_0(k, \delta)$, a subset $S \subseteq \{0, 1, \dots, k-1\}^n$ of density at least δ must contain a combinatorial line of length k . The proof of Furstenberg and Katznelson is ergodic theoretic and gives no quantitative bounds on n_0 . The Polymath project [[Pol12](#)] provided an elementary proof that does give quantitative bounds. For $k = 3$, their result asserts that if $3^{-n}|S| \geq \Omega((\log^* n)^{-1/2})$, then S contains a combinatorial line. In [[BKLM24b](#)], we improve upon this result by showing that if $S \subseteq \{0, 1, 2\}^n$ is such that $3^{-n}|S| \geq \Omega((\log \log \log \log n)^{-c})$ (for some absolute constant $c > 0$), then S contains a combinatorial line.

Remark. [Lemma 1.4](#) is already good enough to get an effective bounds for the density Hales-Jewett theorem in $[3]^n$, except that the bound would have five logs as opposed to four.

Future works: the results of this paper can be seen as an extension of the results in [BKM23a] to the setting of k -ary distributions with no Abelian embeddings. In future works, we plan to investigate the class of k -ary distributions that do admit Abelian embeddings, and prove results along the lines of [BKM24a, BKLM24a] characterizing k -tuples of functions that achieve non-trivial k -wise correlation. This scenario already includes within it, (as special cases) the Gowers' uniformity norms over finite fields [Gow01, GT08, BTZ10, TZ12, GM17], and such inverse theorems may lead to further progress in related problems in additive combinatorics.

2 Preliminaries

Notations: We denote $[n] = \{1, \dots, n\}$, and for $\alpha \in (0, 1)$ we denote by $I \sim_{1-\alpha} [n]$ the distribution over subsets of $[n]$ in which each element $i \in [n]$ is included with probability $1 - \alpha$. For a real number x we write $\|x\|_{\mathbb{R}/\mathbb{Z}} := \min_{z \in \mathbb{Z}} |x - z|$.

2.1 Noise Operators

Let (Σ, ν) be a probability space. Define the inner product on this space by $\langle f, g \rangle_\nu := \mathbb{E}_{x \in \nu} [f(x) \overline{g(x)}]$. We need the following definition of the noise operator.

Definition 2.1. Let Σ be a finite alphabet, and ν be a measure on Σ . For a parameter $\rho \in [0, 1]$, we define the ρ -correlated distribution with respect to ν as follows. For any $y \in \Sigma$, the distribution of inputs that are ρ -correlated with y is denoted by $y' \sim T_\rho y$ and is defined by taking $y' = y$ with probability ρ , and otherwise sampling $y' \sim \nu$.

We also view T_ρ as an operator on functions, mapping $L^2(\Sigma, \nu)$ to $L^2(\Sigma, \nu)$ defined as

$$(T_\rho g)(y) = \mathbb{E}_{y' \sim T_\rho y} [g(y')].$$

We then tensorize this operator, i.e., consider $T_\rho^{\otimes n}$ which acts on functions on n -variables. When clear from context, we drop the $\otimes n$ superscript from notation. The stability of a function $f : \Sigma^n \rightarrow \mathbb{C}$ with respect to the noise parameter ρ is defined as follows.

Definition 2.2. $\text{Stab}_\rho^\nu(g) := \langle g, T_\rho g \rangle_{\nu^{\otimes n}}$. We often drop the superscript ν from $\text{Stab}_\rho^\nu(g)$ when it is clear from the context.

2.2 Inverse Theorems

In this section, we state the inverse theorems that we need to prove our results. Our proof of [Theorem 1](#) is by induction on k , and the base case $k = 3$ was proven in [BKM23a], as follows:

Theorem 4. *Let μ be a distribution over $\Sigma_1 \times \Sigma_2 \times \Sigma_3$ that does not admit any Abelian embedding, and in which the probability of each atom is at least α . Then, for all $\varepsilon > 0$ there is a constant $\delta := \delta(\alpha, \varepsilon)$ such that for all sufficiently large n , if 1-bounded functions $f_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

$$\left| \mathbb{E}_{(x_1, x_2, x_3) \sim \mu^{\otimes n}} \left[\prod_{i=1}^3 f_i(x_i) \right] \right| \geq \varepsilon,$$

then $\text{Stab}_{1-\delta}(f_i) \geq \delta$ for all $i = 1, 2, 3$. Quantitatively, $\delta(\varepsilon) \geq \exp(-\varepsilon^{-O_\alpha(1)})$.

We will also need the following local inverse theorem from [BKL24a]. Here, the only assumption on the distribution μ is that it is pairwise-connected.

Theorem 5. *Let μ be a pairwise-connected distribution over $\Sigma \times \Gamma \times \Phi$, and assume that the probability of each atom is at least α . If 1-bounded functions $f : \Sigma^n \rightarrow \mathbb{C}$, $g : \Gamma^n \rightarrow \mathbb{C}$, $h : \Phi^n \rightarrow \mathbb{C}$ satisfy that*

$$\left| \mathbb{E}_{(x,y,z) \sim \mu^{\otimes n}} [f(x)g(y)h(z)] \right| \geq \varepsilon,$$

then there is a constant $\delta := \delta(\alpha, \varepsilon) > 0$ and distribution ν such that $\mu = (1 - \delta)\nu + \delta U$, where U is uniform over Σ , such that:

$$\Pr_{I \sim_{1-\delta} [n], z \sim \nu^I} \left[\exists \{P_i : \Sigma \rightarrow \mathbb{C}, \|P_i\|_\infty \leq 1\}_{i \in \bar{I}} \quad \text{with} \quad \left| \mathbb{E}_{x \sim \Sigma^{\bar{I}}} \left[f_{I \rightarrow z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geq \delta \right] \geq \delta.$$

Quantitatively, $\delta(\varepsilon) \geq \exp(-\varepsilon^{-O_\alpha(1)})$.

Finally, in order to get improved bounds in Lemma 1.5, instead of Theorem 4, we use the following rest of Mossel [Mos10].

Lemma 2.3 (Lemma 6.2 from [Mos10]). *Let k be a positive integer and μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ such that μ is connected and the probability of each atom is at least α . Then, for all $\varepsilon > 0$ there is a constant $\delta := \delta(\alpha, \varepsilon)$ such that for all sufficiently large n , if 1-bounded functions $f_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then $\text{Stab}_{1-\delta}(f_i) \geq \delta$ for all $i \in [k]$. Quantitatively, $\delta(\varepsilon, \alpha) \geq (\varepsilon \cdot \alpha)^{O(1)}$.

3 k -CSPs Without Abelian Embeddings

In this section we prove Theorem 1. The proof proceeds by induction on k , and critically uses Theorem 4 and Theorem 5 to perform the inductive step.

The structure of the proof and notations: the base case of the proof, namely the case of $k = 3$ in Theorem 1, is given by Theorem 4. We will show that for each k , Theorem 1 for k implies Lemma 1.4 for $k+1$, which in turn implies Theorem 1 for $k+1$. Thus, we denote by $\delta_k(\alpha, \varepsilon)$ the parameter of Theorem 1 for ε , and by $\tau_k(\alpha, \varepsilon)$ the parameter δ from Lemma 1.4 for ε . With this, we have that $\delta_3(\alpha, \varepsilon) = \exp(-\varepsilon^{-O_\alpha(1)})$, and we will show bounds for $\tau_k(\alpha, \varepsilon)$ in terms of $\delta_k(\alpha, \varepsilon)$, and bounds for $\delta_{k+1}(\alpha, \varepsilon)$ in terms of $\tau_k(\alpha, \varepsilon)$. We often suppress the dependency on α to simplify notation.

3.1 Preliminary Observations

We start with the following two simple observations.

Observation 1. If the support of μ admits no Abelian embedding, then μ is pairwise-connected.

Proof. Let μ_{ij} be the restriction of μ to the i, j coordinates. If $(\Sigma_i \cup \Sigma_j, \text{supp}(\mu_{ij}))$ is not connected, then there is a non-trivial partition $\Sigma_i = \Sigma'_i \cup \Sigma''_i$ and $\Sigma_j = \Sigma'_j \cup \Sigma''_j$ such that $\text{supp}(\mu_{ij}) \subseteq (\Sigma'_i \times \Sigma'_j) \cup (\Sigma''_i \times \Sigma''_j)$. Now let $\sigma_t \equiv 0$ for $t \neq i, j$, $\sigma_i(x) = 1_{x \in \Sigma'_i}$, and $\sigma_j(x) = -1_{x \in \Sigma'_j}$. This is clearly an Abelian embedding, and at least one of σ_i or σ_j is nonconstant. \square

Observation 2. If the support of a distribution μ on $\Sigma_1 \times \cdots \times \Sigma_k$ has no Abelian embedding, then the restriction of μ to any $k - 1$ coordinates also has no Abelian embedding.

Proof. Let μ' be the restriction of μ to the coordinates $1, \dots, k - 1$. If μ' has Abelian embeddings $\sigma'_1, \dots, \sigma'_{k-1}$ then μ has the Abelian embeddings $\sigma_k \equiv 0$ and $\sigma_i \equiv \sigma'_i$ for $i = 1, \dots, k - 1$. \square

3.2 Reducing to a $(k - 1)$ -ary Instance

The goal of this subsection and of Section 3.3 is to show that Theorem 1 for $k - 1$ implies Lemma 1.4 for k . The first step is to apply Cauchy-Schwarz to remove the x_k variable, with the goal of reducing to a $(k - 1)$ -ary instance. To do this, we need to set up a bit of notation. Recall that μ_{-k} denoted the marginal distribution of μ on $\Sigma_1 \times \cdots \times \Sigma_{k-1}$. Let $\mu_{-k, -k}$ be the following distribution on $(\Sigma_1 \times \cdots \times \Sigma_{k-1}) \times (\Sigma_1 \times \cdots \times \Sigma_{k-1})$:

- Sample x_k from μ_k .
- Sample (x_1, \dots, x_{k-1}) and (x'_1, \dots, x'_{k-1}) from μ conditioned on x_k .
- $\mu_{-k, -k}$ contains $((x_1, \dots, x_{k-1}), (x'_1, \dots, x'_{k-1}))$.

Note that $\text{supp}(\mu_{-k, -k})$ contains $((x_1, \dots, x_{k-1}), (x_1, \dots, x_{k-1}))$ for any $(x_1, \dots, x_{k-1}) \in \mu_{-k}$. In other words, $\mu_{-k, -k}$ essentially contains μ_{-k} as a submeasure.

Now, using the hypothesis of Lemma 1.4 gives us:

$$\varepsilon^2 \leq \left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^k f_i(x_i) \right] \right|^2 \leq \mathbb{E}_{\substack{(x_1, \dots, x_{k-1}) \\ (x'_1, \dots, x'_{k-1}) \sim \mu_{-k, -k}^{\otimes n}}} \left[\prod_{i=1}^{k-1} f_i(x_i) \overline{f_i(x'_i)} \right].$$

By abuse of notation, identify between μ_{-k} and the measure on $(\Sigma_1 \times \cdots \times \Sigma_{k-1}) \times (\Sigma_1 \times \cdots \times \Sigma_{k-1})$ consisting of points

$$((x_1, \dots, x_{k-1}), (x_1, \dots, x_{k-1})),$$

where (x_1, \dots, x_{k-1}) is distributed as in μ_{-k} . We may write $\mu_{-k, -k} = \alpha^2 \mu_{-k} + (1 - \alpha^2) \nu$ where ν is a distribution. Applying this random restriction to the above equation gives:

$$\varepsilon^2 \leq \mathbb{E}_{I \sim_{1-\alpha^2} [n]} \mathbb{E}_{\substack{(z_1, \dots, z_{k-1}) \\ (z'_1, \dots, z'_{k-1}) \sim \nu^{\otimes I}}} \mathbb{E}_{(x_1, \dots, x_{k-1}) \sim \mu_{-k}^{\otimes \bar{I}}} \left[\prod_{i=1}^{k-1} (f_i)_{I \rightarrow z_i}(x_i) \overline{(f_i)_{I \rightarrow z'_i}(x_i)} \right]. \quad (2)$$

By the hypothesis, μ_{-k} has no Abelian embedding, and thus if

$$\left| \mathbb{E}_{(x_1, \dots, x_{k-1}) \sim \mu_{-k}^{\otimes \bar{I}}} \left[\prod_{i=1}^{k-1} (f_i)_{I \rightarrow z_i}(x_i) \overline{(f_i)_{I \rightarrow z'_i}(x_i)} \right] \right| \geq \varepsilon^2/2, \quad (3)$$

by induction there is a constant $\delta_{k-1} = \delta_{k-1}(\varepsilon^2/2)$ such that

$$\text{Stab}_{1-\delta_{k-1}} \left((f_1)_{I \rightarrow z_1} \overline{(f_1)_{I \rightarrow z'_1}} \right) \geq \delta_{k-1}.$$

Combining this with (2) gives:

$$\mathbb{E}_{I \sim_{1-\alpha^2} [n]} \mathbb{E}_{\substack{(z_1, \dots, z_{k-1}) \\ (z'_1, \dots, z'_{k-1})} \sim \nu^{\otimes I}} \left[\text{Stab}_{1-\delta_{k-1}} \left((f_1)_{I \rightarrow z_1} \overline{(f_1)_{I \rightarrow z'_1}} \right) \right] \geq \varepsilon^2 \delta_{k-1} / 2. \quad (4)$$

3.3 Applying the 3-ary Inverse Theorem

Let $\Sigma := \Sigma_1$. Now we will interpret (4) as a 3-wise correlation over a pairwise-connected distribution, and then apply [Theorem 5](#). Let $\Sigma^+ = (\Sigma \times \Sigma) \cup \{\star\}$, for some symbol \star .

Let ξ be the following distribution over $\Sigma \times \Sigma \times \Sigma^+$:

- With probability $1 - \alpha^2$, sample $(x_1, x'_1) \sim \nu_1$, and include $(x_1, x'_1, (x_1, x'_1))$ in ξ .
- Otherwise, let $x \sim \mu_1$. With probability $1 - \delta_{k-1}$ include $(x, x, (x, x))$ in ξ , and with probability δ_{k-1} include (x, x, \star) in ξ .

Finally, given $x^+ \in \Sigma^+$, define the following way of sampling $(x, x') \in \Sigma^n \times \Sigma^n$, which we denote as $(x, x') \sim_{\star} x^+$. If $x_i^+ \neq \star$, let x_i, x'_i be the first, second coordinates of x_i^+ respectively. Otherwise if $x_i^+ = \star$, sample $x \sim \mu_1$ and set $x_i = x'_i = x$. This lets us define the function $g : (\Sigma^+)^n \rightarrow \mathbb{C}$ as

$$g(x^+) := \mathbb{E}_{(x, x') \sim_{\star} x^+} \left[f_1(x) \overline{f_1(x')} \right].$$

At this point, let us collect two useful observations:

Observation 3. The distribution ξ is pairwise-connected.

Proof. Indeed, the 13-coordinates are pairwise-connected because every $x \in \Sigma$ is connected to $\star \in \Sigma^+$ because (x, x, \star) is in $\text{supp}(\xi)$ for all $x \in \Sigma$. Similarly, the 23-coordinates are pairwise-connected. For the 12-coordinates, note that (x, x') is in the support if there is some $x_k \in \Sigma_k$ such that $(x, x_k), (x', x_k) \in \mu_{1k}$. Thus, the claim follows because the 1k-coordinates were connected in μ by the hypothesis of [Lemma 1.4](#). \square

Observation 4. It holds that

$$\begin{aligned} \mathbb{E}_{(x, x', x^+) \sim \xi^{\otimes n}} \left[f_1(x) \overline{f_1(x')} g(x^+) \right] &= \mathbb{E}_{I \sim_{1-\alpha} [n]} \mathbb{E}_{\substack{(z_1, \dots, z_{k-1}) \\ (z'_1, \dots, z'_{k-1})} \sim \nu^{\otimes I}} \left[\text{Stab}_{1-\delta_{k-1}} \left((f_1)_{I \rightarrow z_1} \overline{(f_1)_{I \rightarrow z'_1}} \right) \right] \\ &\geq \varepsilon^2 \delta_{k-1} / 2. \end{aligned}$$

Proof. The equality is by the definition of ξ and g . The inequality is exactly (4). \square

Given these, we can establish [Lemma 1.4](#) using [Theorem 5](#).

Proof that [Theorem 1](#) with $k - 1$ implies [Lemma 1.4](#) with k . Note that all atoms in ξ have mass at least $\alpha^2 \delta_{k-1}$, and that ξ is pairwise-connected by [Observation 3](#). Thus, we may write $\xi = \zeta \xi' + (1 - \zeta) \xi''$ where ξ', ξ'' are distributions, $\zeta \geq \alpha^2 \delta_{k-1}$, the probability of each atom in ξ' is at least α^2 and ξ' has the same support as ξ . Applying random restrictions, we get that with probability at least $\varepsilon^2 \delta_{k-1} / 4$, choosing $J \sim_{1-\zeta} I$ and $(a, b, c) \sim \xi''^J$ we have that

$$\left| \mathbb{E}_{(x, x', x^+) \sim \xi'^J} \left[(f_1)_{J \rightarrow a}(x) \overline{(f_1)_{J \rightarrow b}(x')} (g)_{J \rightarrow c}(x^+) \right] \right| \geq \varepsilon^2 \delta_{k-1} / 4.$$

For each such J and (a, b, c) , [Theorem 5](#) gives that:

$$\Pr_{I \sim_{1-\delta} J, z \sim \nu_1^I} \left[\exists \{P_i : \Sigma_1 \rightarrow \mathbb{C}, \|P_i\|_\infty \leq 1\}_{i \in \bar{I}} \quad \text{with} \quad \left| \mathbb{E}_{x \sim \Sigma_1^{\bar{I}}} \left[(f_1)_{J \rightarrow a, I \rightarrow z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geq \delta \right] \geq \delta,$$

where $\delta \geq \exp(-(\varepsilon \delta_{k-1})^{-O_\alpha(1)})$. Combining the two random restrictions, we see that we may write $\mu_1 = (1 - \delta\zeta)\nu + \delta\zeta U$ and get that

$$\Pr_{I \sim_{1-\delta\zeta} [n], z \sim \nu_1^I} \left[\exists \{P_i : \Sigma_1 \rightarrow \mathbb{C}, \|P_i\|_\infty \leq 1\}_{i \in \bar{I}} \quad \text{with} \quad \left| \mathbb{E}_{x \sim \Sigma_1^{\bar{I}}} \left[(f_1)_{I \rightarrow z}(x) \prod_{i \in \bar{I}} P_i(x_i) \right] \right| \geq \delta \right] \geq \delta,$$

By applying a further random restriction we can convert ν and U back into μ_1 at the cost of decreasing δ by a constant. This finishes the inductive step, establishing that

$$\tau_k(\varepsilon) \geq \Omega(\varepsilon^2 \delta_{k-1}(\varepsilon^2/2)) \cdot \exp(-(\varepsilon \delta_{k-1}(\varepsilon^2/2))^{-O_\alpha(1)}) \geq \exp(-\delta_{k-1}(\varepsilon^2/2)^{-O_\alpha(1)}). \quad (5)$$

In [Section 3.4](#) we will establish a recursive relation for δ_{k-1} , and then in [Section 3.5](#) we conclude a quantitative estimate for τ_k . \square

We get [Lemma 1.5](#) using a similar argument.

Proof of [Lemma 1.5](#). The proof of this lemma is along the lines of the proof of the [Lemma 1.4](#). The only change is that, instead of relying on the inductive hypothesis to go from (2) to (4), we use [Lemma 2.3](#) where $\delta_{k-1} \geq \varepsilon^{O_\alpha(1)}$. The conclusion then follows from [Observation 4](#) and applying [Theorem 5](#). \square

3.4 Proof of [Theorem 1](#)

Fix functions f_1, \dots, f_k as in the statement of [Theorem 1](#). The goal of this section is to replace each f_i by a product function (potentially after random restriction) by applying [Lemma 1.4](#). Then, by using the fact that μ has no Abelian embedding, we will conclude that the product functions must be nearly constant, which implies that f_i correlates to a constant after random restriction. This is essentially what [Theorem 1](#) claims.

More precisely, the following lemma allows us to replace f_k by a product function:

Lemma 3.1. *Let k be a positive integer and let μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ that has no Abelian embeddings and in which the probability of each atom is at least α . Then, for all $\varepsilon > 0$, if 1-bounded functions $f_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then with probability at least $\tau_k/2 = \tau_k(\alpha, \varepsilon^2)/2$ over $I \sim_{1-\tau_k} [n]$ and $(z_1, \dots, z_k) \sim \mu^I$, there is a 1-bounded product function $P_k : \Sigma_k^{\bar{I}} \rightarrow \mathbb{C}$ satisfying

$$\left| \mathbb{E}_{x \sim \mu^{\otimes \bar{I}}} \left[\prod_{i=1}^{k-1} (f_i)_{I \rightarrow z_i}(x_i) P_k(x_k) \right] \right| \geq \frac{\tau_k}{2}.$$

Proof. Define the function $\tilde{f}_k : \Sigma_k^n \rightarrow \mathbb{C}$ as

$$\tilde{f}_k(x) := \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^{k-1} f_i(x_i) \mid x_k = x \right].$$

Note that \tilde{f}_k is 1-bounded, and by the Cauchy-Schwarz inequality

$$\begin{aligned} \varepsilon^2 &\leq \left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^k f_i(x_i) \right] \right|^2 = \left| \mathbb{E}_{x \sim \mu_k^{\otimes n}} \left[\tilde{f}_k(x) f_k(x) \right] \right|^2 \leq \|\tilde{f}_k\|_2^2 \|f_k\|_2^2 \\ &\leq \|\tilde{f}_k\|_2^2 = \left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^{k-1} f_i(x_i) \overline{\tilde{f}_k(x_k)} \right] \right|. \end{aligned}$$

Note that by [Observation 1](#) and [Observation 2](#), the hypotheses of [Lemma 1.4](#) hold. Thus, [Lemma 1.4](#) gives that for some $\tau_k := \tau_k(\alpha, \varepsilon^2)$, with probability at least τ_k over $I \sim_{1-\tau_k} [n]$ and $z \sim \mu_k^{\otimes I}$ that there is a 1-bounded product function $P_k : \Sigma_k^{\bar{I}} \rightarrow \mathbb{C}$ with

$$\tau_k \leq \left| \mathbb{E}_{x \sim \mu_k^{\otimes \bar{I}}} \left[\overline{(\tilde{f}_k)_{I \rightarrow z}(x)} P_k(x) \right] \right| = \mathbb{E}_{(z_1, \dots, z_k) \sim \mu^{\otimes I} : z_k = z} \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes \bar{I}}} \left[\prod_{i=1}^{k-1} (f_i)_{I \rightarrow z_i}(x_i) \overline{P_k(x_k)} \right].$$

This implies the desired conclusion by an averaging argument. \square

Applying [Lemma 3.1](#) a total of $k - 1$ times gives the following.

Corollary 3.2. *Let k be a positive integer and let μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ that has no Abelian embeddings and in which the probability of each atom is at least α . Then, for all $\varepsilon > 0$ there is a constant $\delta := \delta(\alpha, \varepsilon)$ such that for all sufficiently large n , if 1-bounded functions $f_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

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

then with probability at least δ over $I \sim_{1-\delta} [n]$ and $z \sim \mu_1^I$, there are 1-bounded product functions $P_i : \Sigma_i^{\bar{I}} \rightarrow \mathbb{C}$ for $i = 2, \dots, k$ such that:

$$\left| \mathbb{E}_{x \sim \mu_1^{\otimes n}} \left[(f_1)_{I \rightarrow z}(x) \prod_{i=2}^k P_i(x_i) \right] \right| \geq \delta.$$

Quantitatively, we can take $\delta := \tau_k(\dots(\tau_k(\varepsilon^2))^2)$ where the number of applications of τ_k is $k - 1$.

We want to use this to argue that $(f_1)_{I \rightarrow z}$ has large noise stability. Towards this, we need a somewhat technical statement about when product functions can correlate under distributions with no Abelian embedding. We start with the one-dimensional case.

Lemma 3.3. *Let k be a positive integer and let μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ that has no Abelian embeddings and in which the probability of each atom is at least α . For all $\eta > 0$ and 1-bounded functions $P_i : \Sigma_i \rightarrow \mathbb{C}$ satisfying*

$$\left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu} \left[\prod_{i=1}^k P_i(x_i) \right] \right| \geq 1 - \eta,$$

it holds that $|P_1(x)| \geq 1 - O_\alpha(\eta)$ and $|P_1(x) - P_1(y)| \leq O_\alpha(\sqrt{\eta})$ for all $x, y \in \Sigma_1$.

Proof. We think of α as a constant throughout. We start by reinterpreting what it means for μ to have no Abelian embedding linear-algebraically. If μ has no Abelian embedding, then there are no functions $\sigma_i : \Sigma_i \rightarrow \mathbb{R}$ such that $\sum_{i=1}^k \sigma_i(x_i) \in \mathbb{Z}$ for all $(x_1, \dots, x_k) \in \text{supp}(\mu)$, and $\pi \circ \sigma_i$ are nonconstant, where $\pi : \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$ is the projection map modulo 1. Let $(x_1^*, \dots, x_k^*) \in \text{supp}(\mu)$ be an arbitrary point. We may assume that $\sigma_i(x_i^*) = 0$ for all $i = 1, \dots, k$ by shifting. Let $S = \sum_{i=1}^k |\Sigma_i| - 1$, and denote $\sigma \in \mathbb{R}^S$ as the vector with entries $\sigma_i(x_i)$ for $i = 1, \dots, k$ and $x_i \neq x_i^*$. For a point $x = (x_1, \dots, x_k) \in \text{supp}(\mu)$, let $a_x \in \mathbb{Z}^S$ be the vector such that $\langle a_x, \sigma \rangle = \sum_{i=1}^k \sigma_i(x_i)$. Then, rephrasing the above condition, we get that μ has no Abelian embedding if and only if $\langle a_x, \sigma \rangle \in \mathbb{Z}$ for all $x \in \text{supp}(\mu)$ implies that $\sigma \in \mathbb{Z}^S$. Consider the lattice $L = \text{span}_{\mathbb{Z}}(a_x)$ and let $a^{(1)}, \dots, a^{(S)} \in L$ be a spanning set for L . Let $A \in \mathbb{Z}^{S \times S}$ be the matrix with rows $a^{(1)}, \dots, a^{(S)}$. Then, we know that μ has no Abelian embedding if and only if $A\sigma \in \mathbb{Z}^S$ implies that $\sigma \in \mathbb{Z}^S$. Thus, A^{-1} must have integer entries.

Now, we proceed to the main argument. We may assume that η is sufficiently small, because obviously $|P_1(x) - P_1(y)| \leq 2$ for all $x, y \in \Sigma_1$. We may assume that

$$\mathbb{E}_{(x_1, \dots, x_k) \sim \mu} \left[\prod_{i=1}^k P_i(x_i) \right]$$

is a real number, and thus the real part of $\prod_{i=1}^k P_i(x_i)$ is at least $1 - O(\eta)$ for all $(x_1, \dots, x_k) \in \text{supp}(\mu)$. Because P_j are 1-bounded, this implies that $|P_j(x_j)| \geq 1 - O(\eta)$ for all $x_j \in \Sigma_j$. If $\sigma_j(x_j)$ are such that $P_j(x_j)/|P_j(x_j)| = \exp(2\pi i \sigma_j(x_j))$, we further know that

$$\left\| \sum_{j=1}^k \sigma_j(x_j) \right\|_{\mathbb{R}/\mathbb{Z}} \leq O(\sqrt{\eta})$$

for all $(x_1, \dots, x_k) \in \text{supp}(\mu)$. For a_x defined as above, there are integers k_x such that $|\langle a_x, \sigma \rangle - k_x| \leq O(\sqrt{\eta})$ for $x \in \text{supp}(\mu)$. Because $a^{(1)}, \dots, a^{(S)}$ are finite integer combinations of a_x , we know that there are integers $k^{(i)}$ such that $|\langle a^{(i)}, \sigma \rangle - k^{(i)}| \leq O(\sqrt{\eta})$ for $i = 1, \dots, S$. Because A^{-1} has integral entries, we conclude that $\|\sigma\|_{\mathbb{R}/\mathbb{Z}} \leq O(\sqrt{\eta})$.

This implies the desired conclusion, because

$$\begin{aligned} |P_1(x) - P_1(y)| &\leq O(\eta) + |P_1(x)/|P_1(x)| - P_1(y)/|P_1(y)|| \\ &= O(\eta) + |\exp(2\pi i \sigma_1(x)) - \exp(2\pi i \sigma_1(y))| \leq O(\sqrt{\eta}), \end{aligned}$$

because $\|\sigma_1(x) - \sigma_1(y)\|_{\mathbb{R}/\mathbb{Z}} \leq O(\sqrt{\eta})$. □

Now we state the n -dimensional case that we need.

Lemma 3.4. *Let k be a positive integer and μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ whose support has no Abelian embedding and in which the probability of each atom is at least α . If 1-bounded product functions $P_i : \Sigma_i^n \rightarrow \mathbb{C}$ satisfy*

$$\left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[\prod_{i=1}^k P_i(x_i) \right] \right| \geq \delta,$$

then for any $\gamma \in (0, 1)$ it holds that

$$\mathbb{E}_{\substack{x \sim \mu_1^{\otimes n} \\ y \sim_{1-\gamma} x}} [|P_1(x) - P_1(y)|^2] \leq O_\alpha(\gamma \log(1/\delta)).$$

Proof. We think of α as a constant throughout. Let $P_i(x) = \prod_{j=1}^n P_i^{(j)}(x_j)$ for 1-bounded functions $P_i^{(j)} : \Sigma_i \rightarrow \mathbb{C}$. Let η_j be such that

$$\left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu} \left[\prod_{i=1}^k P_i^{(j)}(x_i) \right] \right| = 1 - \eta_j. \quad (6)$$

Thus, $\prod_{j=1}^n (1 - \eta_j) = \delta$ and hence $\sum_{j=1}^n \eta_j \leq O(\log(1/\delta))$.

The following equality holds for all $x, y \in \Sigma_1^n$:

$$\begin{aligned} & \left| \prod_{j=1}^n P_1^{(j)}(x_j) - \prod_{j=1}^n P_1^{(j)}(y_j) \right|^2 \\ &= \prod_{j=1}^n |P_1^{(j)}(x_j)|^2 + \prod_{j=1}^n |P_1^{(j)}(y_j)|^2 - \prod_{j=1}^n P_1^{(j)}(x_j) \overline{P_1^{(j)}(y_j)} - \prod_{j=1}^n \overline{P_1^{(j)}(x_j)} P_1^{(j)}(y_j). \end{aligned}$$

Note that for

$$\mathbb{E}_{x \sim \mu_1, y \sim_{1-\gamma} x} P_1^{(j)}(x) \overline{P_1^{(j)}(y)}$$

is a real number, and thus equals

$$\begin{aligned} \mathbb{E}_{x \sim \mu_1, y \sim_{1-\gamma} x} P_1^{(j)}(x) \overline{P_1^{(j)}(y)} &= \|P_1^{(j)}\|_2^2 - \frac{1}{2} \mathbb{E}_{x \sim \mu_1, y \sim_{1-\gamma} x} |P_1^{(j)}(x) - P_1^{(j)}(y)|^2 \\ &\geq \|P_1^{(j)}\|_2^2 - O(\gamma \eta_j) \\ &\geq \|P_1^{(j)}\|_2^2 (1 - O(\gamma \eta_j)), \end{aligned} \quad (7)$$

where we used [Lemma 3.3](#). Thus

$$\begin{aligned} \mathbb{E}_{\substack{x \sim \mu_1^{\otimes n} \\ y \sim_{1-\gamma} x}} [|P_1(x) - P_1(y)|^2] &= 2 \prod_{j=1}^n \|P_1^{(j)}\|_2^2 - 2 \prod_{j=1}^n \mathbb{E}_{\substack{x \sim \mu_1 \\ y \sim_{1-\gamma} x}} [P_1^{(j)}(x) \overline{P_1^{(j)}(y)}] \\ &\leq 2 \prod_{j=1}^n \|P_1^{(j)}\|_2^2 \left(1 - \prod_{j=1}^n (1 - O(\gamma \eta_j)) \right) \\ &\leq O(\gamma \log(1/\delta)), \end{aligned}$$

where we have used (7). □

With [Lemma 3.4](#) in hand, we now show that a function satisfying the conclusion of [Corollary 3.2](#) must be noise stable.

Lemma 3.5. *Let k be a positive integer and let μ be a distribution over $\Sigma_1 \times \dots \times \Sigma_k$ that has no Abelian embeddings and in which the probability of each atom is at least α . If 1-bounded function $f : \Sigma_1^n \rightarrow \mathbb{C}$ and product functions $P_i : \Sigma_i^n \rightarrow \mathbb{C}$ for $i = 2, \dots, k$ satisfy*

$$\left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[f(x_1) \prod_{i=2}^k P_i(x_i) \right] \right| \geq \delta$$

then for $\gamma \leq c\delta^2 / \log(1/\delta)$ for sufficiently small $c = c(\alpha) > 0$, it holds that $\text{Stab}_{1-\gamma}(f) \geq \delta^2/4$.

Proof. Define the product function $\tilde{P} : \Sigma_1^n \rightarrow \mathbb{C}$ as

$$\tilde{P}(x) := \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n} : x_1 = x} \left[\prod_{i=2}^k P_i(x_i) \right].$$

Then the following inequalities hold:

$$\delta \leq \left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[f(x_1) \prod_{i=2}^k P_i(x_i) \right] \right| = \left| \mathbb{E}_{x \sim \mu_1^{\otimes n}} \left[f(x) \tilde{P}(x) \right] \right|, \quad (8)$$

and

$$\left| \mathbb{E}_{(x_1, \dots, x_k) \sim \mu^{\otimes n}} \left[\tilde{P}(x_1) \prod_{i=2}^k P_i(x_i) \right] \right| \geq \delta^2.$$

The latter follows by Cauchy-Schwarz (see the proof of [Lemma 3.1](#)). By [Lemma 3.4](#), we know that for $\gamma \leq c\delta^2/\log(1/\delta)$

$$\begin{aligned} \left| \mathbb{E}_{x \sim \mu_1^{\otimes n}, y \sim_{1-\gamma} x} \left[f(x) (\tilde{P}(x) - \tilde{P}(y)) \right] \right| &\leq \mathbb{E}_{x \sim \mu_1^{\otimes n}, y \sim_{1-\gamma} x} \left[|\tilde{P}(x) - \tilde{P}(y)| \right] \\ &\leq \mathbb{E}_{x \sim \mu_1^{\otimes n}, y \sim_{1-\gamma} x} \left[|\tilde{P}(x) - \tilde{P}(y)|^2 \right]^{1/2} \leq O(\sqrt{\gamma \log(1/\delta)}) \leq \delta/2. \end{aligned}$$

Combining this with (8) gives

$$\begin{aligned} \delta/2 &\leq \left| \mathbb{E}_{x \sim \mu_1^{\otimes n}, y \sim_{1-\gamma} x} \left[f(x) \tilde{P}(y) \right] \right| = \left| \mathbb{E}_{x \sim \mu_1^{\otimes n}, y \sim_{1-\gamma} x} \left[f(y) \tilde{P}(x) \right] \right| \\ &\leq \mathbb{E}_{x \sim \mu_1^{\otimes n}} \left| \mathbb{E}_{y \sim_{1-\gamma} x} [f(y)] \right|, \end{aligned}$$

where the equality uses symmetry of x, y and the inequality is because \tilde{P} is 1-bounded. One final application of Cauchy-Schwarz gives that $\text{Stab}_{1-\gamma}(f) \geq \delta^2/4$ as desired. \square

[Theorem 1](#) now readily follows by combining [Corollary 3.2](#) and [Lemma 3.5](#).

Proof of Theorem 1. Applying [Corollary 3.2](#) and [Lemma 3.5](#) gives that for some δ depending on ε ,

$$\text{Stab}_{1-\gamma\delta}(f_1) = \mathbb{E}_{I \sim_{1-\delta}[n], z \sim \mu_1^I} \left[\text{Stab}_{1-\gamma}((f_1)_{I \rightarrow z}) \right] \geq \delta^3/4.$$

for $\gamma = c\delta^2/\log(1/\delta)$. This completes the induction, showing that

$$\delta_k(\varepsilon) \geq c \cdot \tau_k(\tau_k(\dots(\tau_k(\varepsilon^2)^2)^2)^4), \quad (9)$$

where the number of compositions of τ_k is $k-1$. \square

3.5 The Quantitative Bounds for [Lemma 1.4](#) and [Theorem 1](#)

Let $H(k)$ be the number of exponentials needed in [Theorem 1](#) with k . [Theorem 4](#) gives us that $H(3) \leq 1$. For any $k \geq 3$, (5) implies that [Lemma 1.4](#) for k holds with $H(k-1) + 1$ exponentials, and then (5) implies [Theorem 1](#) for k holds with $(k-1)(H(k-1) + 1)$ exponentials. We conclude that $H(k) \leq (k-1)(H(k-1) + 1)$, and solving gives

$$H(k) \leq (k-1) + (k-1)(k-2) + (k-1)(k-2)(k-3) + \dots + (k-1)(k-2)\dots 3 \leq k^k.$$

References

- [ALM⁺98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998. (Preliminary version in *33rd FOCS*, 1992).
- [AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998. (Preliminary version in *33rd FOCS*, 1992).
- [BK21] Amey Bhangale and Subhash Khot. Optimal Inapproximability of Satisfiable k-LIN over Non-Abelian Groups. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, page 1615–1628, 2021.
- [BKLM22] Mark Braverman, Subhash Khot, Noam Lifshitz, and Dor Minzer. An invariance principle for the multi-slice, with applications. In *FOCS 2022*, pages 228–236, 2022.
- [BKLM24a] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. On approximability of satisfiable k -csps: VI. 2024+.
- [BKLM24b] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. Reasonable bounds for combinatorial lines of length three. 2024+.
- [BKM21] Mark Braverman, Subhash Khot, and Dor Minzer. On Rich 2-to-1 Games. In *12th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 185, pages 27:1–27:20, 2021.
- [BKM22] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k -csps: I. In *STOC 2022*, pages 976–988. ACM, 2022.
- [BKM23a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k -csps: II. In *STOC 2023*, pages 632–642. ACM, 2023.
- [BKM23b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k -csps: III. In *STOC 2023*, pages 643–655. ACM, 2023.
- [BKM24a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k -csps: IV. In *STOC 2024*, pages 1423–1434. ACM, 2024.
- [BKM24b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k -csps: V. *CoRR*, abs/2408.15377, 2024.
- [BTZ10] Vitaly Bergelson, Terence Tao, and Tamar Ziegler. An inverse theorem for the uniformity seminorms associated with the action of. *Geometric and Functional Analysis*, 19(6):1539–1596, 2010.
- [Bul17] Andrei A. Bulatov. A Dichotomy Theorem for Nonuniform CSPs. In *IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–330, 2017.
- [FGL⁺96] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, 1996.
- [FK89] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem for $k = 3$. *Discrete Math.*, 75(1-3):227–241, 1989. Graph theory and combinatorics (Cambridge, 1988).
- [FK91] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem. *J. Anal. Math.*, 57:64–119, 1991.
- [FV98] Tomás Feder and Moshe Y. Vardi. The Computational Structure of Monotone Monadic SNP and Constraint Satisfaction: A Study through Datalog and Group Theory. *SIAM Journal on Computing*, 28(1):57–104, 1998.
- [GM17] WT Gowers and Luka Milićević. A quantitative inverse theorem for the U^4 norm over finite fields. *arXiv preprint arXiv:1712.00241*, 2017.

- [Gow01] W. T. Gowers. A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.
- [GT08] Ben Green and Terence Tao. An inverse theorem for the Gowers norm. *Proceedings of the Edinburgh Mathematical Society*, 51(1):73–153, 2008.
- [GW95] Michel X. Goemans and David P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *Journal of the ACM*, 42(6):1115–1145, 1995.
- [Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, July 2001. (Preliminary version in *29th STOC*, 1997).
- [Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM symposium on Theory of computing (STOC)*, pages 767–775. ACM, 2002.
- [KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O’Donnell. Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.
- [MOO05] Elchanan Mossel, Ryan O’Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *Proceedings 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 21–30, 2005.
- [Mos10] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713–1756, 2010.
- [Pol12] DHJ Polymath. A new proof of the density Hales-Jewett theorem. *Annals of Mathematics*, pages 1283–1327, 2012.
- [Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 14th Annual ACM symposium on Theory of computing (STOC)*, pages 245–254, 2008.
- [Sch78] Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the 10th Annual ACM Symposium on Theory of Computing (STOC)*, page 216–226, 1978.
- [TZ12] Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Annals of Combinatorics*, 16(1):121–188, 2012.
- [Zhu20] Dmitriy Zhuk. A Proof of the CSP Dichotomy Conjecture. *Journal of the ACM*, 67(5), August 2020.