Reasonable Bounds for Combinatorial Lines of Length Three

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

Abstract

We prove that any subset $A \subseteq [3]^n$ with $3^{-n}|A| \geqslant (\log\log\log\log n)^{-c}$ contains a combinatorial line of length 3, i.e., $x, y, z \in A$, not all equal, with $x_i = y_i = z_i$ or $(x_i, y_i, z_i) = (0, 1, 2)$ for all $i = 1, 2, \ldots, n$. This improves on the previous best bound of $3^{-n}|A| \geqslant \Omega((\log^* n)^{-1/2})$ of [D.H.J. Polymath, Ann. of Math. 2012].

Contents

1   Introduction

The density Hales-Jewett theorem, proven by Furstenberg and Katznelson [FK89,FK91], asserts that for every positive integer k and $\delta > 0$, that for sufficiently large n, a subset $S \subseteq \{0, 1, \ldots, k-1\}^n$ of size at least δk^n (density at least δ) must contain a *combinatorial line* of length k. Here, we say that points $x^{(1)}, \ldots, x^{(k)} \in S$ form a combinatorial line of length k if not all $x^{(j)}$ are equal, and for each $i = 1, 2, \ldots, n$, either $x_i^{(1)} = \cdots = x_i^{(k)}$ or $(x_i^{(1)}, \ldots, x_i^{(k)}) = (0, 1, \ldots, k-1)$. Henceforth, we refer to this as the $\mathsf{DHJ}[k]$ problem, and let $[k] := \{0, 1, \ldots, k-1\}$. This is the density version of the Hales-Jewett theorem [HJ63], which asserts for any positive integers k and r, for all sufficiently large n it holds that an r-coloring of $[k]^n$ contains a monochromatic combinatorial line. While the original proof of Furstenberg and Katznelson of the $\mathsf{DHJ}[k]$ theorem was based on ergodic theory and thus gave ineffective bounds on n in terms of k and δ, the Polymath project [Pol12] provided an elementary proof that gave effective bounds. For $k = 3$, which is the focus of this paper, [Pol12] proved that a subset $S \subseteq [3]^n$ with density δ contains a combinatorial line as long as $n \geqslant T(O(\delta^{-2}))$, where $T(m) = 2^{T(m-1)}$ and $T(0) = 1$, i.e., a tower of height $O(\delta^{-2})$. Put another way, a subset $S \subseteq [3]^n$ of density at least $\Omega((\log^* n)^{-1/2})$ contains a combiantorial line of length 3.

The main result of this paper is an improvement of this bound to a tower of finite height.

Theorem 1. *There are constants $c, C > 0$ such that for all positive integers n and subsets $|A| \subseteq [3]^n$ with $3^{-n}|A| \geqslant C(\log \log \log \log n)^{-c}$, there are $x, y, z \in A$, not all equal, forming a combinatorial line. In other words, for all $i = 1, \ldots, n$, either $x_i = y_i = z_i$ or $(x_i, y_i, z_i) = (0, 1, 2)$.*

1.1   Connections to Other Problems

The density Hales-Jewett problem has several connections to problems in extremal and additive combinatorics. The van der Waerden [VdW27] theorem says that for any positive integer k, a finite coloring of the positive integers contains a monochromatic arithmetic progression of length k (a k-AP). Szemerédi's theorem [Sze75] is the density analogue of van der Waerden's theorem, and says that for sufficiently large n in terms of δ and k, that any density δ subset of $\{1, \ldots, n\}$ contains a k-AP. In fact, Szemerédi's theorem follows from the $\mathsf{DHJ}[k]$ theorem in the following way. Encode a point $x \in [k]^n$ as $\Pi(x) := \sum_{i=1}^n x_i k^i$. Now note that a combinatorial line corresponds exactly to a k-AP. Another related problem is the corners problem, which asks to find in a dense subset of $[n] \times [n]$, three points that form an axis-aligned isosceles triangle: $(x, y), (x+d, y), (x, y+d)$. This was solved by Ajtai and Szemerédi [AS74]. This can also be deduced from the $\mathsf{DHJ}[k]$ theorem. Even more generally, the multi-dimensional Szemerédi theorem states that one can find even more complicated such patterns in dense subsets of $[n]^d$. This was also initially proven by Furstenberg and Katznelson [FK78] by ergodic methods, but now a combinatorial proof is known via reduction to the $\mathsf{DHJ}[k]$ theorem. The fact that so many problems can be reduced to $\mathsf{DHJ}[k]$ may partly explain the difficulty in obtaining improved bounds for it.

While Szemerédi's original proof did not achieve great quantitative bounds, recent works have introduced several new tools broadly based on Fourier analysis that have greatly improved our understanding of Szemerédi's theorem, while providing significantly improved quantitative bounds. In [Gow98, Gow01], Gowers introduced higher-order Fourier analysis and the Gowers U^k-norms, which have played a crucial role in establishing analytic proofs and improved bounds for special cases of the multidimensional Szemerédi theorem and polynomial Szemerédi theorem [Shk05, Shk06, GTZ11, GTZ12, GT12, GT17, Pel18, Pel19, Pel20, HLY21, PP22, Pel24, LSS23, Len24, LSS24a, LSS24b], as well as several applications to number theory and beyond [Gre05b, GT08, GT10].

In terms of lower bounds, the best known lower bound for the $\mathsf{DHJ}[k]$ problem is density $\Omega(\exp(-(\log n)^{1/\lceil \log_2 k \rceil}))$ [Pol09], using ideas from Behrend's lower bound construction for 3-AP

free sets [Beh46]. For the $k = 3$ case of Szemerédi's theorem, in a remarkable recent result, Kelley and Meka [KM23] achieved an upper bound that nearly matches Behrend's lower bound. The bound has been further refined by Bloom and Sisask [BS23].

1.2 Connections to the Study of CSPs

Motivated by applications to understanding the approximability of satisfiable constraint satisfaction problems (CSPs), Bhangale, Khot, and Minzer [BKM22, BKM23b, BKM23c, BKM24a, BKM24b, BKM23a], studied the following very general problem. Consider finite alphabets $\Sigma_1, \ldots, \Sigma_k$ and a distribution μ supported on a subset of $\Sigma_1 \times \cdots \times \Sigma_k$. If 1-bounded functions $f_i : \Sigma_i^n \to \mathbb{C}$ (so, $|f_i(x)| \leqslant 1$ for all $x \in \Sigma_i^n$) for $i = 1, \ldots, k$ satisfy

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

then what structure can we deduce about the functions f_i? This arose in the study of CSPs because a CSP is formally just a subset of $\Sigma_1 \times \cdots \times \Sigma_k$, which may correspond to the support of the distribution μ being considered.

This generality also captures several problems in additive combinatorics. To see the connection with $\mathsf{DHJ}[k]$, one can consider the setting where $\Sigma_1 = \cdots = \Sigma_k = [k]$ and μ is supported on $(x, \ldots, x)$ for $x \in [k]$ and $(0, 1, \ldots, k - 1)$, and f_i are indicators of a set $S \subseteq [k]^n$. Then (1) exactly measures the density of combinatorial lines with respect to the measure μ. Similarly, many other problems in additive combinatorics can be captured in the same way.

In [BKM24a], the following structural result is proved. Let $k = 3$ and μ be pairwise-connected (the projection of the support of μ to any two coordinates forms a connected graph, see Definition 1.1), and μ has no Abelian embeddings into $\mathbb{Z}$. In this case, there is a finite group H with size depending only on μ, such that for each $i = 1, 2, 3$, f_i correlates to some Fourier character over H times a low-degree function. Equivalently, after randomly setting some coordinates of f_i (see Definition 1.2), f_i correlates to a Fourier character over H on the remaining coordinates. This was used to give reasonable bounds for restricted 3-APs over $\mathbb{F}_p^n$, whose common difference lies in the set $\{0, 1, 2\}^n$ always.

In the companion paper [BKLM24a], we extend the result of [BKM24a] by removing the assumption that μ has no Abelian embeddings into $\mathbb{Z}$, and prove that each function f_i, after random restriction, correlates to a product function (see Theorem 2) as long as μ is pairwise-connected. Similarly, this can be applied to give reasonable bounds for restricted 3-APs over $\mathbb{F}_p^n$, but now with common difference lying in the set $\{0, 1\}^n$. These results do not immediately imply any bounds for the $\mathsf{DHJ}[3]$ problem, because the distribution μ supported on $(0, 0, 0)$, $(1, 1, 1)$, $(2, 2, 2)$, $(0, 1, 2)$ is not pairwise-connected.

To prove Theorem 1, we take inspiration from the Shkredov's approach to the corners problem [Shk06], as well as the combinatorial proof of the $\mathsf{DHJ}[k]$ theorem [Pol12]. As we discuss later, interpreting the corners problem (say over $\mathbb{F}_2^n \times \mathbb{F}_2^n$) in the context of (1) also produces a distribution μ which is not pairwise-connected. However, Shkredov proves strong bounds for the corners problem by a density increment strategy. Additionally, the Polymath paper [Pol12] gives analogies for several ingredients in the corners proof in the context of $\mathsf{DHJ}[3]$ (though their argument is based on the corners proof of Ajtai and Szemerédi and not of Shkredov). At a high level, we manage to combine the new structural results for pairwise-connected correlations from the companion papers of the authors [BKLM24a, BKLM24b] (see Theorems 2 and 3) with Shkredov's corners proof and [Pol12] to give reasonable bounds for the $\mathsf{DHJ}[3]$ problem.

DHJ[k] for $k \geqslant 4$: it is conceivable that an approach in the spirit of the current paper could be used to establish reasonable bounds for DHJ[k] for general k. Such approach would require, at the very least, analogs of the results in [BKLM24a, BKLM24b] for pairwise-connected k-ary CSPs, which we plan to study in future works.

1.3 Preliminaries

General notation. Let $[n] = \{1, \ldots, n\}$. For a subset $I \subseteq [n]$ we write $\overline{I} = [n] \setminus I$. We let $\mathbb{C}$ denote complex numbers. We let $\mathsf{supp}(\mu)$ denote the support of a distribution μ. For a finite set Σ we write $x \sim \Sigma^n$ to denote sampling x uniformly randomly from Σ^n.

Connectivity. We define what it means for a distribution to be pairwise-connected. Let μ_{ij} be the restriction of the distribution μ to the (i, j) coordinates.

Definition 1.1 (Pairwise-connected). We say that a distribution μ on $\Sigma_1 \times \cdots \times \Sigma_k$ is pairwise-connected if for all $1 \leqslant i < j \leqslant k$ the support of μ_{ij} forms a connected graph over vertex set $\Sigma_i \cup \Sigma_j$.

Random restriction. All our methods and results heavily rely on random restrictions. A random restriction takes a subset of the coordinates, and randomly fixes their values. Below, $I \sim_{1-\alpha} [n]$ means that I is a random subset of $[n]$, such that each $i \in [n]$ is included in I with probability $1 - \alpha$, and $z \sim \nu^I$ means that $z \in \Sigma^I$ is such that each z_i is i.i.d. and distributed according to ν.

Definition 1.2 (Random restriction). Let μ be a distribution over Σ^n and let $\mu = (1 - \alpha)\nu + \alpha\mu'$ for distributions ν, μ'. Then for a function $f : \Sigma^n \to \mathbb{C}$, $I \sim_{1-\alpha} [n]$, and $z \sim \nu^I$, we define the random restriction $f_{I \to z} : \Sigma^{[n] \setminus I} \to \mathbb{C}$ as $f_{I \to z}(x) := f(x, z)$.

2 Proof Outline

2.1 Shkredov's Corners Proof

We start by recalling the key ideas behind Shkredov's bound on sets that do not contain a corner [Shk05, Shk06]. For simplicity we consider the finite field version of the problem, where we have a set $S \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$. A corner consists of three points $(x, y), (x + d, y), (x, y + d) \in S$ such that $d \neq 0$. The standard approach to such problems is the density increment method pioneered by Roth [Rot53] to solve the $k = 3$ case of Szemerédi's theorem. At a high level, this method tries to prove that either the set S has about the expected number of corners, or the set S has larger density on some structured subset (for example, a subspace). Then one can pass to this structured object and repeat the argument.

One type of structure that arises in the corners problem is that of combinatorial rectangles: this was observed by Ajtai and Szemerédi [AS74]. Let $\mu(S)$ be the density of S. Then the number of corners in S is given by $\mathrm{Pr}_{x,y,d \in \mathbb{F}_2^n} [(x, y), (x, y + d), (x + d, y) \in S]$. When can this value deviate far from $\mu(S)^3$, which is what it would be if the events $(x, y) \in S$, $(x, y + d) \in S$, $(x + d, y) \in S$ were truly independent? A natural example is when $S \subseteq E_1 \times E_2$ for some $E_1 \subseteq \mathbb{F}_2^n$, $E_2 \subseteq \mathbb{F}_2^n$. Let $\mu' := \mu(S)/\mu(E_1 \times E_2)$ be the relative density of S within $E_1 \times E_2$. Then heuristically, the probability that $(x, y), (x, y + d), (x + d, y) \in S$ is approximately $\mu(E_1)^2 \mu(E_2)^2 (\mu')^3$, which is much larger than $\mu(S)^3 = \mu(E_1)^3 \mu(E_2)^3 (\mu')^3$. This heuristic holds as long as:

1. The probability that $x, x + d \in E_1$ and $y, y + d \in E_2$ is approximately $\mu(E_1)^2 \mu(E_2)^2$.

3

2. The probability that (x, y), $(x, y + d)$, $(x + d, y) \in S$ conditioned on item 1 is about $(\mu')^3$.

To complete the argument, one must understand the precise conditions under which these items hold. It is not too difficult to argue that item 1 holds as long as E_1 and E_2 are U^2-pseudorandom. In other words, all their nontrivial Fourier coefficients are small. Also, a (nontrivial) sequence of Cauchy-Schwarz manipulations proves that item 2 holds if S does not admit a density increment onto a combinatorial subrectangle $E_1' \times E_2' \subseteq E_1 \times E_2$ of nonnegligible size.

From here, Shkredov's proof proceeds via a "double descent" kind of argument. The goal is to maintain $S \subseteq E_1 \times E_2$ such that the relative density of S within $E_1 \times E_2$ increases whenever there are no corners. Given such $S \subseteq E_1 \times E_2$, one first tries to find a sub-instance where E_1 and E_2 are U^2-pseudorandom. This is referred to as the "uniformizing" step. At a high level, this is done by partitioning E_1 and E_2 among subspaces (into smaller combinatorial rectangles) on which they have large Fourier coefficients, until most pieces in the partition are U^2-pseudorandom. By an averaging argument, there exists a piece on which S still has large relative density.

Now, Shkredov argues that either the number of corners is approximately equal to what the heuristic above says, or S admits a relative density increment onto a combinatorial subrectangle $E_1' \times E_2'$. This is proven by several applications of the Cauchy-Schwarz inequality. Of course, now E_1', E_2' are not necessarily U^2-pseudorandom anymore, so one has to uniformize again, and repeat the argument.

2.2 Dictionary between Corners and DHJ[3]

The work [Pol12] gives a combinatorial proof of the density Hales-Jewett theorem by drawing analogies from the corners proof of Ajtai and Szemerédi [AS74]. Here we describe how each aspect of the corners proof described in Section 2.1 can be adapted to the DHJ[3] setting.

One immediate difference in the DHJ[3] setting is that it is not clear under what measure we should be counting triples $(x, y, z) \in S$ that form a combinatorial line. [Pol12] introduces the *equal slices measure* under which the counts of combinatorial lines is sensibly defined. In this work, we take a slightly different approach that is similar in spirit. Regardless, precisely defining the equal slices measure or our variant is not important for this outline – it is only important that there is some sensible way to count combinatorial lines.

Let $S \subseteq [3]^n$ be the set on which we wish to find combinatorial lines. [Pol12] proves that if the count of combinatorial lines is far off from what is expected, then S has increased relative density on the intersection of a 02-insensitive and 01-insensitive set (see also a simpler proof from [DKT14] that avoids the equal slice measure and works with the uniform distribution throughout). Here, a 02-insensitive set is a subset $E_1 \subseteq [3]^n$ satisfying that if $x, y \in [3]^n$ agree on all $i \in [n]$ where $x_i = 1$ or $y_i = 1$, then $x \in E_1$ if and only if $y \in E_1$. In other words, only the locations of the 1's in x determines whether $x \in E_1$. A 01-insensitive set (which we call E_2) is defined similarly. We refer to the intersection of E_1 and E_2 as their *disjoint product* (see Definition 4.1) and denote it as $E_1 \boxtimes E_2$. This is the first entry of the dictionary: $E_1 \boxtimes E_2$ corresponds to the combinatorial rectangles of the corners proof.

From here, [Pol12] argues that $E_1 \boxtimes E_2$ can be split into combinatorial subspaces (i.e., sets isomorphic to $[3]^{n'}$ for $n' \leqslant n$), and thus one obtains a density increment of S onto a combinatorial subspace, and the argument can be repeated. In [Pol12], $n' \leqslant \log n$ due to an application of the multidimensional Szemerédi theorem (for 2 points), and thus the final density bound achieved on S is $(\log^* n)^{-c}$. Similarly, our proof will also use combinatorial subspaces, but for larger $n' \geqslant n^c$.

4

2.3 Pseudorandomness Notion

The main missing ingredient from fully translating Shkredov's proof to the DHJ[3] setting is the correct notion of pseudorandomness. It turns out that the correct notion is what we call *product-pseudorandomness* (see Definition 3.1). While U^2 pseudorandomness of a function f says that $|\langle f, \prod_{i=1}^n \chi_i \rangle|$ is small for any Fourier characters $\chi_1, \dots, \chi_n$, product function pseudorandomness is a stronger notion which says that $|\langle f, \prod_{i=1}^n P_i \rangle|$ is small for any choice of 1-bounded functions $P_1, \dots, P_n : [3] \to \mathbb{C}$, where 1-bounded means that $|P_i(x)| \leqslant 1$ for all $x \in [3]$. In fact, in our definition we also require that random restrictions of f down to even $n^{1/4}$ coordinates correlate to product functions $\prod_{i=1}^n P_i$ with only negligible probability. This notion of pseudorandomness is motivated by the inverse theorems in [BKLM24a, BKLM24b] (see Theorems 2 and 3) which say that if functions f, g, h have large 3-wise correlation over $\mu^{\otimes n}$ for a pairwise-connected distribution μ, then each of f, g, h is not product-pseudorandom.

It remains to discuss how to put all these pieces together. Similarly to the corners proof, our goal is to increment the relative density of S within $E_1 \boxtimes E_2$ while there is no combinatorial line. We design a uniformization algorithm for our notion of product function pseudorandomness – this essentially amounts to understanding how to density increment on sets that correlate to some product function. This is done in Section 7, and is based on the density increment for restricted 3-progressions in [BKM23a, BKLM24a]. Finally we argue that if the number of combinatorial lines (under the proper measure) deviates from what is expected, then S has a relative density increment onto a "subrectangle" $E_1' \boxtimes E_2' \subseteq E_1 \boxtimes E_2$. This is done in Sections 5 and 6.

It is worth discussing where product pseudorandomness enters the picture, more precisely. For example, we wish to argue that $E_1 \boxtimes E_2$ behaves like a combinatorial rectangle. One way to formalize this is that we want to have $\mu(E_1 \boxtimes E_2) \approx \mu(E_1)\mu(E_2)$, i.e., that the events $x \in E_1$ and $x \in E_2$ for $x \in [3]^n$ behave independently. In Lemma 4.3 we argue that product pseudorandomness of E_1 and E_2 suffices for this. In addition, throughout the proof we require more complex correlations of a similar type (e.g., Lemmas 5.8 and 6.1 to 6.3) which all leverage product pseudorandomness. We believe that any form of Fourier pseudorandomness would not suffice to prove these statements.

2.4 Choosing Parameters

Throughout the paper we will use the following parameters and relationships between them.

The parameter α will be the relative density of the set S within the disjoint product $E_1 \boxtimes E_2$. The density of E_1 and E_2 within $[3]^n$ will be δ_1 and δ_2 respectively. These will decrease during density increments, but will always stay above $\exp(-\alpha^{-C})$ (say, for $C = 20$).

The parameter γ will be a pseudorandomness parameter. We require that $\log(1/\gamma)^{-1} < (\alpha\delta_1\delta_2)^C$ for a large constant C. Thus, $\gamma \leqslant \exp(-\exp(\alpha^{-C}))$.

During a uniformization step, n will drop to about $n^{\exp(-\gamma^{-C})}$ for a constant C (see Theorem 6). This informs our choice of α ultimately, because we need

$$\exp(-\gamma^{-C}) \leqslant (\log n)^c, \quad \text{so } \alpha \geqslant \Omega((\log\log\log\log n)^{-c}),$$

by our choice of γ.

Finally, we will set the parameter $\eta = \exp(-\sqrt{\log n})$, where n is the original dimension. Throughout the whole procedure, η will be much smaller than any of $\alpha, \delta_1, \delta_2$, and γ. The parameter η arises sometimes when we are dealing with distributions ν on $[3]$ satisfying $|\nu(x) - 1/3| \leqslant \eta/\sqrt{n}$. In other words, $d_{\mathrm{TV}}(\mu^{\otimes n}, \nu^{\otimes n}) \leqslant O(\eta)$, where μ is the uniform distribution on $[3]$.

3 Correlations to Product Functions

3.1 Product Pseudorandomness

We start by formally introducing the main pseudorandomness notion considered in this paper.

Definition 3.1 $((n', \gamma)$-product pseudorandomness)**.** Let Σ be a finite set, μ be a distribution on Σ, and n a positive integer. For $n' \leqslant n$ and $\gamma > 0$ we say that a 1-bounded function $f : \Sigma^n \to \mathbb{C}$ is (n', γ)-product pseudorandom if for any $\delta \in [n'/n, 1]$, the probability that a random restriction of f down to $\bar{I} \sim_\delta [n]$ γ-correlates to a product function is less than γ. Precisely,

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

Informally, another interpretation of product pseudorandomness is that f does not correlate to a function of the form LP where L has *degree* at most $\widetilde{O}(n/n')$, and P is a product function. This can potentially be formalized by proving a restriction inverse theorem akin to [BKM23c, BKM24a], but it is more convenient to just work with Definition 3.1 in the present work.

It is worth remarking that one can change the measure that a function f is product pseudorandom with respect to by applying random restrictions. Informally, if f is product pseudorandom with respect to μ and $\mu = \beta \nu + (1 - \beta)\nu'$, then applying a random restriction $I \sim_{1-\beta} [n]$ and $z \sim (\nu')^{\otimes I}$, the function $f_{I \to z}$ will be product pseudorandom against ν. A more formal and general statement appears in Lemma 5.5.

3.2 Inverse Theorems for CSPs

In this section we discuss a significant generalization of the U^2 Gowers-norm inverse theorem that we require for our proof. The following theorem was proven in [BKLM24a, Theorems 1, 2].

Theorem 2. *Let μ be a pairwise-connected distribution over $\Sigma \times \Gamma \times \Phi$ for finite sets Σ, Γ, Φ on which the probability of each atom is at least α. If 1-bounded functions $f : \Sigma^n \to \mathbb{C}$, $g : \Gamma^n \to \mathbb{C}$, $h : \Phi^n \to \mathbb{C}$ satisfy that*

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

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

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

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

Put another way, the functions f, g, h are not product pseudorandom. To see the connection to the U^2 inverse theorem, one can consider when $\Sigma = \Gamma = \Phi = H$ for some Abelian group H and μ is uniform over $(x, x + d, x + 2d)$ for $x, d \in H$.

In this present work, we require an inverse theorem for certain 4-ary distributions. This is done by reducing to the 3-ary inverse theorem in Theorem 2. The analogy is that the U^2 Gowers norm is over 4-ary distributions $(x, x + a, x + b, x + a + b)$, but its inverse structure behaves the same as for 3-APs. To formally state this theorem, we need to define what it means for a distribution to be connected.

Definition 3.2. A subset $S \subseteq \Sigma_1 \times \cdots \times \Sigma_k$ is said to be connected if for every $a = (a_1, a_2, \ldots, a_k) \in S$ and $b = (b_1, b_2, \ldots, b_k) \in S$, there exists a sequence $(a, c^{(1)}, \ldots, c^{(t)}, b)$ for some $t \geqslant 1$ such that $c^{(i)} \in S$ for all i and every consecutive pair of tuples in the sequence differs in at most one coordinate. A distribution μ on $\Sigma_1 \times \cdots \times \Sigma_k$ is said to be connected if the support of μ is connected.

The following theorem is from [BKLM24b, Lemma 1.5].

Theorem 3. *Let μ be a distribution over $\Sigma_1 \times \Sigma_2 \times \Sigma_3 \times \Sigma_4$ such that the probability of each atom is at least α and the distribution μ_{123} is connected. There is a constant $\gamma := \gamma(\alpha, \varepsilon) > 0$ such that if 1-bounded functions $f_i : \Sigma_i^n \to \mathbb{C}$ for $i = 1, 2, 3, 4$ satisfy*

$$\left| \mathop{\mathbb{E}}_{(x_1, x_2, x_3, x_4) \sim \mu^{\otimes n}} [f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)] \right| \geqslant \varepsilon,$$

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

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

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

4 Inductive Structure

Let μ be a measure on $[3]^n$. Let $\pi_1 : \{0, 1, 2\}^n \to \{0, 1\}^n$ be defined as $\pi_1(x)_i = 1$ if and only if $x_i = 1$. Similarly, $\pi_2 : \{0, 1, 2\}^n \to \{0, 2\}^n$ is defined as $\pi_2(x)_i = 2$ if and only if $x_i = 2$.

The measure of E_1 is denoted as

$$\mu(E_1) := \mu(\{x \in [3]^n : \pi_1(x) \in E_1\}).$$

$\mu(E_2)$ is defined similarly. By abuse of notation, we let μ denote the distribution on $\{0, 1\}^n$ (respectively $\{0, 2\}^n$) distributed as $\pi_1(x)$ (resp. $\pi_2(x)$) for x distributed according to μ in $[3]^n$.

We now define the disjoint product of subsets $E_1 \subseteq \{0, 1\}^n$ and $E_2 \subseteq \{0, 2\}^n$ which serves as our analogy to combinatorial rectangles.

Definition 4.1 (Disjoint product). Let $E_1 \subseteq \{0, 1\}^n$ and $E_2 \subseteq \{0, 2\}^n$. The disjoint product $E_1 \boxtimes E_2 \subseteq \{0, 1, 2\}^n$ is defined as:

$$E_1 \boxtimes E_2 := \{x \in [3]^n : \pi_1(x) \in E_1 \text{ and } \pi_2(x) \in E_2\}.$$

Our density increment proceeds by maintaining a subset $S \subseteq E_1 \boxtimes E_2$ whose relative density increases over steps of the density increment.

Definition 4.2 (Structure). We define Struct_α to consist of triples (S, E_1, E_2) satisfying $E_1 \subseteq \{0, 1\}^n$, $E_2 \subseteq \{0, 2\}^n$ with $S \subseteq E_1 \boxtimes E_2$ and for μ uniform over $[3]^n$:

1. $\mu(S) \geqslant \alpha \cdot \mu(E_1 \boxtimes E_2)$.

2. Let $\delta_1 = \mu(E_1)$ and $\delta_2 = \mu(E_2)$. The functions $1_{E_1} - \delta_1 : \{0, 1\}^n \to \mathbb{R}$ and $1_{E_2} - \delta_2 : \{0, 2\}^n \to \mathbb{R}$ are (n', γ)-product pseudorandom for $n' = n^{1/4}$ and $\gamma := \gamma(\alpha, \delta_1, \delta_2)$ chosen later.

7

At this point it is worth noting that the pseudorandomness condition implies that $\mu(E_1 \boxtimes E_2) \approx \mu(E_1)\mu(E_2)$. In fact, this holds as long as the measures of E_1 and E_2 do not change significantly under random restriction – this was previously observed by Austin [Aus11, Lemma 6.1] in a different language.

Lemma 4.3. *Let μ be a distribution on $[3]$ with mass $\Omega(1)$ on each atom. If E_1, E_2 are $(n^{0.99}, \gamma)$ product pseudorandom with respect to μ then*

$$|\mu(E_1 \boxtimes E_2) - \mu(E_1)\mu(E_2)| \leqslant 2\gamma^{1/2}.$$

Proof. Let $\overline{E_1} := 1_{E_1} - \mu(E_1)$ and $\overline{E_2} := 1_{E_2} - \mu(E_2)$. Then

$$\mu(E_1 \boxtimes E_2) = \mathop{\mathbb{E}}_{x \sim \mu^{\otimes n}} [1_{E_1}(x) 1_{E_2}(x)] = \mu(E_1)\mu(E_2) + \mathop{\mathbb{E}}_{x \sim \mu^{\otimes n}} \left[\overline{E_1}(x)\overline{E_2}(x) \right]$$

$$= \mu(E_1)\mu(E_2) + \mathop{\mathbb{E}}_{x \sim \mu^{\otimes n}} \left[\overline{E_1}(\pi_1(x))\overline{E_2}(\pi_2(x)) \right].$$

Assume for contradiction that $\left| \mathbb{E}_{x \sim \mu^{\otimes n}} \left[\overline{E_1}(x)\overline{E_2}(x) \right] \right| > 2\sqrt{\gamma}$. Then by Cauchy-Schwarz:

$$4\gamma \leqslant \mathop{\mathbb{E}}_{\pi_2(x)} \left| \mathop{\mathbb{E}}_{x : \pi_2(x)} \overline{E_1}(\pi_1(x)) \right|^2 = \mathop{\mathbb{E}}_{\pi_2(x)} \mathop{\mathbb{E}}_{x, x' : \pi_2(x)} \overline{E_1}(\pi_1(x))\overline{E_1}(\pi_1(x')),$$

where the notation means that we first sample $\pi_2(x)$ for $x \sim \mu^{\otimes n}$, and then sample x, x' conditioned on $\pi_2(x)$. Because μ has full support on $[3]$, one can check that $(\pi_1(x), \pi_1(x'))$ has full support on $(0,0), (0,1), (1,0), (1,1)$. Let ν be the distribution on $(\pi_1(x), \pi_1(x'))$ and let $\nu = \beta U + (1-\beta)\nu'$ where U is uniform over $\{0,1\} \times \{0,1\}$, and $\beta = \Omega(1)$. Then we conclude that

$$4\gamma \leqslant \mathop{\mathbb{E}}_{I \sim_{1-\beta} [n]} \mathop{\mathbb{E}}_{(z,z') \sim (\nu')^{\otimes I}} \mathop{\mathbb{E}}_{(x,x') \sim U^{\otimes \overline{I}}} \left[(\overline{E_1})_{I \to z}(x)(\overline{E_1})_{I \to z'}(x') \right]$$

$$= \mathop{\mathbb{E}}_{I \sim_{1-\beta} [n]} \mathop{\mathbb{E}}_{(z,z') \sim (\nu')^{\otimes I}} \left(\mathop{\mathbb{E}}_{x \sim U^{\otimes \overline{I}}} \left[(\overline{E_1})_{I \to z}(x) \right] \right) \left(\mathop{\mathbb{E}}_{x \sim U^{\otimes \overline{I}}} \left[(\overline{E_1})_{I \to z'}(x) \right] \right)$$

$$\leqslant \mathop{\mathbb{E}}_{I \sim_{1-\beta} [n]} \mathop{\mathbb{E}}_{(z,z') \sim (\nu')^{\otimes I}} \left| \mathop{\mathbb{E}}_{x \sim U^{\otimes \overline{I}}} \left[(\overline{E_1})_{I \to z}(x) \right] \right|.$$

By an averaging argument and the fact that $\beta n \geqslant n^{0.99}$, this contradicts that $\overline{E_1}$ is $(n^{0.99}, \gamma)$ product pseudorandom. $\qquad\square$

5 Moving between Measures

In this section we prove that if S contains no combinatorial lines, then some particular 3-wise correlation over a product distribution whose atoms are at least $\Omega(1)$ in each coordinate is large. Showing this involves picking a suitable measure under which to count combinatorial lines.

5.1 Lower Bounding the Expected Count

Let ν be a measure on $[3]^n \times [3]^n \times [3]^n$, such that for all $i \in [n]$, ν_i is supported on triples $(0,0,0), (1,1,1), (2,2,2), (0,1,2)$. Let $S \subseteq E_1 \boxtimes E_2 \subseteq [3]^n$ with relative density α in $E_1 \boxtimes E_2$, E_1 has density δ_1, and E_2 has density δ_2 (here, we are being a bit informal in that we are not specifying the density we are using here, but it may not be uniform). The following discussion holds for any ν, but we will pick convenient ones to work with soon.

8

We can count the density of combinatorial lines contained in S with respect to ν. This is:

$$\underset{(x,y,z)\sim\nu}{\mathbb{E}}[1_S(x)1_S(y)1_S(z)] = \underset{(x,y,z)\sim\nu}{\mathbb{E}}[(1_S(x) - \alpha 1_{E_1 \boxtimes E_2}(x))1_S(y)1_S(z)]$$
$$+ \alpha \underset{(x,y,z)\sim\nu}{\mathbb{E}}[1_{E_1 \boxtimes E_2}(x)1_S(y)1_S(z)].$$

On the right hand side of the equation, the first term is the "error" term, and the second is the "main term". Studying the main term, one can note that $1_{E_1 \boxtimes E_2}(x)1_S(y)1_S(z) = 1_S(y)1_S(z)$ in fact, because $\pi_2(y) = \pi_2(x)$ and $\pi_1(z) = \pi_1(x)$ for all $(x, y, z) \in \mathsf{supp}(\nu)$. Thus, the main term can we written more succinctly as: $\alpha \mathbb{E}_{(x,y,z)\sim\nu}[1_S(y)1_S(z)]$. We would like to pick a measure ν on which we can lower bound this quantity. This is essentially asking for a measure ν on which the counting version of $\mathsf{DHJ}[2]$ holds. A key observation made in [Pol12, Theorem 3.1] is that this holds when ν is the *equal-slices measure*. We take a somewhat different approach here which is similar in spirit, but allows us to work only with product distributions throughout the arguments.

Let η be sufficiently small (depending on $\alpha, \delta_1, \delta_2$), and let $K \in \mathbb{Z}_{\geqslant 1}, \eta' \in \mathbb{R}_{>0}$ satisfy $K\eta' \leqslant \eta/100$. Define the following distribution over sequences in $([3]^n)^{K+1}$.

Definition 5.1. Define the distribution $\mathcal{D}$ over $(y^{(0)}, \dots, y^{(K)}) \in ([3]^n)^{K+1}$ as follows. $y^{(0)}$ is drawn uniformly in $[3]^n$. For $i = 0, \dots, K - 1$, define for $t \in [n]$ independently:

$$y_t^{(i+1)} := \begin{cases} y_t^{(i)}, & \text{if } y_t^{(i)} \in \{0, 2\}, \\ 1 \text{ w.p. } 1 - \frac{\eta'}{\sqrt{n}}, & \text{if } y_t^{(i)} = 1, \\ 2 \text{ w.p. } \frac{\eta'}{\sqrt{n}}, & \text{if } y_t^{(i)} = 1. \end{cases}$$

In other words, to get $y^{(i+1)}$ from $y^{(i)}$, for each index $t \in [n]$ where $y_t^{(i)} = 1$, we change it to a 2 with probability $\eta'/\sqrt{n}$. We now record an observation about what the distribution of $y^{(i)}$ and joint distribution $(y^{(i)}, y^{(j)})$ look like.

Observation 1. For all $i = 0, \dots, K$, $y^{(i)}$ is distributed as $(\nu^{(i)})^{\otimes n}$ for a distribution $\nu^{(i)}$ on $[3]$ satisfying that $\nu^{(i)}(0) = 1/3$ and $|\nu^{(i)}(x) - 1/3| \leqslant \eta/\sqrt{n}$ for $x \in \{1, 2\}$. For all pairs $0 \leqslant i < j \leqslant K$, $(y^{(i)}, y^{(j)})$ is distributed as $(\xi^{(ij)})^{\otimes n}$ for some distribution $\xi^{(ij)} \in [3] \times [3]$ supported on $(0, 0), (1, 1), (2, 2)$, and $(1, 2)$ satisfying that:

1. $\xi^{(ij)}((0, 0)) = 1/3$ and $|\xi^{(ij)}((x, x)) - 1/3| \leqslant \eta/\sqrt{n}$ for $x \in \{1, 2\}$.

2. $\xi^{(ij)}((1, 2)) \geqslant \frac{\eta'}{10\sqrt{n}}$.

The first item is actually redundant with the claim about the distribution of $y^{(i)}$, but we list it for completeness. Finally, we argue that there is some pair (i, j) under which $\mathbb{E}_{(y,z)\sim(\xi^{(ij)})^{\otimes n}}[1_S(y)1_S(z)]$ can be lower bounded.

Lemma 5.2. *Let μ be uniform on $[3]^n$. For every subset $S \subseteq [3]^n$, there is $0 \leqslant i < j \leqslant K$ such that*

$$\underset{(y,z)\sim(\xi^{(ij)})^{\otimes n}}{\mathbb{E}}\left[1_S(y)1_S(z)\right] \geqslant \mu(S)^2 - 6\eta - \frac{\mu(S)}{K}.$$

Proof. Note that

$$\sum_{i \neq j} 1_S(y^{(i)})1_S(y^{(j)}) = \left(\sum_{i=0}^{K} 1_S(y^{(i)})\right)^2 - \sum_{i=0}^{K} 1_S(y^{(i)}).$$

9

Also, because $d_{TV}(\nu(i), \mu) \leqslant 2\eta$ by Observation 1 for all $i = 0, \ldots, K$ we know that

$$(K+1)(\mu(S) + 2\eta) \geqslant \mathop{\mathbb{E}}_{(y^{(0)}, \ldots, y^{(K)}) \sim \mathcal{D}} \left[\sum_{i=0}^{K} 1_S(y^{(i)}) \right] \geqslant (K+1)(\mu(S) - 2\eta).$$

Thus by Cauchy-Schwarz

$$\sum_{i \neq j} \mathop{\mathbb{E}}_{(y^{(0)}, \ldots, y^{(K)}) \sim \mathcal{D}} \left[1_S(y^{(i)}) 1_S(y^{(j)}) \right] \geqslant (K+1)^2 (\mu(S) - 2\eta)^2 - (K+1)(\mu(S) + 2\eta).$$

By averaging, there is a pair (i, j) such that

$$\mathop{\mathbb{E}}_{(y^{(0)}, \ldots, y^{(K)}) \sim \mathcal{D}} \left[1_S(y^{(i)}) 1_S(y^{(j)}) \right] \geqslant \frac{(K+1)^2 (\mu(S) - 2\eta)^2 - (K+1)(\mu(S) + 2\eta)}{(K+1)K}$$

$$\geqslant (\mu(S) - 2\eta)^2 - \frac{\mu(S) + 2\eta}{K} \geqslant \mu(S)^2 - 6\eta - \frac{\mu(S)}{K}.$$

This is exactly the desired conclusion, given the definition of $\xi^{(ij)}$ in Observation 1. $\qquad\square$

Let $\xi^{(ij)'}$ be the unique distribution supported on $(0,0,0), (1,1,1), (2,2,2), (0,1,2)$ such that $(\xi^{(ij)'})_{yz} = \xi^{(ij)}$. We will use $\xi^{(ij)'}$ to argue that if a set $S \subseteq E_1 \boxtimes E_2$ has no combinatorial lines, then some specific 3-wise correlation is large.

Lemma 5.3. *Let $S \subseteq E_1 \boxtimes E_2$ be such that $\mu(E_1 \boxtimes E_2) = \delta$, $\mu(S) \geqslant \alpha\delta$, and S contains no combinatorial line. For $K \geqslant 100\mu(S)^{-1}$ and $\eta \leqslant \mu(S)^2/100$ there is some $0 \leqslant i < j \leqslant K$ such that*

$$\left| \mathop{\mathbb{E}}_{(x,y,z) \sim (\xi^{(ij)'})^{\otimes n}} \left[(1_S - \alpha 1_{E_1 \boxtimes E_2})(x) 1_S(y) 1_S(z) \right] \right| \geqslant \alpha^3 \delta^2/2.$$

Proof. By Lemma 5.2 there are $0 \leqslant i < j \leqslant K$ such that

$$\mathop{\mathbb{E}}_{(y,z) \sim (\xi^{(ij)})^{\otimes n}} \left[1_S(y) 1_S(z) \right] \geqslant \mu(S)^2 - 6\eta - \frac{\mu(S)}{K} \geqslant \frac{3}{4} \mu(S)^2 \geqslant \frac{3}{4} \alpha^2 \delta^2.$$

If S has no combinatorial line, then

$$\exp(-\Omega(\eta' \sqrt{n})) \geqslant \mathop{\mathbb{E}}_{(x,y,z) \sim (\xi^{(ij)'})^{\otimes n}} \left[1_S(x) 1_S(y) 1_S(z) \right]$$

$$= \mathop{\mathbb{E}}_{(x,y,z) \sim (\xi^{(ij)'})^{\otimes n}} \left[(1_S - \alpha 1_{E_1 \boxtimes E_2})(x) 1_S(y) 1_S(z) \right]$$

$$+ \alpha \mathop{\mathbb{E}}_{(x,y,z) \sim (\xi^{(ij)'})^{\otimes n}} \left[1_{E_1 \boxtimes E_2}(x) 1_S(y) 1_S(z) \right]$$

$$\geqslant \mathop{\mathbb{E}}_{(x,y,z) \sim (\xi^{(ij)'})^{\otimes n}} \left[(1_S - \alpha 1_{E_1 \boxtimes E_2})(x) 1_S(y) 1_S(z) \right] + \frac{3}{4} \alpha^3 \delta^2.$$

The first inequality followed because $\xi^{(ij)'}((0,1,2)) \geqslant \frac{\eta'}{10\sqrt{n}}$ by Item 2 of Observation 1. $\qquad\square$

10

5.2 Harsh Random Restrictions

Our goal will now be to take a random restriction down to a distribution supported on $(0,0,0)$, $(1,1,1)$, $(2,2,2)$, $(0,1,2)$ whose mass on all atoms is $\Omega(1)$. Towards this, we first prove that taking such random restrictions preserves product pseudorandomness. In fact, if the restriction is sufficiently harsh, the function becomes product pseudorandom against any fixed distribution with $\Omega(1)$ mass on each with high probability.

First, we record the simple observation that if f is 1-bounded and correlates to a 1-bounded product function P, then f correlates with a 1-bounded product function with noticeable probability under random restriction.

Lemma 5.4. *Let $f : \Sigma^n \to \mathbb{C}$ be a 1-bounded function, and P is a 1-bounded product function. Let $\mu = \beta\mu' + (1-\beta)\nu$ be such that $\left| \mathbb{E}_{x\sim\mu^{\otimes n}}[f(x)P(x)] \right| \geqslant \varepsilon$. Then*

$$
\Pr_{I\sim_{1-\beta}[n],z\sim\nu^{\otimes I}} \left[\left| \mathbb{E}_{x\sim(\mu')^{\otimes\overline{I}}}[f_{I\to z}(x)P_{I\to z}(x)] \right| \geqslant \varepsilon/2 \right] \geqslant \varepsilon/2.
$$

Proof. Follows by a simple averaging argument, as f, P are 1-bounded. $\qquad\square$

Lemma 5.5. *Let Σ be a finite set and let $f : \Sigma^n \to \mathbb{C}$ be a 1-bounded function which is $(n^{1/4}, \gamma)$ product pseudorandom against a distribution μ with $\mu(x) \geqslant \alpha$ for all $x \in \Sigma$. Let $\eta \geqslant 0$ and ν be a distribution on Σ with $|\nu(x) - \mu(x)| \leqslant \eta/\sqrt{n}$ for all $x \in \Sigma$. Then for every distribution μ' on Σ with mass at least $\beta \geqslant n^{-1/20}$ on every atom and $n^{-11/20} \leqslant \delta \leqslant \eta/\sqrt{n}$, the following holds for $\gamma' = 2(\gamma + 3\eta/\alpha)^{1/3}$:*

$$
\Pr_{I\sim_{1-\delta}[n],z\sim\nu^{\otimes I}} \left[f_{I\to z} \text{ is } (n^{1/3}, \gamma') \text{ product pseudorandom against } \mu' \right] \geqslant 1 - \gamma'.
$$

Proof. Throughout this proof, all inner products are with respect to distributions $\mu^{\otimes n'}$ for some n'. For clarity, we denote this by $\langle f, g\rangle_\mu := \mathbb{E}_{x\sim\mu^{\otimes n'}}[f(x)g(x)]$.

Let $n' = n - |I|$, and note that $n' \geqslant n^{2/5}$ with high probability. If $f_{I\to z}$ is not $(n^{1/3}, \gamma')$ product pseudorandom with respect to μ', then for $\delta' = n^{1/3}/n'$, with probability at least γ' over $I' \sim_{1-\delta'} ([n] \setminus I)$ and $z' \sim (\mu')^{\otimes I'}$, there is a 1-bounded product function P with $|\langle f_{(I,I')\to(z,z')}, P\rangle_{\mu'}| \geqslant \gamma'$. We may write $\mu' = \beta\mu + (1-\beta)\xi$ where ξ is some distribution. If $|\langle f_{(I,I')\to(z,z')}, P\rangle_{\mu'}| \geqslant \gamma'$ holds, then by Lemma 5.4, over $I'' \sim_{1-\beta} ([n] \setminus (I \cup I'))$ and $z'' \sim \xi^{\otimes I''}$ the probability that

$$
|\langle f_{(I,I',I'')\to(z,z',z'')}, P_{I''\to z''}\rangle_\mu| \geqslant \gamma'/2
$$

is at least $\gamma'/2$. Putting this together gives that

$$
\Pr_{I',I'',z',z''} \left[|\langle f_{(I,I',I'')\to(z,z',z'')}, P\rangle_\mu| \geqslant \gamma'/2 \text{ for 1-bounded } P \right] \geqslant (\gamma')^2/2.
$$

If the conclusion of the theorem is false, then

$$
\Pr_{\substack{I\sim_{1-\delta}[n],z\sim\nu^{\otimes I} \\ I'\sim_{1-\delta'}([n]\setminus I),z'\sim(\mu')^{\otimes I'} \\ I''\sim_{1-\beta}([n]\setminus(I\cup I')),z''\sim\xi^{\otimes I''}}} \left[|\langle f_{(I,I',I'')\to(z,z',z'')}, P\rangle_\mu| \geqslant \gamma'/2 \text{ for 1-bounded } P \right] \geqslant (\gamma')^3/2 \geqslant 4(\gamma + 3\eta/\alpha).
$$

Consider the joint restriction $(I, I', I'') \to (z, z', z'')$. This is equivalent to taking a restriction $f_{J\to u}$ for $J \sim_{1-\delta''} [n], u \sim (\mu'')^{\otimes J}$ for $\delta'' = \delta\beta\delta' \geqslant n^{-3/4}$. Also $|\mu''(x) - \mu(x)| \leqslant \eta/\sqrt{n} + \delta \leqslant 2\eta/\sqrt{n}$ for

11

all $x \in \Sigma$. This implies that $\mathrm{d}_{\mathrm{TV}}(\mu^{\otimes n}, (\mu'')^{\otimes n}) \leqslant 10\eta/\alpha$ because $\mu(x) \geqslant \alpha$ for all $x \in \Sigma$. Applying this to the previous displayed equation gives

$$\Pr_{J \sim_{1-\delta\beta\delta'}[n], u \sim \mu^{\otimes J}} \left[|\langle f_{J \to u}, P \rangle_\mu| \geqslant \gamma'/2 \text{ for 1-bounded } P \right] \geqslant 4(\gamma + 3\eta/\alpha) - 10\eta/\alpha \geqslant \gamma,$$

which contradicts that f is $(n^{1/4}, \gamma)$ product pseudorandom. $\qquad\square$

In the following arguments in this section and the next, we will heavily leverage product pseudorandomness with respect to several different distributions on [3]. To formally set this up, we define the following large class of distributions.

Definition 5.6. Let $\mathcal{Q}$ be the set of all distributions ν on a finite alphabet Σ such that $\nu(x)$ is a positive rational number with denominator at most 2^{1000} for all $x \in \Sigma$.

The idea is that by a union bound, Lemma 5.5 implies that harsh random restrictions of E_1, E_2 will be product pseudorandomn against all $\nu \in \mathcal{Q}$ with probability at least $1 - O(\gamma')$.

Finally we random restrict the conclusion of Lemma 5.3 to obtain a 3-wise correlation against a distribution whose mass on all atoms is $\Omega(1)$. Let $\xi := \xi^{(ij)'}$ be as guaranteed in Lemma 5.3.

Lemma 5.7. Let μ be the uniform distribution on [3], and let ν on $[3]^3$ be the distribution with mass $1/6$ on $(0, 0, 0)$ and $(0, 1, 2)$, and mass $1/3$ on $(1, 1, 1)$ and $(2, 2, 2)$. Let S, E_1, E_2 satisfying that $\mu(E_1) = \delta_1$, $\mu(E_2) = \delta_2$, and $\mu(S) \geqslant \alpha\mu(E_1 \boxtimes E_2)$. Assume that $1_{E_1} - \delta_1$ and $1_{E_2} - \delta_2$ are $(n^{1/4}, \gamma)$ product pseudorandom, and let $\widetilde{\tau} > 0$. If S has no combinatorial line, at least one of the following holds. There is some $S' \subseteq E_1' \boxtimes E_2' \subseteq [3]^{n'}$ satisfying either
(Case 1):

1. $n' \geqslant n^{2/5}$, $\mu(S') \geqslant (\alpha + \alpha^3\widetilde{\tau}/1000)\mu(E_1' \boxtimes E_2')$, and $\mu(E_1' \boxtimes E_2') \geqslant \delta_1\delta_2/2$,

2. If S' has a combinatorial line, then S does too,

or **(Case 2):**

1. $n' \geqslant n^{2/5}$, $\mu(S') \geqslant (\alpha - \widetilde{\tau})\mu(E_1' \boxtimes E_2')$, and $\mu(E_1' \boxtimes E_2') \geqslant \delta_1\delta_2/2$,

2. $1_{E_1'} - \delta_1$ and $1_{E_2'} - \delta_2$ are $(n^{1/3}, \gamma')$ product pseudorandom for $\gamma' = 2(\gamma + 1000\eta)^{1/3}$ with respect to all distributions $\mu' \in \mathcal{Q}$ supported on $\{0, 1\}$ or $\{0, 2\}$,

3. $\mathbb{E}_{(x,y,z) \sim \nu^{\otimes n'}} \left[(1_{S'} - \alpha 1_{E_1' \boxtimes E_2'})(x) 1_{S'}(y) 1_{S'}(z) \right] \geqslant \alpha^3\delta_1^2\delta_2^2/4$.

4. If S' has a combinatorial line, then S does too.

Proof. Let $\beta = 6\xi((0, 1, 2)) \leqslant 20\eta/\sqrt{n}$, and let $\xi = \beta\nu + (1 - \beta)\nu'$ where ν' is supported on $(0, 0, 0)$, $(1, 1, 1)$, $(2, 2, 2)$. Note that $|\nu'(x) - 1/3| \leqslant 100\eta/\sqrt{n}$ for all $x \in [3]$. Let $I \sim_{1-\beta} [n]$, $z \sim (\nu')^{\otimes I}$, and let $S' = S_{I \to z}$, $E_1' = (E_1)_{I \to z}$, and $E_2' = (E_2)_{I \to z}$. Note that $n - |I| \geqslant n^{2/5}$ with high probability. We first handle the case when $\mu(S')$ drops significantly under random restriction with nonnegligible probability.

Note that with probability at least $1 - O(\gamma')$ it holds that $1_{E_1'} - \delta_1$ and $1_{E_2'} - \delta_2$ are $(n^{1/3}, \gamma')$ product pseudorandom, by Lemma 5.5. In one case,

$$\Pr \left[\mu(S') \leqslant (\alpha - \widetilde{\tau}/2)\delta_1\delta_2 \right] \geqslant \alpha^3/100. \tag{2}$$

12

We combine this with (using Lemma 4.3)

$$\mathbb{E}[\mu(S')] = \xi_1(S) \geqslant \mu(S) - \mathrm{d_{TV}}(\mu^{\otimes n}, \xi_1^{\otimes n})$$
$$\geqslant \alpha\mu(E_1 \boxtimes E_2) - 200\eta \geqslant \alpha\delta_1\delta_2 - O((\gamma')^c) - 200\eta,$$

to get that there is some S' with

$$\mu(S') \geqslant (\alpha + \alpha^3\widetilde{\tau}/500)\delta_1\delta_2 - O((\gamma')^c) - 200\eta \geqslant (\alpha + \alpha^3\widetilde{\tau}/1000)\mu(E_1' \boxtimes E_2'),$$

as desired, where we have applied Lemma 4.3 again.

We now consider the other case (when (2) does not hold). In this case let $\mathcal{E}$ be the event that $1_{E_1'} - \delta_1$ and $1_{E_2'} - \delta_2$ are $(n^{1/3}, \gamma')$ product pseudorandom against all distributions in $\mathcal{Q}$. By Lemma 5.5, $\Pr[\mathcal{E}] \geqslant 1 - O(\gamma')$. When $\mathcal{E}$ holds, Lemma 5.8 proven below gives

$$\left| \mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes\overline{I}}} \left[(1_{S'} - \alpha 1_{E_1'\boxtimes E_2'})(x)1_{S'}(y)1_{S'}(z) \right] \right| \leqslant \mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes\overline{I}}} \left[1_{E_1'\boxtimes E_2'}(y)1_{E_1'\boxtimes E_2'}(z) \right] \leqslant 2\delta_1^2\delta_2^2. \quad (3)$$

Because $\Pr[\mathcal{E}] \geqslant 1 - O(\gamma')$, applying Lemmas 4.3 and 5.3 gives

$$\mathop{\mathbb{E}}_{\substack{S',E_1',E_2' \\ \mathcal{E} \text{ holds}}} \mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes n'}} \left[(1_{S'} - \alpha 1_{E_1'\boxtimes E_2'})(x)1_{S'}(y)1_{S'}(z) \right] \geqslant \alpha^3\delta_1^2\delta_2^2/2 - O(\gamma').$$

Combining the previous equation with (3) and an averaging argument (over $\mathcal{E}$) gives

$$\mathop{\Pr}_{S',E_1',E_2'} \left[\mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes n'}} \left[(1_{S'} - \alpha 1_{E_1'\boxtimes E_2'})(x)1_{S'}(y)1_{S'}(z) \right] \geqslant \alpha^3\delta_1^2\delta_2^2/4 \right] \geqslant \alpha^3/20. \quad (4)$$

We turn to verifying the conditions of the lemma. By a union bound, $\mu(S') \geqslant (\alpha - \widetilde{\tau}/2)\delta_1\delta_2$, (4), and $\mathcal{E}$ all hold for some S', E_1', E_2'. When this holds, $\mu(S') \geqslant (\alpha - \widetilde{\tau}/2)\delta_1\delta_2 \geqslant (\alpha - \widetilde{\tau})\mu(E_1' \boxtimes E_2')$ and $\mu(E_1' \boxtimes E_2') \geqslant \delta_1\delta_2/2$, verifying condition 1. Condition 2 holds because E_1' and E_2' are chosen so that $\mathcal{E}$ holds. Condition 3 holds because (4) holds. Condition 4 holds because $S' = S_{I\to z}$. $\quad\square$

We now prove Lemma 5.8.

Lemma 5.8. *If $E_1 \subseteq \{0,1\}^n$ and $E_2 \subseteq \{0,2\}^n$ satisfy that $1_{E_1} - \delta_1$ and $1_{E_2} - \delta_2$ are $(n^{0.99}, \gamma')$ product pseudorandom with respect to ν, then*

$$\mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes n}} [1_{E_1\boxtimes E_2}(y)1_{E_1\boxtimes E_2}(z)] \leqslant \delta_1^2\delta_2^2 + O(\log(1/\gamma')^{-c}).$$

Proof. We can rewrite

$$\mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes n}} [1_{E_1\boxtimes E_2}(y)1_{E_1\boxtimes E_2}(z)] = \mathop{\mathbb{E}}_{(x,y,z)\sim\nu^{\otimes n}} [1_{E_1}(y)1_{E_2}(y)1_{E_1}(z)1_{E_2}(z)].$$

The support of the distribution $(\pi_1(y), \pi_2(y), \pi_1(z), \pi_2(z))$ is

$$\{(0,0,0,0), (1,0,1,0), (0,2,0,2), (1,0,0,2)\}.$$

It can be checked that the projections to the coordinates 234 and 123 are connected. Thus, the result follows expanding $1_{E_1} = \delta_1 + (1_{E_1} - \delta_1)$, $1_{E_2} = \delta_2 + (1_{E_2} - \delta_2)$, and applying Theorem 3. $\quad\square$

6 Density Increment

6.1 Obtaining a Large "Box Norm"

In Lemma 5.7 we proved that if S has no combinatorial lines, then after random restriction, the function $1_S - \alpha 1_{E_1 \boxtimes E_2}$ has a nontrivial 3-wise correlation over the $\mathsf{DHJ}[3]$ distribution. In this section, we will use this to obtain a density increment. To start we apply a sequence of Cauchy-Schwarz manipulations to prove the following.

Theorem 4. *Let μ be a distribution supported on $(0,0,0), (1,1,1), (2,2,2), (0,1,2)$ whose mass on each atom is a rational number with denominator at most 1000, so that $\mu_x \in \mathcal{Q}$. Let $f, g, h : [3]^n \to \mathbb{R}$ be 1-bounded functions supported on $E_1 \boxtimes E_2$, and $\delta_1, \delta_2 > 0$ such that $1_{E_1} - \delta_1, 1_{E_2} - \delta_2$ are $(\gamma' n, \gamma')$ product pseudorandom with respect to every measure $\nu \in \mathcal{Q}$. Also,*

$$\left| \mathop{\mathbb{E}}_{(x,y,z) \sim \mu^{\otimes n}} \left[f(x) g(y) h(z) \right] \right| \geqslant \alpha^3 \delta_1^2 \delta_2^2 / 4.$$

Then there is a distribution μ_2 supported on $[3]^4$ (see Table 2) such that the marginal of μ_2 onto any variable is μ_x and

$$\mathop{\mathbb{E}}_{(x,x',x'',x''') \sim \mu_2^{\otimes n}} \left[f(x) f(x') f(x'') f(x''') \right] \geqslant \alpha^{12} \delta_1^2 \delta_2^2 / 2^{10}. \tag{5}$$

Tables 1 to 3 contain distributions that we encounter in the proof.

x	y	x'	y'	z
0	0	0	0	0
1	1	1	1	1
2	2	2	2	2
0	1	0	1	2
2	2	0	1	2
0	1	2	2	2

Table 1: The support of the distribution μ_1

x	x'	x''	x'''	y	y'	z	z'
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
0	0	0	0	1	1	2	2
0	2	0	2	1	2	2	2
2	0	2	0	2	1	2	2
1	1	0	0	1	1	1	2
0	0	1	1	1	1	2	1

Table 2: The support of the distribution μ_2

14

$\pi_1(y)$	$\pi_1(y')$	$\pi_1(y'')$	$\pi_1(y''')$	$\pi_2(z)$	$\pi_2(z')$	$\pi_2(z'')$	$\pi_2(z''')$
0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0
0	0	0	0	2	2	2	2
1	1	1	1	2	2	2	2
1	0	1	0	2	2	2	2
0	1	0	1	2	2	2	2
1	1	1	1	0	2	0	2
1	1	1	1	2	0	2	0
0	0	1	1	0	0	2	2
1	1	0	0	2	2	0	0

Table 3: The support of the distribution μ_3

Proof. Throughout this proof we write $E_1(x)$ and $E_2(x)$ in place of $1_{E_1}(x)$ and $1_{E_2}(x)$. Applying the Cauchy-Schwarz inequality gives

$$\alpha^6 \delta_1^4 \delta_2^4 / 16 \leqslant \Big(\mathop{\mathbb{E}}_{z \sim \mu_z^{\otimes n}} h(z) E_1(z) E_2(z) \mathop{\mathbb{E}}_{x,y|z} f(x) g(y) E_1(z) E_2(z) \Big)^2$$

$$\leqslant \Big(\mathop{\mathbb{E}}_{z \sim \mu_z^{\otimes n}} E_1(z) E_2(z) \Big) \Big(\mathop{\mathbb{E}}_{(z,x,y,x',y') \sim \mu_1^{\otimes n}} f(x) f(x') g(y) g(y') E_1(z) E_2(z) \Big)$$

$$\leqslant 1.1 \delta_1 \delta_2 \Big(\mathop{\mathbb{E}}_{(z,x,y,x',y') \sim \mu_1^{\otimes n}} f(x) f(x') g(y) g(y') E_1(z) E_2(z) \Big),$$

by Lemma 4.3. Square again and apply Cauchy-Schwarz to get

$$\alpha^{12} \delta_1^8 \delta_2^8 / 256 \leqslant 1.3 \delta_1^2 \delta_2^2 \Big(\mathop{\mathbb{E}}_{y,y' \sim (\mu_1)_{yy'}^{\otimes n}} g(y) g(y') E_1(y) E_2(y) E_1(y') E_2(y') \mathop{\mathbb{E}}_{x,x',z|y,y'} f(x) f(x') \prod_{u \in \{y,y',z\}} E_1(u) E_2(u) \Big)^2$$

$$\leqslant 1.3 \delta_1^2 \delta_2^2 \Big(\mathop{\mathbb{E}}_{y,y' \sim (\mu_1)_{yy'}^{\otimes n}} E_1(y) E_2(y) E_1(y') E_2(y') \Big)$$

$$\Big(\mathop{\mathbb{E}}_{(x,x',x'',x''',y,y',z,z') \sim \mu_2^{\otimes n}} f(x) f(x') f(x'') f(x''') \prod_{u \in \{x,x',x'',x''',y,y',z,z'\}} E_1(u) E_2(u) \Big).$$

By Lemma 6.1 below (whose proof uses product pseudorandomness of E_1 and E_2), the first term is bounded by $1.1 \delta_1^2 \delta_2^2$. Also, by examining Table 2 we can check that

$$\prod_{u \in \{x,x',x'',x''',y,y',z,z'\}} E_1(u) E_2(u) = E_1(x) E_2(x) E_2(x') E_1(x'') E_1(y) E_1(y') E_2(z) E_2(z').$$

Now define

$$\omega(x,x',x'',x''') := \mathop{\mathbb{E}}_{y,y',z,z'|x,x',x'',x'''} E_1(y) E_1(y') E_2(z) E_2(z'),$$

where the expectation is over (y,y',z,z') from μ_2 conditioned on x,x',x'',x'''. By Lemma 6.2 below,

15

we can write:

$$\alpha^{12}\delta_1^4\delta_2^4/2^9 \leqslant \mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}} f(x)f(x')f(x'')f(x''')\omega(x,x',x'',x''')$$

$$\leqslant \delta_1^2\delta_2^2 \mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}} \left[f(x)f(x')f(x'')f(x''') \right] + \mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}} \left[|\omega(x,x',x'',x''') - \delta_1^2\delta_2^2| \right]$$

$$\leqslant \delta_1^2\delta_2^2 \mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}} \left[f(x)f(x')f(x'')f(x''') \right] + O((\log(1/\gamma'))^{-c}).$$

Rearranging this implies the claim in the lemma. $\qquad\square$

Now we establish the two lemmas that were required for Theorem 4.

Lemma 6.1. *It holds that*

$$\mathop{\mathbb{E}}_{y,y'\sim(\mu_1)_{yy'}^{\otimes n}} E_1(y)E_2(y)E_1(y')E_2(y') \leqslant \delta_1^2\delta_2^2 + O((\log(1/\gamma'))^{-c}).$$

Proof. Write

$$\mathop{\mathbb{E}}_{y,y'\sim(\mu_1)_{yy'}^{\otimes n}} E_1(y)E_2(y)E_1(y')E_2(y') = \mathop{\mathbb{E}}_{y,y'\sim(\mu_1)_{yy'}^{\otimes n}} \delta_1 E_2(y)E_1(y')E_2(y')$$

$$+ \mathop{\mathbb{E}}_{y,y'\sim(\mu_1)_{yy'}^{\otimes n}} (E_1 - \delta_1)(y)E_2(y)E_1(y')E_2(y').$$

We claim that the term on the second line is bounded by $O((\log(1/\gamma'))^{-c})$. This follows from Theorem 3 once we verify that the corresponding 4-ary distribution

$$(\pi_1(y), \pi_2(y), \pi_1(y'), \pi_2(y')) \in \{0,1\} \times \{0,2\} \times \{0,1\} \times \{0,2\}$$

for $(y, y') \sim (\mu_1)_{yy'}$ is pairwise-connected and has connected marginals onto any 3 of the coordinates, because $E_1 - \delta_1$ is product pseudorandom. By examining Table 1, $(\pi_1(y), \pi_2(y), \pi_1(y'), \pi_2(y')) \in \{0,1\} \times \{0,2\} \times \{0,1\} \times \{0,2\}$ has $\Omega(1)$ mass on each of

$$\{(0,0,0,0), (1,0,1,0), (0,2,0,2), (0,2,1,0), (1,0,0,2))\}.$$

It can be checked that the projection of the distribution to any 3 coordinates is connected. Now the lemma follows from repeating the decomposition three more times. $\qquad\square$

The following lemma proves that $\omega(x,x',x'',x''')$ concentrates around $\delta_1^2\delta_2^2$.

Lemma 6.2. *It holds that*

$$\mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}} \left[|\omega(x,x',x'',x''') - \delta_1^2\delta_2^2| \right] \leqslant O((\log(1/\gamma'))^{-c}).$$

Proof. By Cauchy-Schwarz, it suffices to bound

$$\mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}}\left[|\omega(x,x',x'',x''')-\delta_1^2\delta_2^2|^2\right]$$

$$=\mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}}\left[\left|\mathop{\mathbb{E}}_{y,y',z,z'|x,x',x'',x'''}E_1(y)E_1(y')E_2(z)E_2(z')-\delta_1^2\delta_2^2\right|^2\right]$$

$$=\mathop{\mathbb{E}}_{(x,x',x'',x''')\sim\mu_2^{\otimes n}}\left[\left|\sum_{\substack{\overline{E}_1,\overline{E}_1'\in\{\delta_1,E_1-\delta_1\}\\\overline{E}_2,\overline{E}_2'\in\{\delta_2,E_2-\delta_2\}\\(\overline{E}_1,\overline{E}_1',\overline{E}_2,\overline{E}_2')\neq(\delta_1,\delta_1,\delta_2,\delta_2)}}\mathop{\mathbb{E}}_{y,y',z,z'|x,x',x'',x'''}\overline{E}_1(y)\overline{E}_1'(y')\overline{E}_2(z)\overline{E}_2'(z')\right|^2\right]$$

$$\leqslant 15\max_{x,x',x'',x'''}\mathop{\mathbb{E}}_{\substack{y,y',y'',y'''\\z,z',z'',z'''|x,x',x'',x'''}}\overline{E}_1(y)\overline{E}_1(y'')\overline{E}_1'(y')\overline{E}_1'(y''')\overline{E}_2(z)\overline{E}_2(z'')\overline{E}_2'(z')\overline{E}_2'(z''')$$

$$=15\max\mathop{\mathbb{E}}_{\substack{y,y',y'',y'''\\z,z',z'',z'''\sim\mu_3^{\otimes n}}}\overline{E}_1(y)\overline{E}_1(y'')\overline{E}_1'(y')\overline{E}_1'(y''')\overline{E}_2(z)\overline{E}_2(z'')\overline{E}_2'(z')\overline{E}_2'(z'''),$$

where μ_3 is the distribution in Table 3. In the previous expression we mean that $y,y',y'',y'''\in\{0,1\}$ and $z,z',z'',z'''\in\{0,2\}$ because the y's and z's are in $\overline{E}_1$ and $\overline{E}_2$ terms respectively. Also, the max is over all $\overline{E}_1,\overline{E}_1'\in\{\delta_1,E_1-\delta_1\}$ and $\overline{E}_2,\overline{E}_2'\in\{\delta_2,E_2-\delta_2\}$, except for $\overline{E}_1=\overline{E}_1'=\delta_1$ and $\overline{E}_2=\overline{E}_2'=\delta_2$. Thus, the first inequality in the above display follows by expanding $E_1=\delta_1+(E_1-\delta_1)$ and $E_2=\delta_2+(E_2-\delta_2)$ and applying Cauchy-Schwarz. We consider the case where $\overline{E}_1=E_1-\delta_1$ – the remaining cases are similar.

$$\mathop{\mathbb{E}}_{\substack{y,y',y'',y'''\\z,z',z'',z'''\sim\mu_3^{\otimes n}}}\overline{E}_1(y)\overline{E}_1(y'')\overline{E}_1'(y')\overline{E}_1'(y''')\overline{E}_2(z)\overline{E}_2(z'')\overline{E}_2'(z')\overline{E}_2'(z''')$$

$$\leqslant\mathop{\mathbb{E}}_{\substack{y',y'''\\z,z',z'',z'''\sim\mu_3^{\otimes n}}}\left|\mathop{\mathbb{E}}_{y,y''|y',y''',z,z',z'',z'''}\overline{E}_1(y)\overline{E}_1(y'')\right|$$

$$\leqslant\left(\mathop{\mathbb{E}}_{\substack{y',y'''\\z,z',z'',z'''\sim\mu_3^{\otimes n}}}\left|\mathop{\mathbb{E}}_{y,y''|y',y''',z,z',z'',z'''}\overline{E}_1(y)\overline{E}_1(y'')\right|^2\right)^{1/2}$$

$$=\left(\mathop{\mathbb{E}}_{\substack{y',y'''\\z,z',z'',z'''\sim\mu_3^{\otimes n}}}\mathop{\mathbb{E}}_{\substack{y,y''\\\widetilde{y},\widetilde{y}''|y',y''',z,z',z'',z'''}}\overline{E}_1(y)\overline{E}_1(y'')\overline{E}_1(\widetilde{y})\overline{E}_1(\widetilde{y}'')\right)^{1/2}$$

$$=\left(\mathop{\mathbb{E}}_{(y,y'',\widetilde{y},\widetilde{y}'')\sim\mu_4^{\otimes n}}\overline{E}_1(y)\overline{E}_1(y'')\overline{E}_1(\widetilde{y})\overline{E}_1(\widetilde{y}'')\right)^{1/2}, \tag{6}$$

for a distribution μ_4 over 4-tuples $(y,y'',\widetilde{y},\widetilde{y}'')$ that has mass $\Omega(1)$ on:

$$\{(0,0,0,0),(1,1,1,1),(0,1,0,1),(1,0,1,0),(0,0,1,1),(1,1,0,0)\},$$

which can be seen by inspecting Table 3. Specifically, $(0,0,0,0)$, $(1,1,1,1)$, $(0,1,0,1)$, and $(1,0,1,0)$ come from using rows $1,2,9,10$ of Table 3 respectively, and $(0,0,1,1)$, $(1,1,0,0)$ come from combining rows 4 and 6. It can be checked that the projection of the distribution to any 3 coordinates is

17

connected and hence satisfies the hypotheses of Theorem 3, and thus the quantity in (6) is bounded by $O((\log(1/\gamma'))^{-c})$ because $\overline{E}_1$ is product pseudorandom. One can check that the marginals of μ_2 are all μ_x due to how the sequence of Cauchy-Schwarz manipulations were applied. This completes the proof. $\qquad\square$

6.2 Random Restrictions for a Density Increment

In this subsection we will take random restrictions of (5) to obtain a density increment, for $f := S' - \alpha E_1' \boxtimes E_2'$. Note that f is nonzero on only about a $\mu(E_1' \boxtimes E_2') \approx \delta_1 \delta_2$ fraction of x. Below we argue that when we restrict to these x that the measures of $S'(x''')$ and $E_1' \boxtimes E_2'(x''')$ are not affected much.

Lemma 6.3. *Let μ be the marginal of μ_2 onto any of x, x', x'', x'''. Consider $S \subseteq E_1 \boxtimes E_2 \subseteq [3]^n$ such that $\mu(E_1) = \delta_1$, $\mu(E_2) = \delta_2$, such that $E_1 - \delta_1$ and $E_2 - \delta_2$ are $(\gamma' n, \gamma')$ product pseudorandom. Then*

$$\left| \mathbb{E}_{(x,x''')\sim\mu_2^{\otimes n}} \left[E_1(x) E_2(x) S(x''') \right] - \delta_1 \delta_2 \mu(S) \right| \leqslant O(\log(1/\gamma')^{-c}),$$

and

$$\left| \mathbb{E}_{(x,x''')\sim\mu_2^{\otimes n}} \left[E_1(x) E_2(x) E_1(x''') E_2(x''') \right] - \delta_1^2 \delta_2^2 \right| \leqslant O(\log(1/\gamma')^{-c}).$$

Proof. By inspecting μ_2, the corresponding distribution $(\pi_1(x), \pi_2(x), x''')$ is supported on

$$\{(0,0,0), (1,0,1), (0,2,2), (0,0,2), (0,2,0), (1,0,0), (0,0,1)\},$$

which can easily be seen to be pairwise-connected. Thus the first result follows from Theorem 2, after expanding $E_1 = \delta_1 + (E_1 - \delta_1)$ and $E_2 = \delta_2 + (E_2 - \delta_2)$.

For the second result, the corresponding distribution $(\pi_1(x), \pi_2(x), \pi_1(x'''), \pi_2(x'''))$ has support

$$\{(0,0,0,0), (1,0,1,0), (0,2,0,2), (0,0,0,2), (0,2,0,0), (1,0,0,0), (0,0,1,0)\},$$

which is again easily seen to be connected when restricted to any 3 coordinates. Thus the result follows from Theorem 3 after expanding $E_1 = \delta_1 + (E_1 - \delta_1)$ and $E_2 = \delta_2 + (E_2 - \delta_2)$. $\qquad\square$

Finally we show how to obtain a density increment if S has no combinatorial lines.

Theorem 5. *Let μ denote the uniform distribution on $[3]$. Let $S \subseteq E_1 \boxtimes E_2 \subseteq [3]^n$ such that $\mu(E_1) = \delta_1$, $\mu(E_2) = \delta_2$, and $\mu(S) = \alpha \mu(E_1 \boxtimes E_2)$. Assume that $1_{E_1} - \delta_1$ and $1_{E_2} - \delta_2$ are $(n^{1/4}, \gamma)$ product pseudorandom, and that S has no combinatorial lines. Then there is some $S' \subseteq E_1' \boxtimes E_2' \subseteq [3]^{n'}$ satisfying:*

1. *$n' \geqslant n^{1/3}$,*

2. *$\mu(S') \geqslant (\alpha + \Omega(\alpha^{15}))\mu(E_1' \boxtimes E_2')$ and $\mu(E_1' \boxtimes E_2') \geqslant \Omega(\alpha^{12} \delta_1 \delta_2)$,*

3. *If S' has a combinatorial line, then S does too.*

Proof. Set $\widetilde{\tau} = c\alpha^{12}$ for sufficiently small constant c. Now we apply Lemma 5.7. If Case 1 of Lemma 5.7 holds, then we are done by the choice of $\widetilde{\tau}$. Otherwise, Case 2 holds. In this case we may apply Theorem 4 to get

$$\mathbb{E}_{(x,x',x'',x''')\sim(\mu_2)^{\otimes n'}} \left[f(x) f(x') f(x'') f(x''') \right] \geqslant \alpha^{12} \delta_1^2 \delta_2^2 / 2^{10},$$

18

for $f = 1_{S'} - \alpha 1_{E'_1 \boxtimes E'_2}$ for some $\mu(S') \geqslant (\alpha - \widetilde{\tau})\mu(E'_1 \boxtimes E'_2)$. Let ν be uniform over $(0,0,0,0)$, $(0,2,0,2)$, $(0,0,1,1)$ and write $\mu_2 = \beta\nu + (1-\beta)\nu'$ for some $\beta = \Omega(1)$ and a distribution ν'. Note that ν is equivalent to $(0, \pi_2(y), \pi_1(y), y)$ for $y \sim [3]$.

Taking this random restriction gives

$$\mathop{\mathbb{E}}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I}}} \mathop{\mathbb{E}}_{y \sim [3]^{([n'] \backslash I)}} \left[f_{I \to z}(0) f_{I \to z'}(\pi_2(y)) f_{I \to z''}(\pi_1(y)) f_{I \to z'''}(y) \right] \geqslant \alpha^{12} \delta_1^2 \delta_2^2 / 2^{10}.$$

Let $\mathcal{E}$ be the event that $f_{I \to z}(0) \neq 0$, i.e., that $\pi_1(z,0) \in E'_1$ and $\pi_2(z,0) \in E'_2$, where by $(z,0) \in [3]^{n'}$ we mean that z is assigned to the coordinates in I, and 0 is assigned to the coordinates in $\overline{I}$.

Recall that Lemma 4.3 gives us $|\mu(E'_1 \boxtimes E'_2) - \delta_1\delta_2| \leqslant O((\gamma')^c)$. Thus, conditioning on $\mathcal{E}$ and using $|f_{I \to z}(0)| \leqslant 1$ gives

$$\mathop{\mathbb{E}}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \left| \mathop{\mathbb{E}}_{y \sim [3]^{([n'] \backslash I)}} \left[f_{I \to z'}(\pi_2(y)) f_{I \to z''}(\pi_1(y)) f_{I \to z'''}(y) \right] \right| \geqslant \alpha^{12} \delta_1 \delta_2 / 2^{11}. \qquad (7)$$

For the following sequence of equations, the first equality is by the definition of conditioned on $\mathcal{E}$, the second equality is because μ is the uniform measure, the third equality is by properties of random restriction, and the final inequalities apply Lemma 6.3 and $|\Pr[\mathcal{E}] - \delta_1\delta_2| \leqslant O(\gamma^c)$:

$$\mathop{\mathbb{E}}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \left[\mu(S'_{I \to z'''}) - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu((E'_1)_{I \to z'''} \boxtimes (E'_2)_{I \to z'''}) \right]$$

$$= \frac{1}{\Pr[\mathcal{E}]} \mathop{\mathbb{E}}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I}}} \left[(E'_1)_{I \to z}(0)(E'_2)_{I \to z}(0) \Big(\mu(S'_{I \to z'''}) - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu((E'_1)_{I \to z'''} \boxtimes (E'_2)_{I \to z'''}) \Big) \right]$$

$$= \frac{1}{\Pr[\mathcal{E}]} \mathop{\mathbb{E}}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I}}} \mathop{\mathbb{E}}_{(y,y',y'',y''') \sim \nu^{\otimes \overline{I}}} \left[(E'_1)_{I \to z}(0)(E'_2)_{I \to z}(0) S'_{I \to z'''}(y''') \right.$$

$$\left. - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})(E'_1)_{I \to z}(0)(E'_2)_{I \to z}(0)(E'_1)_{I \to z'''}(y''')(E'_2)_{I \to z'''}(y''') \right]$$

$$= \frac{1}{\Pr[\mathcal{E}]} \mathop{\mathbb{E}}_{(x,x''') \sim \mu_2^{\otimes n}} \left[E'_1(x)E'_2(x)S'(x''') - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})E'_1(x)E'_2(x)E'_1(x''')E'_2(x''') \right]$$

$$\geqslant -\alpha^{12}\delta_1\delta_2/2^{14} - O((\log(1/\gamma')^{-c})) \geqslant -\alpha^{12}\delta_1\delta_2/2^{13}.$$

Looking at the expectation on the left hand side of (7), we observe that the first two terms are analogous to new disjoint products that may be correlated with a restriction of f. Indeed, we note that as f receives two values in $E'_1 \boxtimes E'_2$, the first two terms form a partition of $E'_1 \boxtimes E'_2$ into four disjoint products, and (7) asserts that f is correlated with this partition.

We now formalize this idea. Note that $f(x) \in \{-\alpha, 0, 1-\alpha\}$, with $f(x) \neq 0$ if and only if $x \in E'_1 \boxtimes E'_2$. For fixed (z', z'') define

$$F_1^+ := \{y \in \{0,1\}^{[n] \backslash I} : f_{I \to z''}(y) = 1 - \alpha\} \quad \text{and} \quad F_1^- := \{y \in \{0,1\}^{[n] \backslash I} : f_{I \to z''}(y) = -\alpha\}.$$

Define

$$F_2^+ := \{y \in \{0,2\}^{[n] \backslash I} : f_{I \to z'}(y) = 1 - \alpha\} \quad \text{and} \quad F_2^- := \{y \in \{0,2\}^{[n] \backslash I} : f_{I \to z'}(y) = -\alpha\}.$$

By examining μ_2 one can check that $F_1^+ \cup F_1^- = (E_1')_{I \to z'''}$ an $F_2^+ \cup F_2^- = (E_2')_{I \to z'''}$.

Now the above equation gives us:

$$\mathbb{E}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \mathbb{E}_{\substack{F_1 \in \{F_1^+, F_1^-\} \\ F_2 \in \{F_2^+, F_2^-\}}} \left[\mu(S'_{I \to z'''} \cap (F_1 \boxtimes F_2)) - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu(F_1 \boxtimes F_2) \right]$$

$$= \frac{1}{4} \mathbb{E}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \left[\mu(S'_{I \to z'''}) - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu((E_1')_{I \to z'''} \boxtimes (E_2')_{I \to z'''}) \right]$$

$$\geqslant -\alpha^{12}\delta_1\delta_2/2^{15}. \tag{8}$$

$f_{I \to z''}$ is constant on any $F_1 \in \{F_1^+, F_1^-\}$, and we denote its value by $f_{I \to z''}(F_1)$. We similarly define $f_{I \to z'}(F_2)$. Applying the triangle inequality, and then (7) and Lemma 6.3 gives

$$\mathbb{E}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \mathbb{E}_{\substack{F_1 \in \{F_1^+, F_1^-\} \\ F_2 \in \{F_2^+, F_2^-\}}} \left| \mu(S'_{I \to z'''} \cap (F_1 \boxtimes F_2)) - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu(F_1 \boxtimes F_2) \right|$$

$$\geqslant \mathbb{E}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \left| \mathbb{E}_{\substack{F_1 \in \{F_1^+, F_1^-\} \\ F_2 \in \{F_2^+, F_2^-\}}} f_{I \to z'}(F_2) f_{I \to z''}(F_1) \left(\mu(S'_{I \to z'''} \cap (F_1 \boxtimes F_2)) - \alpha \cdot \mu(F_1 \boxtimes F_2) \right) \right|$$

$$- \frac{\alpha^{12}}{2^{14}} \mathbb{E}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \mathbb{E}_{\substack{F_1 \in \{F_1^+, F_1^-\} \\ F_2 \in \{F_2^+, F_2^-\}}} \mu(F_1 \boxtimes F_2)$$

$$= \frac{1}{4} \mathbb{E}_{\substack{I \sim_{1-\beta} [n'] \\ (z,z',z'',z''') \sim (\nu')^{\otimes I} \\ \mathcal{E} \text{ holds}}} \left| \mathbb{E}_{y \sim [3]^{([n'] \setminus I)}} \left[f_{I \to z'}(\pi_2(y)) f_{I \to z''}(\pi_1(y)) f_{I \to z'''}(y) \right] \right|$$

$$- \frac{1}{\Pr[\mathcal{E}]} \cdot \alpha^{12}/2^{16} \cdot \mathbb{E}_{(x,x''') \sim \mu_2^{\otimes n}} \left[E_1(x) E_2(x) E_1(x''') E_2(x''') \right]$$

$$\geqslant \alpha^{12}\delta_1\delta_2/2^{13} - \alpha^{12}\delta_1\delta_2/2^{15} \geqslant \alpha^{12}\delta_1\delta_2/2^{14}. \tag{9}$$

Consider the random variable, defined on $I \sim_{1-\beta} [n']$, $(z, z', z'', z''') \sim (\nu')^{\otimes I}$, and $F_1 \in \{F_1^+, F_1^-\}$, $F_2 \in \{F_2^+, F_2^-\}$,

$$X = \mu(S'_{I \to z'''} \cap (F_1 \boxtimes F_2)) - (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu(F_1 \boxtimes F_2).$$

(8) and (9) together imply that

$$\mathbb{E}[X + |X|] \geqslant -\alpha^{12}\delta_1\delta_2/2^{15} + \alpha^{12}\delta_1\delta_2/2^{14} = \alpha^{12}\delta_1\delta_2/2^{15}.$$

Hence there is a realization of $X \geqslant 0$ with $X = \frac{1}{2}(X + |X|) \geqslant \alpha^{12}\delta_1\delta_2/2^{16}$. By the definition of X, there is some choice of (z, z', z'', z''') and F_1, F_2 such that

$$\mu(S'_{I \to z'''} \cap (F_1 \boxtimes F_2)) \geqslant (\alpha - \widetilde{\tau} + \alpha^{12}/2^{14})\mu(F_1 \boxtimes F_2) + \Omega(\alpha^{12}\delta_1\delta_2).$$

This implies the conditions because $\beta n' \gg n^{1/3}$, $\widetilde{\tau} = c\alpha^{12}$ for small constant c, and

$$\mu(F_1 \boxtimes F_2) \geqslant \mu(S'_{I \to z'''} \cap (F_1 \boxtimes F_2)) \geqslant \Omega(\alpha^{12}\delta_1\delta_2).$$

$\square$

7 Uniformization

In this section we explain how to start with a subset $S \subseteq E_1 \boxtimes E_2$ of density α, where E_1, E_2 are not necessarily product pseudorandom in the sense of Definition 3.1, and to reach a situation where E_1, E_2 are, without decreasing the measure of S or n by too much. The proof will use two types of random restrictions: the standard one introduced in Definition 1.2 as well as one which fixes many coordinates to take the same value (see Definition 7.2). Note that applying such restrictions cannot create new combinatorial lines.

Theorem 6. *Let $S \subseteq E_1 \boxtimes E_2 \subseteq [3]^n$ with $\mu(E_1 \boxtimes E_2) = \delta$ and $\mu(S) \geqslant (\alpha + \tau)\delta$, and $0 < \gamma < (\delta\tau)^{10}$. There is some $n' \geqslant n^{\exp(-O(\delta^{-1}\tau^{-1}\gamma^{-4}))}$ and $S' \subseteq E_1' \boxtimes E_2' \subseteq [3]^{n'}$ such that:*

1. *$\mu(E_1' \boxtimes E_2') \geqslant \delta\tau/3$.*

2. *$(S', E_1', E_2') \in \mathsf{Struct}_{\alpha+\tau/2}$.*

3. *If S' contains a combinatorial line, then S does too.*

We mimic the proof of the uniformization step in Shkredov's bound for the corners problem [Shk05, Shk06] (also see [Gre04, Gre05a] for notes from which we borrowed notation). The high level idea is that if E_1 (or E_2) is not product function pseudorandom, then we can "partition" $[3]^n$ into smaller sets of the form $E_1' \boxtimes E_2'$ so that the measures of these E_1' have higher variance than before. This process eventually terminates, and we find a partition piece under which S still has large enough density. One step of this process amounts to being able to density increment sets that correlate to product functions (after random restriction), which was exactly done in [BKLM24a, Section 8].

Now we move towards formally proving Theorem 6. Let ξ be some distribution over triples (S', E_1', E_2') with $S' \subseteq E_1' \boxtimes E_2' \subseteq [3]^{n'}$ for some n'. Define the *index* of ξ to be:

$$\mathcal{I}(\xi) := \mathop{\mathbb{E}}_{(S',E_1',E_2')\sim\xi}[\mu(E_1')^2 + \mu(E_2')^2].$$

Note that $\mathcal{I}(\xi) \leqslant 2$ always. The following lemma shows how to perform one round of partitioning.

Lemma 7.1. *Let $S \subseteq E_1 \boxtimes E_2 \subseteq [3]^n$ where $1_{E_1} - \mu(E_1)$ or $1_{E_2} - \mu(E_2)$ is not $(n^{1/4}, \gamma)$ product-pseudorandom. Then there is a distribution ξ over triples (S', E_1', E_2') satisfying, for $c = 1/1000$:*

1. *$S' \subseteq E_1' \boxtimes E_2' \subseteq [3]^{n'}$ for $n' \geqslant n^c$.*

2. *$\left| \mathbb{E}_{(S',E_1',E_2')\sim\xi}[\mu(S')] - \mu(S) \right| \leqslant n^{-c}$.*

3. *$\left| \mathbb{E}_{(S',E_1',E_2')\sim\xi}[\mu(E_i')] - \mu(E_i) \right| \leqslant n^{-c}$ for $i = 1, 2$.*

4. *$\left| \mathbb{E}_{(S',E_1',E_2')\sim\xi}[\mu(E_1' \boxtimes E_2')] - \mu(E_1 \boxtimes E_2) \right| \leqslant n^{-c}$.*

5. *For all $(S', E_1', E_2') \in \mathsf{supp}(\xi)$, if S' has a combinatorial line, then S does too.*

6. *(Increment) $\mathcal{I}(\xi) \geqslant (\mu(E_1)^2 + \mu(E_2)^2) + \gamma^4/2$.*

Condition 1 simply says that the new instances we produce are in a polynomially large number of dimensions. Conditions 2-4 say that the densities of the objects we care about, namely $S, E_1, E_2, E_1 \boxtimes E_2$, only change by a negligible amount. Condition 5 says that if the new instances

21

have combinatorial lines, then the original instance did too. Finally, Condition 6 says that the index of the set has increased.

Beyond standard random restrictions, the proof of Lemma 7.1 also requires another operation to generate the triples (S', E_1', E_2') that we increment onto. This operation takes a subset $T \subseteq [n]$ and forces that $x_t = x_{t'}$ for all $t, t' \in T$.

Definition 7.2. For a function $g : [3]^n \to \mathbb{C}$ and $T \subseteq [n]$, define $g_{=T} : [3]^{n-|T|+1} \to \mathbb{C}$ as follows. For $x \in [3]$ and $y \in [3]^{[n] \setminus T}$ let $v \in [3]^n$ be the vector where $v_i = y_i$ for $i \in [n] \setminus T$ and $v_i = x$ otherwise. Then $g_{=T}(x, y) := g(v)$. For disjoint sets $T_1, \ldots, T_N$ we also write $g_{=T_1, \ldots, T_N} = (\ldots (g_{=T_1}) \ldots)_{=T_N}$.

Our next lemma argues that if T is a random subset of $S \subseteq [n]$ of size much smaller than $\sqrt{|S|}$, then performing the operation described in Definition 7.2 does not significantly affect the TV distance on the uniform distribution. This is more or less a special case of [Pol12, Lemma 6.3], but we provide a proof sketch for this easier case for completeness.

Lemma 7.3. *Let ν be the distribution on $[3]^n$ over the v vector (as defined in Definition 7.2) obtained by sampling $T \subseteq S, |T| = k$ and $x \in [3]^{n-k+1}$ uniformly. Then $\mathrm{d_{TV}}(\nu, \mu) \leqslant \frac{10k}{\sqrt{|S|}}$, where μ is the uniform distribution on $[3]^n$.*

Proof. Further let ν' to be the distribution where we sample $T \subseteq S$ and set $x_t = a$ for all $t \in T$, where a is some fixed element of $[3]$. We will prove that $\mathrm{d_{TV}}(\nu', [3]^n) \leqslant \frac{10k}{\sqrt{|S|}}$. When $k = 1$ the claim follows by a direct calculation. For $k > 1$ note that sampling $T \subseteq S$ can be achieved by sampling $T' \subseteq S$ with $|T| = k - 1$ and then sampling $t \in S \setminus T'$ and setting $T = T' \cup \{t\}$. Thus the result follows by induction. $\square$

We are now in a position to prove Lemma 7.1.

Proof of Lemma 7.1. We consider the case where $1_{E_1} - \mu(E_1)$ is not product pseudorandom – the case of E_2 is similar. Throughout this proof we will use $E_1(x)$ in place of $1_{E_1}(x)$ (and similar for E_2, S). Define the sets $(E_1)_{I \to z} \subseteq \{0,1\}^{[n] \setminus I}$ and $(E_1)_{=T} \subseteq \{0,1\}^{n-|T|+1}$ in the expected way.

Let $f := E_1 - \mu(E_1)$. The first step is to take some $I \subseteq [n]$ with $|I| \leqslant n - n^{1/4}/2$ such that

$$\Pr_{z \sim \mu^{\otimes I}} \left[\exists \{P_i : [3] \to \mathbb{C}, \|P_i\|_\infty \leqslant 1\}_{i \in \overline{I}} \quad \text{with} \quad \left| \underset{x \sim \mu^{\otimes \overline{I}}}{\mathbb{E}} \left[f_{I \to z}(x) \prod_{i \in \overline{I}} P_i(x_i) \right] \right| \geqslant \gamma \right] \geqslant \gamma,$$

which is guaranteed to exist because f is not $(n^{1/4}, \gamma)$ product pseudorandom.

For $z \in [3]^I$ define $S' = S_{I \to z}$, $E_1' = (E_1)_{I \to z}$, $E_2' = (E_2)_{I \to z}$, and $m = n - |I| \geqslant n^{1/4}/2$. Relabel $\overline{I}$ as $[m]$. For γ fraction of such z, there are functions $P_1, \ldots, P_m$ such that:

$$\left| \underset{x \sim \mu^{\otimes m}}{\mathbb{E}} \left[f_{I \to z}(x) \prod_{i=1}^m P_i(x_i) \right] \right| \geqslant \gamma. \tag{10}$$

We focus on such a z and perform further operations to (S', E_1', E_2'). Let $v_1, \ldots, v_m : [3] \to \mathbb{R}/\mathbb{Z}$ be such that $P_j(x) = e^{2\pi i v_j(x)}$ for all $j \in [m], x \in [3]$. Let $\zeta = \frac{1}{72}$, and let $N = m^\zeta$. For $x \in (\mathbb{R}/\mathbb{Z})^n$ let $\|x\|_\infty := \max_{i \in [n]} |x_i|_{\mathbb{R}/\mathbb{Z}}$. By the Pigeonhole principle, we can find disjoint sets $S_1, \ldots, S_N$ of size $|S_i| = \sqrt{m}/2$ such that for every $i \in [N]$ and $j, j' \in S_i$, it holds that $\|v_j - v_{j'}\|_\infty \leqslant m^{-\frac{1}{6}}$. Let v be an arbitrary vector in S_i, and let $1 \leqslant k_i \leqslant m^{\frac{1}{12}}$ be such that $\|k_i v\|_\infty \leqslant m^{-\frac{1}{36}}$ – such a k_i exists by Dirichlet approximation.

22

Now let T_i be a uniformly random subset of S_i of size k_i for all $i \in [N]$ and perform the following sequence of restrictions. For all $i \in [N]$ force that $x_j = x_{j'}$ for all $j, j' \in T_i$, and then uniformly random restrict all coordinates in $J := [m] \setminus (T_1 \cup \cdots \cup T_N)$. Let $P = P_1 \ldots P_m$ and $Q = (P_{=T_1,\ldots,T_N})_{J \to u}$. We will argue that Q is nearly constant. For $x, y \in [3]^N$,

$$Q(x) - Q(y) = \exp\left(2\pi i \sum_{j=1}^{N} \sum_{t \in T_j} v_t(y)\right) \left(\exp\left(2\pi i \sum_{j=1}^{N} \sum_{t \in T_j} (v_t(x) - v_t(y))\right) - 1\right) \prod_{i \in ([m] \setminus \cup_i T_i)} P_i(u_i).$$

For $j \in [N]$ and let $v \in T_j$ be such that $\|k_j v\|_\infty \leqslant m^{-\frac{1}{36}}$. Then

$$\left\|\sum_{t \in T_j} v_t\right\|_\infty \leqslant \|k_j v\|_\infty + \sum_{t \in T_j} \|v_t - v\|_\infty \leqslant m^{-\frac{1}{36}} + k_j m^{-\frac{1}{6}} \leqslant 2m^{-\frac{1}{36}}.$$

Combining this with the above yields that $|Q(x) - Q(y)| \leqslant 100 N m^{-\frac{1}{36}} = 100 m^{-\frac{1}{72}}$.

By Lemma 7.3 we know that

$$\mathop{\mathbb{E}}_{\substack{T_1,\ldots,T_N \\ u \sim [3]^J}} \left| \mathop{\mathbb{E}}_{x \sim [3]^N} \left[(f_{I \to z, J \to u})_{=T_1,\ldots,T_N}(x) Q(x)\right]\right|$$

$$\geqslant \left| \mathop{\mathbb{E}}_{x \sim [3]^m} f_{I \to z}(x) P(x) - \frac{10 \sum_{i=1}^{N} k_i}{(\sqrt{m}/2)^{1/2}} \right| \geqslant \left| \mathop{\mathbb{E}}_{x \sim [3]^m} f_{I \to z}(x) P(x) - \frac{10 N m^{1/12}}{m^{1/4}/2} \right|$$

$$\geqslant \gamma - m^{-1/8}.$$

By using that Q is nearly constant above, the previous inequality becomes:

$$\mathop{\mathbb{E}}_{\substack{T_1,\ldots,T_N \\ u \sim [3]^J}} |\mu(((E_1)_{I \to z, J \to u})_{=T_1,\ldots,T_N}) - \mu(E_1)| \geqslant \gamma - m^{-1/8} - 100 m^{-1/72} \geqslant \gamma - 101 m^{-1/72}. \quad (11)$$

By Lemma 7.3 again we also know that:

$$\left| \mathop{\mathbb{E}}_{\substack{T_1,\ldots,T_N \\ u \sim [3]^J}} \left[\mu(((E_1)_{I \to z, J \to u})_{=T_1,\ldots,T_N})\right] - \mu((E_1)_{I \to z}) \right| \leqslant m^{-1/8}. \quad (12)$$

Finally, define distribution ξ over triples (S'', E_1'', E_2'') as follows. Let $z \sim [3]^I$. If z did not satisfy (10) for any product function $P = P_1 \ldots P_m$ then set $(S'', E_1'', E_2'') = (S', E_1', E_2')$, where recall that $S' = S_{I \to z}$, $E_1' = (E_1)_{I \to z}$, and $E_2' = (E_2)_{I \to z}$. If z satisfies (10) then set $S'' = (S_{I \to z, J \to u})_{=T_1,\ldots,=T_N}$, $E_1'' = ((E_1)_{I \to z, J \to u})_{=T_1,\ldots,=T_N}$, and $E_2'' = ((E_2)_{I \to z, J \to u})_{=T_1,\ldots,=T_N}$ for $T_1, \ldots, T_N$ and u sampled as described above. We verify the conditions 1-6 of the lemma.

Condition 1 follows by construction, as $N := m^{1/72} \geqslant (n^{1/4}/2)^{1/72} \geqslant n^{1/300}$. Conditions 2-4 follow by applications of Lemma 7.3 in the same way as (12). Condition 5 follows by constrction: restrictions and $= T$ operations can only reduce the number of combinatorial lines.

To see Condition 6, we write

$$\mathbb{E}[\mu(E_1'')^2] = \mathbb{E}\left[\left|\mu(E_1'') - \mathbb{E}[\mu(E_1'')]\right|^2\right] + \mathbb{E}[\mu(E_1'')]^2$$

$$\geqslant \left(\gamma(\gamma - 101 m^{-1/72})\right)^2 + (\mu(E_1) - m^{-1/8})^2 \geqslant \mu(E_1)^2 + 0.6\gamma^4,$$

23

where the first inequality uses (11) and the fact that a γ fraction of z's satisfy (10), and (11). Also by Jensen, $\mathbb{E}[\mu(E_2'')^2] \geqslant \mathbb{E}[\mu(E_2'')]^2 \geqslant \mu(E_2)^2 - m^{-1/8}$, where we have used Condition 3. Condition 6 follows by combining these. $\qquad\square$

To prove Theorem 6 we repeatedly apply Lemma 7.1 until the mass of the triples (S', E_1', E_2') that are not product pseudorandom is small.

Proof of Theorem 6. Call a triple (S', E_1', E_2') with $S' \subseteq E_1' \boxtimes E_2' \subseteq [3]^{n'}$ *good* if both $1_{E_1'} - \mu(E_1')$ and $1_{E_2'} - \mu(E_2')$ are $((n')^{1/4}, \gamma)$ product pseudorandom.

Let $\xi^{(0)}$ have mass 1 on (S, E_1, E_2). Do the following algorithm: for $t \geqslant 0$ terminate if

$$\Pr_{(S', E_1', E_2') \sim \xi^{(t)}}[(S', E_1', E_2') \text{ is not good}] \leqslant \delta\tau/100.$$

Otherwise, define $\xi^{(t+1)}$ as follows. Sample $(S', E_1', E_2') \sim \xi^{(t)}$. If (S', E_1', E_2') is good, then keep that mass in $\xi^{(t+1)}$. Otherwise, sample (S'', E_1'', E_2'') from the distribution guaranteed by Lemma 7.1 and put that mass in $\xi^{(t+1)}$. By Lemma 7.1 we know that $\mathcal{I}(\xi^{(t+1)}) \geqslant \mathcal{I}(\xi^{(t)}) + \delta\tau\gamma^4/200$, and hence the algorithm terminates within $T := O(\delta^{-1}\tau^{-1}\gamma^{-4})$ steps.

Let $n' = n^{1000^{-T}}$. By applying Lemma 7.1 inductively we know that

$$\left|\mathbb{E}_{(S', E_1', E_2') \sim \xi^{(T)}}[\mu(S')] - \mu(S)\right| \leqslant T/n' \quad \text{and} \quad \left|\mathbb{E}_{(S', E_1', E_2') \sim \xi^{(T)}}[\mu(E_1' \boxtimes E_2')] - \mu(E_1 \boxtimes E_2)\right| \leqslant T/n'.$$

Thus we get that for $c' = 1000^{-T}$,

$$\mathbb{E}_{(S', E_1', E_2') \sim \xi^{(T)}}\left[1_{(S', E_1', E_2') \text{ good}}\left(\mu(S') - (\alpha + \tau/2)\mu(E_1' \boxtimes E_2')\right)\right]$$

$$\geqslant \mu(S) - (\alpha + \tau/2)\mu(E_1 \boxtimes E_2) - 2Tn^{-c'} - \delta\tau/100 \geqslant \delta\tau/3,$$

for our choice of parameters. Thus there is $(S', E_1', E_2') \in \mathsf{supp}(\xi^{(T)})$ that is good, $\mu(S') \geqslant (\alpha + \tau/2)\mu(E_1' \boxtimes E_2')$. For such choice, $\mu(E_1' \boxtimes E_2') \geqslant \mu(S') \geqslant \delta\tau/3$ as $S' \subseteq E_1' \boxtimes E_2'$. Also, $S' \subseteq [3]^{n'}$, which concludes the proof. $\qquad\square$

Proof of Theorem 1. Assume for contradiction that S has no combinatorial line. Initially, for $E_1 = \{0,1\}^n$ and $E_2 = \{0,2\}^n$ and $|S| = \alpha_0 \cdot 3^n$, it holds that $(S, E_1, E_2) \in \mathsf{Struct}_{\alpha_0}$. If $(S, E_1, E_2) \in \mathsf{Struct}_{\alpha}$ for $S \subseteq [3]^n$ and $\delta := \mu(E_1 \boxtimes E_2)$, then applying Theorems 5 and 6 gives a new $(S', E_1', E_2') \in \mathsf{Struct}_{\alpha+\Omega(\alpha^{15})}$ where $S' \subseteq [3]^{n'}$, and $n' \geqslant n^{1000^{-O(\gamma^{-4}\delta^{-1}\alpha^{-O(1)})}}$. Also if S' has a combinatorial line, then S does too and $\mu(E_1' \boxtimes E_2') \geqslant \Omega(\alpha^{O(1)})\delta$.

Thus the process terminates in $O(\alpha_0^{-15})$ steps, and $\delta = \mu(E_1 \boxtimes E_2) \geqslant \exp(-\alpha_0^{-O(1)})$ throughout the process. Throughout, we have required that $\log(1/\gamma)^{-c} < \delta^{O(1)}$, which holds for $\gamma = \exp(-\exp(\alpha_0^{-O(1)}))$. For $\alpha_0 = (\log\log\log\log n)^{-c}$ for sufficiently small c, the dimension during the iterative process is always at least

$$n^{1000^{-O(\gamma^{-4}\delta^{-1}\alpha_0^{-O(1)})}} \geqslant \exp((\log n)^{0.999}).$$

This completes the proof. $\qquad\square$

Acknowledgments

References

[AS74]   M. Ajtai and E. Szemerédi. Sets of lattice points that form no squares. *Studia Sci. Math. Hungar.*, 9:9–11, 1974.

[Aus11]   Tim Austin. Deducing the density Hales–Jewett theorem from an infinitary removal lemma. *Journal of Theoretical Probability*, 24:615–633, 2011.

[Beh46]   F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U.S.A.*, 32:331–332, 1946.

[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. On approximability of satisfiable k-csps: VII. 2024.

[BKM22]   Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: 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. Effective bounds for restricted 3-arithmetic progressions in $\mathbb{F}_p^n$. *Electron. Colloquium Comput. Complex.*, TR23-116, 2023. To appear in Discrete Analysis.

[BKM23b]   Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: 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.

[BKM23c]   Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: III. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 643–655. ACM, 2023.

[BKM24a]   Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: IV. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1423–1434. ACM, 2024.

[BKM24b]   Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: V. *CoRR*, abs/2408.15377, 2024.

[BS23]   Thomas F Bloom and Olof Sisask. An improvement to the Kelley-Meka bounds on three-term arithmetic progressions. *arXiv preprint arXiv:2309.02353*, 2023. Available at https://arxiv.org/pdf/2309.02353.

[DKT14]   Pandelis Dodos, Vassilis Kanellopoulos, and Konstantinos Tyros. A simple proof of the density Hales-Jewett theorem. *Int. Math. Res. Not. IMRN*, (12):3340–3352, 2014.

[FK78]   H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.*, 34:275–291, 1978.

[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.

[Gow98]   W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998.

[Gow01]   W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[Gre04]   Ben Green. Finite field models in additive combinatorics. *arXiv preprint math/0409420*, 2004. Available at https://arxiv.org/pdf/math/0409420.

[Gre05a] Ben Green. An argument of Shkredov in the finite field setting. *Preprint*, 2005. Available at https://people.maths.ox.ac.uk/greenbj/papers/corners.pdf.

[Gre05b] Ben Green. Roth's theorem in the primes. *Ann. of Math. (2)*, 161(3):1609–1636, 2005.

[GT08]   Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.

[GT10]   Benjamin Green and Terence Tao. Linear equations in primes. *Ann. of Math. (2)*, 171(3):1753–1850, 2010.

[GT12]   Ben Green and Terence Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math. (2)*, 175(2):465–540, 2012.

[GT17]   Ben Green and Terence Tao. New bounds for Szemerédi's theorem, III: a polylogarithmic bound for $r_4(N)$. *Mathematika*, 63(3):944–1040, 2017.

[GTZ11]   Ben Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers U^4-norm. *Glasg. Math. J.*, 53(1):1–50, 2011.

[GTZ12]   Ben Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math. (2)*, 176(2):1231–1372, 2012.

[HJ63]   A. W. Hales and R. I. Jewett. Regularity and positional games. *Trans. Amer. Math. Soc.*, 106:222–229, 1963.

[HLY21]   Rui Han, Michael T. Lacey, and Fan Yang. A polynomial Roth theorem for corners in finite fields. *Mathematika*, 67(4):885–896, 2021.

[KM23]   Zander Kelley and Raghu Meka. Strong bounds for 3-progressions. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science—FOCS 2023*, pages 933–973. IEEE Computer Soc., Los Alamitos, CA, [2023] ©2023.

[Len24]   James Leng. A quantitative bound for Szemerédi's theorem for a complexity one polynomial progression over $\mathbb{Z}/N\mathbb{Z}$. *Discrete Anal.*, pages Paper No. 3, 33, 2024.

[LSS23]   James Leng, Ashwin Sah, and Mehtaab Sawhney. Improved bounds for five-term arithmetic progressions. *arXiv preprint arXiv:2312.10776*, 2023. Available at https://arxiv.org/pdf/2312.10776.

[LSS24a] James Leng, Ashwin Sah, and Mehtaab Sawhney. Improved bounds for Szemerédi's theorem. *arXiv preprint arXiv:2402.17995*, 2024. Available at https://arxiv.org/pdf/2402.17995.

[LSS24b] James Leng, Ashwin Sah, and Mehtaab Sawhney. Quasipolynomial bounds on the inverse theorem for the Gowers $U^{s+1}[N]$-norm. *arXiv preprint arXiv:2402.17994*, 2024. Available at https://arxiv.org/pdf/2402.17994v3.

[Pel18]   Sarah Peluse. Three-term polynomial progressions in subsets of finite fields. *Israel J. Math.*, 228(1):379–405, 2018.

[Pel19]   Sarah Peluse. On the polynomial Szemerédi theorem in finite fields. *Duke Math. J.*, 168(5):749–774, 2019.

[Pel20]   Sarah Peluse. Bounds for sets with no polynomial progressions. *Forum Math. Pi*, 8:e16, 55, 2020.

[Pel24]   Sarah Peluse. Subsets of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ without L-shaped configurations. *Compos. Math.*, 160(1):176–236, 2024.

[Pol09]   DHJ Polymath. Density Hales–Jewett and Moser numbers in low dimensions. *Unpublished, http://michaelnielsen.org/polymath1*, 2009.

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

[PP22]   Sarah Peluse and Sean Prendiville. A polylogarithmic bound in the nonlinear Roth theorem. *Int. Math. Res. Not. IMRN*, (8):5658–5684, 2022.

[Rot53]   K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.

[Shk05]   I. D. Shkredov. On a problem of Gowers. *Izvestiya: Mathematics*, 71(1):46–48, 2005.

[Shk06]   I. D. Shkredov. On a generalization of Szemerédi's theorem. *Proc. London Math. Soc. (3)*, 93(3):723–760, 2006.

[Sze75]   E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. In *Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974), Vol. 2*, pages 503–505. Canad. Math. Congr., Montreal, QC, 1975.

[VdW27]   Bartel Leendert Van der Waerden. Beweis einer baudetschen vermutung. *Nieuw Arch. Wiskunde*, 15:212–216, 1927.