

POPULAR DIFFERENCES FOR MATRIX PATTERNS

AARON BERGER, ASHWIN SAH, MEHTAAB SAWHNEY, AND JONATHAN TIDOR

ABSTRACT. The following combinatorial conjecture arises naturally from recent ergodic-theoretic work of Ackelsberg, Bergelson, and Best. Let M_1, M_2 be $k \times k$ integer matrices, G be a finite abelian group of order N , and $A \subseteq G^k$ with $|A| \geq \alpha N^k$. If $M_1, M_2, M_1 - M_2$, and $M_1 + M_2$ are automorphisms of G^k , is it true that there exists a popular difference $d \in G^k \setminus \{0\}$ such that

$$\#\{x \in G^k : x, x + M_1d, x + M_2d, x + (M_1 + M_2)d \in A\} \geq (\alpha^4 - o(1))N^k.$$

We show that this conjecture is false in general, but holds for $G = \mathbb{F}_p^n$ with p an odd prime given the additional spectral condition that no pair of eigenvalues of $M_1M_2^{-1}$ (over $\overline{\mathbb{F}_p}$) are negatives of each other. In particular, the “rotated squares” pattern does not satisfy this eigenvalue condition, and we give a construction of a set of positive density in $(\mathbb{F}_5^n)^2$ for which that pattern has no nonzero popular difference. This is in surprising contrast to three-point patterns, which we handle over all compact abelian groups and which do not require an additional spectral condition.

1. INTRODUCTION

1.1. Popular patterns and past results. Using an argument of Varnavides [21], it is well-known that Roth’s theorem [17] on three-term arithmetic progressions can be strengthened to guarantee at least $c_\alpha N^2$ arithmetic progressions in a set $A \subseteq [N]$ of size αN . The constant c_α is known not to be polynomial in α ; in particular, modifying a well-known construction of Behrend [2] allows one to construct sets with $\alpha^{c \log(1/\alpha)} N^2$ three-term arithmetic progressions. However, Green [9], showed that one has a “popular” common difference $d \neq 0$, i.e., a value $d \in [N]$ such that

$$\#\{a : a, a + d, a + 2d \in A\} \geq (\alpha^3 - o(1))N.$$

That is, the set behaves like a random set along certain structured differences, if not all of them. Green’s proof involves an arithmetic regularity lemma, which is essentially equivalent to arithmetic regularity for the Gowers U^2 -norm.

One can ask if this phenomenon holds for longer arithmetic progressions. The analogous result for four-term arithmetic progressions with $\alpha^4 - o(1)$ on the right-hand side was proved by [13] relying on a remarkable “positivity” identity [13] (see [10] for a version over \mathbb{F}_p^n for $p \geq 5$) in combination with the U^3 -arithmetic regularity results of Green and Tao [13]. However, surprisingly, an $\alpha^k - o(1)$ (or any polynomial) bound does not hold for k -term arithmetic progressions for $k \geq 5$ due to a construction of Ruzsa [3, Appendix]. These results were motivated by corresponding ergodic results of Bergelson, Host, and Kra [3], although the theorems do not directly transfer when studying popular differences (as opposed to Furstenberg’s correspondence theorem for Szemerédi’s theorem).

One may ask about popularity of more general patterns, for example $\{0, 1, 2, 4\}$. (We use a set to refer to the pattern consisting of homothetic copies of that set; in this case, the pattern is $(a, a + d, a + 2d, a + 4d)$.) The proof of Green and Tao [13] for four-term arithmetic progressions (and [10] over finite fields) immediately extends to patterns of the form $\{0, k_1, k_2, k_1 + k_2\}$ for $k_1 k_2 (k_1 + k_2) \neq 0$. Work of the second and third authors and Zhao [18] shows that two-point, three-point, and these specific “parallelogram” four-point patterns are the only popular patterns over \mathbb{Z} .

Berger, Sah, Sawhney, and Tidor were supported by NSF Graduate Research Fellowship Program DGE-1745302.

Popularity of higher-dimensional patterns such as corners, $\{(0, 0), (1, 0), (0, 1)\}$, was first studied by Mandache [15] in the combinatorial setting (see [6, 7] for related work in the ergodic theory setting), who showed over \mathbb{F}_p^n that they are not α^3 -popular but do satisfy a weakened bound with α^4 instead. Fox, the second and third authors, Stoner, and Zhao [8] showed that the tight bound is of the form $\alpha^4\tau(\alpha)$, where τ grows as $\alpha \rightarrow 0$, but is of the form $\alpha^{o(1)}$. Finally, the first author [4] showed the same behavior over \mathbb{Z}^2 . The second and third authors and Zhao [18] studied higher-dimensional patterns which are homothetic copies of a set and provide a nearly comprehensive classification.

1.2. Our contributions. The standard toolset of arithmetic regularity in higher-order Fourier analysis, which can prove popular difference results for three and four-point single-dimensional patterns, necessarily breaks to some extent when handling higher-dimensional corners (as pointed out in [18]), and has not yet been successfully applied to four-point patterns such as squares for which the question remains open. However, it was noted by Prendiville [16] that classic single-dimensional techniques extend if one considers *full-rank* matrix patterns (a collection which excludes corners and squares but includes a wide class of multidimensional configurations such as “rotated corners” – also known as “right isosceles triangles” – and “rotated squares”), and he achieves versions of Szemerédi’s theorem (for $k \leq 4$ points) with good quantitative bounds in this setting. We continue in this line of work, achieving popular difference results of strength equal to the single-dimensional case, illustrating by comparison the suitability of these methods to full-rank patterns.

The main novelty of this paper lies in the popular difference results for four-point patterns, where we exhibit further behavior that does not appear even in Prendiville’s work. In order to properly handle popularity of four-point patterns, we show that one must apply the method of arithmetic regularity in a manner that sees the spectral properties of the matrices defining the pattern. In particular, the counting lemma (which for four-point patterns over \mathbb{F}_p^n relies heavily on equidistribution over parts in quadratic factors) becomes qualitatively distinct depending on the spectral structure of the matrices in the pattern (see [Theorem 4.4](#)). This subtlety is not present in earlier counting lemmas for scalar-valued patterns. This also translates concretely to an additional restriction that no pair of eigenvalues of an associated matrix can be negatives of each other for our method to produce a popular difference result (see [Theorem 1.2](#)). To confirm that this behavior is genuine and not an artifact of the proof, in [Theorem 1.3](#) we exhibit a full-rank matrix pattern which does not satisfy the additional spectral condition imposed by [Theorem 1.2](#) and for which the conclusion of the theorem is false. In particular, we show that rotated squares in \mathbb{F}_5^n do not satisfy a popular difference result, at least with popularity α^4 .

1.3. Summary of results. We first prove a popular differences result for all full-rank three-point patterns. A three-point pattern is full rank if it can be expressed in the form $\vec{x}, \vec{x} + M_1\vec{d}, \vec{x} + M_2\vec{d}$ where $M_1, M_2, M_1 - M_2$ are invertible. One such example is “rotated corners,” which are of the form $(x, y), (x+a, y+b), (x+b, y-a)$. (By contrast, standard corners $(x, y), (x+a, y), (x, y+a)$ are not full rank.) As a special case, this resolves a conjecture of Ackelsberg, Bergelson, and Best [1, Question 1.21], which concerns the case of rotated corners specifically. Kovač [14] has independently proved this rotated corners conjecture with similar methods.

Theorem 1.1. *Let M_1, M_2 be $k \times k$ invertible integer matrices so that $M_1 - M_2$ is invertible. For any $\alpha, \epsilon > 0$ there exists $N_0(\alpha, \epsilon, M_1, M_2)$ so that the following holds. If $N \geq N_0$, then for any $A \subseteq [N]^k$, $|A| \geq \alpha N^k$, there is a popular difference $\vec{d} \neq 0$ so that*

$$\#\{\vec{x} \in [N]^k : \vec{x}, \vec{x} + M_1\vec{d}, \vec{x} + M_2\vec{d} \in A\} \geq (\alpha^3 - \epsilon)N^k.$$

We additionally prove an analogous version of the result where the interval $[N]$ is replaced by an arbitrary compact abelian group G . See [Section 7](#) for the precise statement and proof of this result.

We turn next to four-point patterns of matrices. There are a few natural restrictions on generic patterns $\vec{x}, \vec{x} + M_1\vec{d}, \vec{x} + M_2\vec{d}, \vec{x} + M_3\vec{d}$ that arise when trying to prove a popular differences result. First, we impose $M_3 = M_1 + M_2$, which is a generalization of the “parallelogram” condition in the popular differences result of Green and Tao [13]. Second, we require that $M_1, M_2, M_1 - M_2, M_1 + M_2$ are all invertible; in this case we call the pattern full rank.¹ The combination of these two conditions is analogous to the “admissibility” condition of [1], and essentially appears in [16]. One might guess that they are sufficient to guarantee popular differences. In this paper we show that this guess is incorrect by demonstrating the necessity of an additional spectral condition on the pattern. In the spirit of the finite field philosophy advocated by Green [11], we restrict attention to the finite field model $G = \mathbb{F}_p^n$ with p an odd prime. We suspect our methods can be extended to handle more general abelian groups, but choose to avoid the complexity of the inverse theorems for the U^3 -norm over general abelian groups.

Theorem 1.2. *Fix $k \geq 1$ and p an odd prime. Let M_1, M_2 be $k \times k$ matrices with coefficients in \mathbb{F}_p such that $M_1, M_2, M_1 - M_2$, and $M_1 + M_2$ are invertible and no pair of eigenvalues of $M_1M_2^{-1}$ (viewed over $\overline{\mathbb{F}}_p$) are negatives of each other. For $\alpha, \epsilon > 0$, there exists $n_0(\alpha, \epsilon, p)$ such that the following holds. If $n \geq n_0$, then for any $A \subseteq (\mathbb{F}_p^n)^k$, $|A| \geq \alpha p^{nk}$, there is a popular difference $\vec{d} \neq 0$ so that*

$$\#\{\vec{x} \in (\mathbb{F}_p^n)^k : \vec{x}, \vec{x} + M_1\vec{d}, \vec{x} + M_2\vec{d}, \vec{x} + (M_1 + M_2)\vec{d} \in A\} \geq (\alpha^4 - \epsilon)p^{nk}.$$

In fact, there are $\Omega_{\alpha, \epsilon, p}(p^{nk})$ values of \vec{d} that work.

Furthermore, we show that one cannot completely remove the spectral condition.

Theorem 1.3. *There is an absolute constant $c > 0$ such that the following holds. If $\alpha \in (0, c)$, then for all sufficiently large n (depending on α) there is a set $A \subseteq (\mathbb{F}_5^n)^2$ satisfying $|A| \geq \alpha 5^{2n}$ and*

$$\max_{(a,b) \neq 0} \#\{(x, y) : (x, y), (x + a, y + b), (x + b, y - a), (x + a + b, y + b - a) \in A\} \leq (1 - c)\alpha^4 5^{2n}.$$

Here the associated matrices are

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Note that the eigenvalues of $M_1M_2^{-1}$ indeed are negatives of each other. We believe it is likely that one can construct a counterexample for all $k \times k$ matrices with some pair of negated eigenvalues by lifting the ideas involved in this construction.

Although there is no direct implication, this can be seen as a combinatorial finite field analogue of [1, Question 1.11] and we expect our counterexample can be extended to the ergodic setting. In particular, we answer the combinatorial analogue in the negative but point to a potential new condition under which their question might be resolved positively.

1.4. Notation and outline. We use O, o, Ω as standard asymptotic notation. Subscripts in said notation denote dependence of the implicit constants on those subscripts.

The majority of this paper, Sections 2 to 5, is devoted to the proof of Theorem 1.2, the popular difference result for four-point patterns. See Section 2 for an outline of that argument. In Section 6, we construct the counterexample that proves Theorem 1.3. Finally, in Section 7, we show Theorem 1.1, the three-point pattern result.

Acknowledgements. We thank our advisor Yufei Zhao for introducing us to the study of popular differences in additive combinatorics.

¹Axis-aligned squares are an example of a four-point pattern that is not full rank, and for which the version of popular differences we would like to prove is known to be false; see [18, Theorem 3.1].

2. GOWERS NORMS AND ARITHMETIC REGULARITY

The proof of [Theorem 1.2](#) proceeds in three steps, following the now-standard framework of the arithmetic regularity method.

First, we show that the matrix patterns we are interested in are controlled by an appropriate Gowers U^s -norm. Results of this nature are sometimes referred to as “generalized von Neumann theorems”. The definition of the Gowers norms and the proof of this result are given in this section.

Second, we prove an arithmetic regularity lemma, which gives a decomposition of an arbitrary function $f: G \rightarrow \mathbb{C}$ as $f = f_{\text{str}} + f_{\text{sml}} + f_{\text{psr}}$ into a “structured”, “small”, and “pseudorandom” piece. For our application in groups G^k , it will be necessary to carefully define “structured” in a way that is adapted to the product structure of G^k . This definition and the proof of this result is given in [Section 3](#).

Third, we prove novel equidistribution results in order to understand the counts of matrix patterns inside the structured piece, f_{str} . These results occur in [Section 4](#). Combining these three steps, we prove [Theorem 1.2](#) in [Section 5](#).

Definition 2.1. Fix an integer $s \geq 1$ and a finite abelian group G . For a function $f: G \rightarrow \mathbb{C}$, the Gowers U^s -norm is defined by

$$\|f\|_{U^s(G)} = \left(\mathbb{E}_{x, h_1, \dots, h_s \in G} \prod_{\omega \in \{0,1\}^s} \mathcal{C}^{|\omega|} f(x + \omega_1 h_1 + \dots + \omega_s h_s) \right)^{1/2^s},$$

where \mathcal{C} denotes the complex conjugation operator and $|\omega| = \omega_1 + \dots + \omega_s$.

It is well-known that the above is indeed a norm when $s \geq 2$. (For $s = 1$ it is the seminorm $f \mapsto |\mathbb{E}_{x \in G} f(x)|$, so the term “Gowers norm” is a slight misnomer.) A useful equivalent definition is that

$$\|f\|_{U^s(G)}^{2^s} = \mathbb{E}_{h \in G} \|\partial_h f\|_{U^{s-1}(G)}^{2^{s-1}}$$

where the *multiplicative derivative* $\partial_h f$ is defined by $(\partial_h f)(x) = f(x) \overline{f(x+h)}$.

We now prove that full-rank matrix patterns are controlled by an appropriate U^s -norm. The typical setup in this paper is to consider a pattern of the form $\vec{x} + M_1 \vec{d}, \vec{x} + M_2 \vec{d}, \dots, \vec{x} + M_s \vec{d}$ in G^k where M_1, \dots, M_s are $k \times k$ matrices with certain non-degeneracy conditions. In particular we assume that M_i and $M_i - M_j$ are invertible for each $i \neq j$.

This lemma is true even in the general setting where we replace the matrix M_i acting on G^k by an arbitrary automorphism A_i acting on G^k . In this general setting, the product structure on G^k is no longer important. As the proof of the more general version is no more difficult than the original result, we include it here. The proof follows by an application of the Cauchy–Schwarz inequality; similar results for specific patterns are implicit in the literature (e.g., [\[16\]](#)).

Lemma 2.2. *Let $s \geq 2$, and G be a finite abelian group. Let A_1, \dots, A_s be automorphisms of G such that $A_i - A_j$ is an automorphism for each $i \neq j$. Then for functions $f_i: G \rightarrow \mathbb{C}$ satisfying $\|f_i\|_\infty \leq 1$ we have*

$$|\mathbb{E}_{x, d \in G} f_1(x + A_1 d) \cdots f_s(x + A_s d)| \leq \min_{i \in [s]} \|f_i\|_{U^{s-1}(G)}.$$

Proof. We induct on s . For $s = 2$, note that

$$\begin{aligned} |\mathbb{E}_{x, d \in G} f_1(x + A_1 d) f_2(x + A_2 d)| &= |\mathbb{E}_{x, d \in G} f_1(x) f_2(x + (A_2 - A_1) d)| \\ &= |\mathbb{E}_{x, y \in G} f_1(x) f_2(y)| \\ &= \|f_1\|_{U^1(G)} \|f_2\|_{U^1(G)}. \end{aligned}$$

Since $\|f_i\|_\infty \leq 1$, the result follows in this case. Now suppose $s \geq 3$. We have

$$\begin{aligned}
& |\mathbb{E}_{x,d \in G} f_1(x + A_1 d) \cdots f_s(x + A_s d)| \\
&= |\mathbb{E}_{x,d \in G} f_1(x + (A_1 - A_s)d) \cdots f_{s-1}(x + (A_{s-1} - A_s)d) f_s(x)| \\
&\leq \mathbb{E}_x |\mathbb{E}_d f_1(x + (A_1 - A_s)d) \cdots f_{s-1}(x + (A_{s-1} - A_s)d)| \\
&\leq (\mathbb{E}_x |\mathbb{E}_d f_1(x + (A_1 - A_s)d) \cdots f_{s-1}(x + (A_{s-1} - A_s)d)|^2)^{1/2} \\
&= (\mathbb{E}_x \mathbb{E}_{d,h} f_1(x + (A_1 - A_s)d) \cdots f_{s-1}(x + (A_{s-1} - A_s)d) \\
&\quad \cdot \overline{f_1}(x + (A_1 - A_s)d + (A_1 - A_s)h) \cdots \overline{f_{s-1}}(x + (A_{s-1} - A_s)d + (A_{s-1} - A_s)h))^{1/2}.
\end{aligned}$$

To bound the last expression, we apply the induction hypothesis with the maps $A_1 - A_s, \dots, A_{s-1} - A_s$ and the functions $\partial_{(A_i - A_s)h} f_i$. Note that by hypothesis, the maps $A_i - A_s$ are automorphisms as are $(A_i - A_s) - (A_j - A_s)$ for $i \neq j$. Therefore we obtain

$$\begin{aligned}
|\mathbb{E}_{x,d \in G} f_1(x + A_1 d) \cdots f_s(x + A_s d)| &\leq (\mathbb{E}_h \|\partial_{(A_1 - A_s)h} f_1\|_{U^{s-2}(G)})^{1/2} \\
&\leq \left(\mathbb{E}_h \|\partial_{(A_1 - A_s)h} f_1\|_{U^{s-2}(G)}^{2^{s-2}} \right)^{1/2^{s-1}} \\
&= \|f_1\|_{U^{s-1}(G)}.
\end{aligned}$$

The last equality comes from the recursive definition of the Gowers norms as well as the fact that $A_1 - A_s$ is an automorphism on G . By symmetry, the same holds for f_2, \dots, f_s , completing the proof. \square

3. THE U^3 -ARITHMETIC REGULARITY LEMMA

From now on until [Section 7](#), we restrict our attention to the case where $G = \mathbb{F}_p^n$ and p is an odd prime. The goal of this section is to prove a U^3 -arithmetic regularity lemma for functions $f: G^k \rightarrow \mathbb{C}$. Since $G^k \cong \mathbb{F}_p^{nk}$, we could apply a standard result (say [\[10, Proposition 3.12\]](#)) to deduce some U^3 -regularity statement. However such regularity statement would ignore the product structure on G^k which will become very important in our application.

The main novelty of this section is our definition of a k -symmetrized quadratic factor which gives an appropriate notion of structured function adapted to the product structure of G^k . We then prove [Theorem 3.2](#), our k -symmetrized U^3 -arithmetic regularity lemma. The structure of the proof closely follows [\[10\]](#).

An element $\vec{x} \in G^k$ is a tuple $\vec{x} = (x_1, \dots, x_k)$ with $x_1, \dots, x_k \in \mathbb{F}_p^n$. It will simplify the following arguments to introduce the following slightly awkward notation: we view the elements of G^k as $k \times n$ matrices X where the rows of X correspond to the elements of the k -tuple. In particular, the element $\vec{x} \in G^k$ is alternatively represented as

$$X = \begin{pmatrix} \text{-----} x_1^\top \text{-----} \\ \vdots \\ \text{-----} x_k^\top \text{-----} \end{pmatrix}.$$

Finally, define \mathcal{S}_k (respectively, \mathcal{S}'_k) to be the set of symmetric (respectively, skew-symmetric) matrices in $\mathbb{F}_p^{k \times k}$.

Definition 3.1. A (k) -symmetrized quadratic factor $\mathfrak{B} = (\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ is given by a list $\mathfrak{B}_1 = (r_1, \dots, r_{d_1})$ of column vectors in \mathbb{F}_p^n , a list $\mathfrak{B}_2 = (M_1, \dots, M_{d_2})$ of symmetric matrices in $\mathbb{F}_p^{n \times n}$, and a list $\mathfrak{B}_3 = (N_1, \dots, N_{d_3})$ of skew-symmetric matrices in $\mathbb{F}_p^{n \times n}$. The *complexity* of \mathfrak{B} is (d_1, d_2, d_3) .

We say that \mathfrak{B} has *rank* at least r if r_1, \dots, r_{d_1} are linearly independent and all nontrivial linear combinations

$$\sum_{i=1}^{d_2} a_i M_i + \sum_{j=1}^{d_3} b_j N_j$$

have \mathbb{F}_p -rank at least r . (This is equivalent to the same condition on M_1, \dots, M_{d_2} and N_1, \dots, N_{d_3} separately up to an absolute multiplicative constant in the rank.)

A k -symmetrized quadratic factor \mathfrak{B} defines maps $\mathbf{B}_{1,i}: G^k \rightarrow \mathbb{F}_p^k$, $\mathbf{B}_{2,i}: G^k \rightarrow \mathcal{S}_k$, and $\mathbf{B}_{3,i}: G^k \rightarrow \mathcal{S}'_k$ given by

$$\mathbf{B}_{1,i}(X) = X r_i, \quad \mathbf{B}_{2,i}(X) = X M_i X^\top, \quad \mathbf{B}_{3,i}(X) = X N_i X^\top.$$

We additionally define

$$\begin{aligned} \mathbf{B}_1(X) &= (\mathbf{B}_{1,i}(X))_{i \in [d_1]}, & \mathbf{B}_2(X) &= (\mathbf{B}_{2,i}(X))_{i \in [d_2]}, & \mathbf{B}_3(X) &= (\mathbf{B}_{3,i}(X))_{i \in [d_3]}, \\ \mathbf{B}(X) &= (\mathbf{B}_1(X), \mathbf{B}_2(X), \mathbf{B}_3(X)). \end{aligned}$$

For a function $f: G^k \rightarrow \mathbb{C}$, we use the notation $\mathbb{E}[f|\mathfrak{B}]$ to represent the condition expectation of f with respect to \mathfrak{B} , or equivalently the projection of f onto \mathfrak{B} . Here we abuse notation and use \mathfrak{B} to denote the σ -algebra generated by the fibers of \mathbf{B} in G^k . Explicitly, $\mathbb{E}[f|\mathfrak{B}]: G^k \rightarrow \mathbb{C}$ is defined by $\mathbb{E}[f|\mathfrak{B}](X) = \mathbb{E}_{Y \in \mathbf{B}^{-1}(\mathbf{B}(X))}[f(Y)]$.

Finally, we say that a factor \mathfrak{B}' *refines* a factor \mathfrak{B} if the σ -algebra corresponding to \mathfrak{B}' refines the σ -algebra corresponding to \mathfrak{B} .

The main result of this section is the following arithmetic regularity statement which guarantees that the desired factor is k -symmetrized.

Theorem 3.2 (Arithmetic regularity lemma). *Fix $k \geq 1$. Let $\delta > 0$ and let $\omega_1, \omega_2: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be arbitrary growth functions (which may depend on δ). Let $G = \mathbb{F}_p^n$, let $f: G^k \rightarrow [0, 1]$ be a function, and let $(\mathfrak{B}_1^{(0)}, \mathfrak{B}_2^{(0)}, \mathfrak{B}_3^{(0)})$ be a k -symmetrized quadratic factor of complexity $(d_1^{(0)}, d_2^{(0)}, d_3^{(0)})$. Then there is a refinement $(\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ of complexity (d_1, d_2, d_3) and a decomposition $f = f_{\text{str}} + f_{\text{sml}} + f_{\text{psr}}$ such that:*

1. $f_{\text{str}} = \mathbb{E}[f|\mathfrak{B}]$;
2. $\|f_{\text{sml}}\|_2 \leq \delta$;
3. $\|f_{\text{psr}}\|_{U^3(G^k)} \leq 1/\omega_2(d_1 + d_2 + d_3)$;
4. f_{str} and $f_{\text{str}} + f_{\text{sml}}$ take values in $[0, 1]$ and $f_{\text{psr}}, f_{\text{sml}}$ take values in $[-1, 1]$;
5. the complexity of \mathfrak{B} is (d_1, d_2, d_3) where

$$d_1, d_2, d_3 \leq C(k, \delta, \omega_1, \omega_2, d_1^{(0)}, d_2^{(0)}, d_3^{(0)})$$

for a fixed function C ;

6. the rank of \mathfrak{B} is at least $\omega_1(d_1 + d_2 + d_3)$.

The proof closely follows the proof of arithmetic regularity given in [9]; the only additional ingredient is guaranteeing at each stage that the factor introduced is k -symmetrized.

Lemma 3.3. *Fix $k \geq 1$. Let $\delta > 0$. There exists $\epsilon > 0$ such that the following holds. Let $\mathfrak{B}^{(0)} = (\mathfrak{B}_1^{(0)}, \mathfrak{B}_2^{(0)}, \mathfrak{B}_3^{(0)})$ be a k -symmetrized quadratic factor with complexity (d_1, d_2, d_3) and let $f: G^k \rightarrow [-1, 1]$ be a function such that*

$$\|f - \mathbb{E}[f|\mathfrak{B}^{(0)}]\|_{U^3(G^k)} \geq \delta.$$

Then there exists a refinement $\mathfrak{B} = (\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ with complexity at most $(d_1 + k, d_2 + \binom{k+1}{2}, d_3 + \binom{k}{2})$ such that

$$\|\mathbb{E}[f|\mathfrak{B}]\|_2^2 \geq \|\mathbb{E}[f|\mathfrak{B}^{(0)}]\|_2^2 + \epsilon^2.$$

Proof. By the inverse theorem for the Gowers U^3 -norm applied to $G^k \cong \mathbb{F}_p^{nk}$ (see [10, 12]), there exist $\epsilon > 0$ (only depending on δ), a vector $r \in \mathbb{F}_p^{nk}$, and a symmetric matrix $M \in \mathbb{F}_p^{nk \times nk}$ such that

$$\left| \mathbb{E}_{x \in \mathbb{F}_p^{nk}} \left(f(x) - \mathbb{E}[f|\mathfrak{B}^{(0)}](x) \right) e_p(r^\top x + x^\top Mx) \right| \geq \epsilon.$$

Say $r = (r_1, \dots, r_k) \in \mathbb{F}_p^{nk}$ where $r_1, \dots, r_k \in \mathbb{F}_p^n$ and $M = (M_{ij})_{i,j \in [k]} \in \mathbb{F}_p^{nk \times nk}$ where $M_{ij} \in \mathbb{F}_p^{n \times n}$. Note that the matrices M_{ii} are symmetric, while $M_{ij} = M_{ji}^\top$. For $i < j$, write $M_{ij} = M'_{ij} + M''_{ij}$ where M'_{ij} is symmetric and M''_{ij} is skew-symmetric. (Here we use that $p > 2$.)

We define the factor \mathfrak{B} by appending the vectors r_1, \dots, r_k to the list $\mathfrak{B}_1^{(0)}$, appending the symmetric matrices $(M_{ii})_{i \in [k]}$ and $(M'_{ij})_{i < j}$ to the list $\mathfrak{B}_2^{(0)}$, and appending the skew-symmetric matrices $M_{ij})_{i < j}$ to the list $\mathfrak{B}_3^{(0)}$. To conclude, all that remains to show is that

$$\|\mathbb{E}[f|\mathfrak{B}]\|_2^2 \geq \|\mathbb{E}[f|\mathfrak{B}^{(0)}]\|_2^2 + \epsilon^2.$$

Define $g: (\mathbb{F}_p^n)^k \rightarrow \mathbb{C}$ by $g(x) = e_p(r^\top x + x^\top Mx)$. Note that $\|g\|_2 = 1$ and crucially that g is \mathfrak{B} -measurable by the simple equality

$$g(x_1, \dots, x_k) = e_p \left(\sum_{i=1}^k r_i^\top x_i + \sum_{i=1}^k x_i^\top M_{ii} x_i + 2 \sum_{i < j} x_i^\top M'_{ij} x_j + 2 \sum_{i < j} x_i^\top M''_{ij} x_j \right).$$

Now the desired inequality follows from the Pythagorean theorem and Cauchy–Schwarz inequality since

$$\begin{aligned} \|\mathbb{E}[f|\mathfrak{B}]\|_2^2 - \|\mathbb{E}[f|\mathfrak{B}^{(0)}]\|_2^2 &= \|\mathbb{E}[f|\mathfrak{B}] - \mathbb{E}[f|\mathfrak{B}^{(0)}]\|_2^2 \\ &\geq \left| \left\langle \mathbb{E}[f|\mathfrak{B}] - \mathbb{E}[f|\mathfrak{B}^{(0)}], g \right\rangle \right|^2 \\ &= \left| \left\langle f - \mathbb{E}[f|\mathfrak{B}^{(0)}], \mathbb{E}[g|\mathfrak{B}] \right\rangle \right|^2 \\ &\geq \epsilon^2. \end{aligned} \quad \square$$

Lemma 3.4. *Fix $k \geq 1$. Let $\omega: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be an arbitrary growth function. There exists a growth function $\tau: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that the following holds. Let $\mathfrak{B} = (\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ be a k -symmetrized quadratic factor with complexity (d_1, d_2, d_3) . There exists a refinement $\mathfrak{B}' = (\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ with complexity (d'_1, d'_2, d'_3) that satisfies the following:*

1. the rank of \mathfrak{B}' is at least $\omega(d'_1 + d'_2 + d'_3)$;
2. $d'_2 \leq d_2$ and $d'_3 \leq d_3$ and $d'_1 \leq \tau(d_1 + d_2 + d_3)$.

Proof. Consider a k -symmetrized quadratic factor $\mathfrak{B} = (\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ defined by $\mathfrak{B}_1 = (r_1, \dots, r_{d_1})$ and $\mathfrak{B}_2 = (M_1, \dots, M_{d_2})$ and $\mathfrak{B}_3 = (N_1, \dots, N_{d_3})$. (Recall that the r_i are vectors of length n while the M_i are symmetric $n \times n$ matrices and the N_i are skew-symmetric $n \times n$ matrices.)

If the rank of \mathfrak{B} is less than r , then either the r_1, \dots, r_{d_1} are linearly dependent or there exists a non-trivial linear combination

$$\sum_{i=1}^{d_2} a_i M_i + \sum_{j=1}^{d_3} b_j N_j$$

that has \mathbb{F}_p -rank less than r .

We do the following. First if there is some linear combination with rank less than r , then choose vectors $s_1, \dots, s_{r-1}, t_1, \dots, t_{r-1}$ such that

$$\sum_{i=1}^{r-1} s_i t_i^\top = \sum_{i=1}^{d_2} a_i M_i + \sum_{j=1}^{d_3} b_j N_j.$$

Add $s_1, \dots, s_{r-1}, t_1, \dots, t_{r-1}$ to \mathfrak{B}_1 and remove the first M_i or N_j with nonzero coefficient (i.e., if $a_1 = \dots = a_{i-1} = 0$ while $a_i \neq 0$, then remove M_i from \mathfrak{B}_2 ; if $a_1 = \dots = a_{d_2} = b_1 = \dots = b_{j-1} = 0$ and $b_j \neq 0$, then remove N_j from \mathfrak{B}_3). Then remove any element of the modified \mathfrak{B}_1 that is linearly dependent on the previous vectors in \mathfrak{B}_1 . Note that the factor produced refines the original factor.

We iterate the above process, producing a sequence of k -symmetrized quadratic factors $\mathfrak{B} = \mathfrak{B}^{(0)}, \mathfrak{B}^{(1)}, \dots, \mathfrak{B}^{(M)}$ as follows. Suppose that $\mathfrak{B}^{(m)}$ has complexity $(d_1^{(m)}, d_2^{(m)}, d_3^{(m)})$. If $\mathfrak{B}^{(m)}$ has rank at least $\omega(d_1^{(m)} + d_2^{(m)} + d_3^{(m)})$, then halt and set $m = M$. Otherwise refine $\mathfrak{B}^{(m)}$ to $\mathfrak{B}^{(m+1)}$ as described above. Note that $M \leq d_2 + d_3 + 1$ since every step (except possibly the first) reduces $d_2^{(m)} + d_3^{(m)}$ by 1. Furthermore, one can easily see that $d_1^{(M)}$ is bounded by some function of d_1, d_2, d_3 and ω , as desired. \square

Lemma 3.5. *Fix $k \geq 1$. Let $\delta > 0$ and let $\omega: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be an arbitrary growth function. Let $\mathfrak{B}^{(0)} = (\mathfrak{B}_1^{(0)}, \mathfrak{B}_2^{(0)}, \mathfrak{B}_3^{(0)})$ be a k -symmetrized quadratic factor with complexity $(d_1^{(0)}, d_2^{(0)}, d_3^{(0)})$ and let $f: G^k \rightarrow [0, 1]$ be a function. Then there exists a refinement $\mathfrak{B} = (\mathfrak{B}_1, \mathfrak{B}_2, \mathfrak{B}_3)$ and a decomposition $f = f_{\text{str}} + f_{\text{psr}}$ such that:*

1. $f_{\text{str}} = \mathbb{E}[f|\mathfrak{B}]$;
2. $\|f_{\text{psr}}\|_{U^3(G^k)} \leq \delta$;
3. f_{str} takes values in $[0, 1]$ and f_{psr} takes values in $[-1, 1]$;
4. the complexity of \mathfrak{B} is (d_1, d_2, d_3) where

$$d_1, d_2, d_3 \leq C(k, \delta, d_1^{(0)}, d_2^{(0)}, d_3^{(0)})$$

for a fixed function C ;

5. the rank of \mathfrak{B} is at least $\omega(d_1 + d_2 + d_3)$.

Proof. This follows immediately by iterating Lemma 3.3 and Lemma 3.4 at most $\epsilon(\delta)^{-2}$ times.

In particular, we construct a sequence of k -symmetrized quadratic factors $\mathfrak{B}^{(0)}, \mathfrak{B}^{(1)}, \dots, \mathfrak{B}^{(M)}$ each refining the last as follows. If $\|f - \mathbb{E}[f|\mathfrak{B}^{(m)}]\|_{U^3(G^k)} < \delta$, halt the process and set $M = m$. Otherwise, let $\tilde{\mathfrak{B}}^{(m+1)}$ be the factor produced by applying Lemma 3.3 to $\mathfrak{B}^{(m)}$ and f with parameter δ . Then let $\mathfrak{B}^{(m+1)}$ be the factor produced by applying Lemma 3.4 to $\tilde{\mathfrak{B}}^{(m+1)}$ with parameter ω . By definition, at every step of this process, the rank of $\mathfrak{B}^{(m)}$ is at least $\omega_1(d_1^{(m)} + d_2^{(m)} + d_3^{(m)})$ where $(d_1^{(m)}, d_2^{(m)}, d_3^{(m)})$ is the complexity of \mathfrak{B} .

Since

$$\|\mathbb{E}[f|\mathfrak{B}^{(m+1)}]\|_2^2 \geq \|\mathbb{E}[f|\tilde{\mathfrak{B}}^{(m+1)}]\|_2^2 \geq \|\mathbb{E}[f|\mathfrak{B}^{(m)}]\|_2^2 + \epsilon(\delta)^2$$

and this quantity is bounded between 0 and 1, we conclude that the process must stop after $M \leq \epsilon(\delta)^{-2}$ steps.

At the conclusion of this process, we have produced a k -symmetrized quadratic factor $\mathfrak{B}^{(M)}$ that refines $\mathfrak{B}^{(0)}$ such that $\|f - \mathbb{E}[f|\mathfrak{B}^{(M)}]\|_{U^3(G^k)} < \delta$. Defining $f_{\text{str}} = \mathbb{E}[f|\mathfrak{B}^{(M)}]$ and $f_{\text{psr}} = f - \mathbb{E}[f|\mathfrak{B}^{(M)}]$ gives the desired result. \square

Proof of Theorem 3.2. The desired result follows by iterating Lemma 3.5 at most δ^{-2} times.

In particular, we construct a sequence of k -symmetrized quadratic factors $\mathfrak{B}^{(0)}, \mathfrak{B}^{(1)}, \dots, \mathfrak{B}^{(M)}$ each refining the last as follows. If $\|\mathbb{E}[f|\mathfrak{B}^{(m)}] - \mathbb{E}[f|\mathfrak{B}^{(m-1)}]\|_2^2 < \delta^2$, halt the process and set $M = m$. Otherwise, let $\mathfrak{B}^{(m+1)}$ be the factor produced by applying Lemma 3.5 to $\mathfrak{B}^{(m)}$ and f with parameter $1/\omega_2(d_1^{(m)} + d_2^{(m)} + d_3^{(m)})$ and growth function ω_1 .

Note that by the Pythagorean theorem,

$$\|\mathbb{E}[f|\mathfrak{B}^{(m)}] - \mathbb{E}[f|\mathfrak{B}^{(m-1)}]\|_2^2 = \|\mathbb{E}[f|\mathfrak{B}^{(m)}]\|_2^2 - \|\mathbb{E}[f|\mathfrak{B}^{(m-1)}]\|_2^2.$$

Since these L^2 -norms are bounded between 0 and 1, we see that the process must stop after $M \leq \delta^{-2}$ steps.

At the conclusion of this process, we have produced a k -symmetrized quadratic factor $\mathfrak{B}^{(M-1)}$ that refines $\mathfrak{B}^{(0)}$ with complexity $(d_1^{(M-1)}, d_2^{(M-1)}, d_3^{(M-1)})$ and rank at least $\omega_1(d_1^{(M-1)} + d_2^{(M-1)} + d_3^{(M-1)})$. Defining $f_{\text{str}} = \mathbb{E}[f|\mathfrak{B}^{(M-1)}]$ and $f_{\text{sm1}} = \mathbb{E}[f|\mathfrak{B}^{(M)}] - \mathbb{E}[f|\mathfrak{B}^{(M-1)}]$ and $f_{\text{psr}} = f - \mathbb{E}[f|\mathfrak{B}^{(M)}]$ gives the desired result. \square

4. EQUIDISTRIBUTION AND COUNTING LEMMA

The goal of this section is to study the counts of matrix patterns of the form $\{0, M_1, M_2, M_1 + M_2\}$ in the “structured term” f_{str} . Recall that a k -symmetrized quadratic factor \mathfrak{B} defines a map $\mathbf{B}: G^k \rightarrow (\mathbb{F}_p^k)^{d_1} \times \mathcal{S}_k^{d_2} \times \mathcal{S}'_k^{d_3}$ where \mathcal{S}_k and \mathcal{S}'_k are the spaces of $k \times k$ symmetric (resp. skew-symmetric) matrices. We call the fibers of this map *atoms* of \mathfrak{B} .

Understanding the counts of patterns in f_{str} is equivalent to understanding how occurrences of these patterns are distributed among tuples of atoms. The first result of this section is simply that the atoms of \mathfrak{B} are approximately the same size; in other words, as $X \in G^k$ varies, $\mathbf{B}(X)$ is equidistributed in $(\mathbb{F}_p^k)^{d_1} \times \mathcal{S}_k^{d_2} \times \mathcal{S}'_k^{d_3}$.

The main result of this section describes how the 4-tuple $(\mathbf{B}(X), \mathbf{B}(X + M_1 D), \mathbf{B}(X + M_2 D), \mathbf{B}(X + (M_1 + M_2) D))$ is distributed as $X, D \in G^k$ vary. This 4-tuple is not equidistributed across all possible 4-tuples of atoms, instead it is equidistributed on a certain linear subspace. We need a somewhat unfortunate amount of notation in this section to describe this linear subspace.

Note that this is also the place where the “mysterious” spectral condition that $M_1 M_2^{-1}$ has no pair of eigenvalues that are negatives of each other appears. It turns out that the dimension of the space that the relevant 4-tuples are equidistributed over changes depending on whether or not this spectral condition is satisfied.

Finally in this section we restrict our attention to matrix patterns of the form $\{0, I, J, I + J\}$ where $I = I_{k \times k}$ is the identity, $J, I - J, I + J$ are invertible, and that J satisfies the spectral condition (that no pair of eigenvalues of J over $\overline{\mathbb{F}}_p$ are negatives of each other). By a change of variables, all cases can be reduced to this one.

4.1. Equidistribution results. We first quote the following result on the equidistribution in \mathbb{F}_p^n .

Proposition 4.1 ([10, Lemma 4.2]). *Define $\Gamma(x) = (r_1^\top x, \dots, r_{d_1}^\top x)$ and $\Phi(x) = (x^\top M_1 x, \dots, x^\top M_{d_2} x)$ where the M_i are symmetric. Furthermore suppose that $\{r_i\}_{i \in [d_1]}$ are linearly independent and for any nonzero vector $(\lambda_1, \dots, \lambda_{d_2})$ in $\mathbb{F}_p^{d_2}$ we have $\text{rank}(\sum_{i=1}^{d_2} \lambda_i M_i) \geq r$. Then for any $a \in \mathbb{F}_p^{d_1}$ and $b \in \mathbb{F}_p^{d_2}$ we have*

$$\mathbb{P}_{x \in \mathbb{F}_p^n}[\Gamma(x) = a, \Phi(x) = b] = p^{-d_1 - d_2} + O(p^{-r/2}).$$

Note [10] only states the above for \mathbb{F}_5^n but the proof in general is completely analogous. Given this we can immediately derive the necessary equidistribution result on factors for the specialized factors constructed in the previous section.

Proposition 4.2. *Let \mathfrak{B} be a k -symmetrized quadratic factor with rank at least r . Then*

$$\mathbb{P}_{X \in \mathbb{F}_p^{k \times n}}[\mathbf{B}(X) = ((v_i)_{i \in [d_1]}, (U_i)_{i \in [d_2]}, (V_i)_{i \in [d_3]})] = p^{-kd_1 - \binom{k+1}{2}d_2 - \binom{k}{2}d_3} + O(p^{-r/2})$$

for all $v_i \in \mathbb{F}_p^k$, $U_i \in \mathcal{S}_k$, and $V_i \in \mathcal{S}'_k$.

Proof. This is immediate if one treats $X \in \mathbb{F}_p^{k \times n}$ as a kn -dimensional vector. In particular for each U_i consider the family M_i of $\binom{k+1}{2}$ block matrices where all n by n blocks are zero except for either a diagonal block labeled U_i or a pair of block symmetric with respect to the diagonal such that blocks are labeled U_i . Similarly for each V_i consider the family N_i of $\binom{k}{2}$ block matrices where all n by n blocks are zero for a pair of block symmetric with respect to the diagonal such that the block above

the diagonal is labeled V_i and below the diagonal is labeled $-V_i$. Note that the resulting quadratic forms are easily seen to be high rank using that the U_i and V_i initially were high rank. Now the desired equidistribution statement is equivalent to equidistribution of $X^\top W X$ for all $W \in M_i, N_i$ as well as the linear forms specified by v_i . This now follows immediately from [Proposition 4.1](#). \square

Say that a random variable is ϵ -*equidistributed* if it takes each value in its range with equal probability within a *multiplicative* error of ϵ . A convenient property of this definition is that it is preserved under linear maps.

Lemma 4.3. *Suppose \mathbf{x} is a random variable taking values in \mathbb{F}_p^r satisfying*

$$\sup_{a \in \mathbb{F}_p^r} |p^r \mathbb{P}[\mathbf{x} = a] - 1| \leq \epsilon$$

and $L: \mathbb{F}_p^r \rightarrow \mathbb{F}_p^s$ is a linear map with image of dimension t . Then for any $a \in L\mathbb{F}_p^n$ we have $|p^t \mathbb{P}[L\mathbf{x} = a] - 1| \leq \epsilon$, whereas if $a \notin L\mathbb{F}_p^n$ then $\mathbb{P}[L\mathbf{x} = a] = 0$.

Proof. This follows immediately from the fact that the preimage of every point in $L\mathbb{F}_p^r$ has size p^{r-t} . \square

We now explicitly define the relevant lattice that the image of our pattern under \mathbf{B} will equidistribute over, in order to state the main result of this section. Recall we have a given $J \in \mathbb{F}_p^{k \times k}$. Let

$$\Xi_J = \{A \in \mathbb{F}_p^{k \times k} : (JA)^\top = JA\},$$

and let

$$\begin{aligned} \Lambda_J &= \{(-A, -A(I+J)(I-J)^{-1}, A(I+J)(I-J)^{-1}, A) : A^\top = +A, A \in \Xi_J\}, \\ \Lambda'_J &= \{(-A, -A(I+J)(I-J)^{-1}, A(I+J)(I-J)^{-1}, A) : A^\top = -A, A \in \Xi_J\}. \end{aligned}$$

Also, let

$$\Psi_J = \{(x_1, x_2, x_3, x_4) \in (\mathbb{F}_p^k)^4 : x_1 - x_2 - x_3 + x_4 = 0, x_4 - x_2 = J(x_2 - x_1)\}.$$

We make $\mathbb{F}_p^{k \times k}$ an inner product space with the standard inner product

$$\langle A, B \rangle = \langle A, B \rangle_{\text{HS}} = \text{tr}(A^\top B)$$

on $\mathbb{F}_p^{k \times k}$. We extend this inner product to $(\mathbb{F}_p^{k \times k})^4$ in the natural way, that is,

$$\langle (X_1, X_2, X_3, X_4), (Y_1, Y_2, Y_3, Y_4) \rangle = \langle X_1, Y_1 \rangle + \langle X_2, Y_2 \rangle + \langle X_3, Y_3 \rangle + \langle X_4, Y_4 \rangle.$$

We wish to study the equidistribution of the tuple

$$(\mathbf{B}(X), \mathbf{B}(X+D), \mathbf{B}(X+JD), \mathbf{B}(X+(I+J)D))$$

as X, D range over $\mathbb{F}_p^{k \times n}$, for a k -symmetrized quadratic factor \mathfrak{B} . Ultimately, we will find that the components corresponding to each $\mathbf{B}_{1,i}, \mathbf{B}_{2,i}, \mathbf{B}_{3,i}$ are all “independent”, and that each equidistributes in the following way:

$$\begin{aligned} (\mathbf{B}_{1,i}(X), \mathbf{B}_{1,i}(X+D), \mathbf{B}_{1,i}(X+JD), \mathbf{B}_{1,i}(X+(I+J)D)) & \text{ equidistributes on } \Psi_J, \\ (\mathbf{B}_{2,i}(X), \mathbf{B}_{2,i}(X+D), \mathbf{B}_{2,i}(X+JD), \mathbf{B}_{2,i}(X+(I+J)D)) & \text{ equidistributes on } \Lambda_J^\perp \cap (\mathcal{S}_k)^4, \\ (\mathbf{B}_{3,i}(X), \mathbf{B}_{3,i}(X+D), \mathbf{B}_{3,i}(X+JD), \mathbf{B}_{3,i}(X+(I+J)D)) & \text{ equidistributes on } \Lambda'_J{}^\perp \cap (\mathcal{S}'_k)^4. \end{aligned}$$

Here the \perp means the orthogonal subspace with respect to the inner product defined above. For ease of notation, we will write Λ_J^\perp for $\Lambda_J^\perp \cap (\mathcal{S}_k)^4$ and $\Lambda'_J{}^\perp$ for $\Lambda'_J{}^\perp \cap (\mathcal{S}'_k)^4$.

Theorem 4.4. *Suppose $J \in \mathbb{F}_p^{k \times k}$ is such that $J, I - J, I + J$ are invertible and J has no pair of eigenvalues that are negatives of each other (over $\overline{\mathbb{F}}_p$), let $G = \mathbb{F}_p^n$, and suppose \mathfrak{B} is a k -symmetrized quadratic factor of rank r . Then for any $a \in \Psi_J^{d_1} \times (\Lambda_J^\perp)^{d_2} \times (\Lambda_J^{\perp})^{d_3}$*

$$\begin{aligned} & \mathbb{P}_{X, D \in G^k}[(\mathbf{B}(X), \mathbf{B}(X + D), \mathbf{B}(X + JD), \mathbf{B}(X + (I + J)D)) = a] \\ &= p^{-d_1 \dim(\Psi_J) - d_2 \dim(\Lambda_J^\perp) - d_3 \dim(\Lambda_J^{\perp})} (1 + O(p^{-r/2 + 2kd_1 + (2\binom{k+1}{2} + k^2)d_2 + (2\binom{k}{2} + k^2)d_3})) \end{aligned}$$

The approach is similar to the proof of [Proposition 4.2](#). We want to consider (X, D) as a $2kn$ -dimensional vector and apply [Proposition 4.1](#), but now some linear dependencies will appear.² We will instead apply equidistribution on a set of “abstractly independent” forms to which we can indeed apply [Proposition 4.1](#). Then we realize $(\mathbf{B}(X), \mathbf{B}(X + D), \mathbf{B}(X + JD), \mathbf{B}(X + (I + J)D))$ as the image of those elements under a linear map, and apply [Lemma 4.3](#).

As a first step, we state the necessary equidistribution over these “abstract atoms”. For convenience, given a k -symmetrized quadratic factor \mathfrak{B} , define an attached map

$$\mathbf{B}'(X, D) = ((XM_i D^\top)_{i \in [d_2]}, (XN_i D^\top)_{i \in [d_3]}).$$

Proposition 4.5. *Suppose \mathfrak{B} is a k -symmetrized quadratic factor of rank r . Then for any $a \in ((\mathbb{F}_p^k)^{d_1} \times (\mathcal{S}_k)^{d_2} \times (\mathcal{S}'_k)^{d_3})^2 \times (\mathbb{F}_p^{k \times k})^{d_2 + d_3}$ we have*

$$\mathbb{P}_{X, D \in G^k}[(\mathbf{B}(X), \mathbf{B}(D), \mathbf{B}'(X, D)) = a] = p^{-2kd_1 - (2\binom{k+1}{2} + k^2)d_2 - (2\binom{k}{2} + k^2)d_3} + O(p^{-r/2}).$$

This follows immediately by applying [Proposition 4.1](#) to (X, D) viewed as an element of \mathbb{F}_p^{2kn} . The proof is exactly analogous to the proof of [Proposition 4.2](#) from [Proposition 4.1](#).

To complete the proof of [Theorem 4.4](#), we note that the desired map, $(X, D) \mapsto (\mathbf{B}(X), \mathbf{B}(X + D), \mathbf{B}(X + JD), \mathbf{B}(X + (I + J)D))$, can be written as the map $(X, D) \mapsto (\mathbf{B}(X), \mathbf{B}(D), \mathbf{B}'(X, D))$ composed with a linear transformation. For example,

$$(X + JD)M_i(X + JD)^\top = XM_i X^\top + J(DM_i X^\top) + XM_i D^\top J^\top + J(DM_i D^\top)J^\top$$

and $XM_i D^\top = (DM_i X^\top)^\top$. Therefore it suffices to understand the linear constraints induced by this last linear transformation.

4.2. Deriving the linear constraints. To this end we prove the following abstract linear algebra statement, which essentially encodes the eigenvalue condition in [Theorem 1.2](#).

Lemma 4.6. *If $A \in \mathbb{F}_p^{k \times k}$ is invertible and has no pair of eigenvalues (over $\overline{\mathbb{F}}_p$) which are negatives of each other, then $A \in \mathbb{F}_p[A^2]$.*

Proof. Given any matrix M , let $Q_M \in \mathbb{F}_p[t]$ be the monic polynomial of minimum degree satisfying $Q_M(M) = 0$ (this exists by the Cayley–Hamilton theorem and the fact that $\mathbb{F}_p[t]$ is a principal ideal domain). Then $\mathbb{F}_p[M] \cong \mathbb{F}_p[t]/(Q_M(t))$ as \mathbb{F}_p -algebras, and the dimension as an \mathbb{F}_p -vector space is $\deg Q_M$.

We clearly have $A \in \mathbb{F}_p[A^2]$ if and only if $\mathbb{F}_p[A] = \mathbb{F}_p[A^2]$, which will certainly follow from

$$\dim_{\mathbb{F}_p} \mathbb{F}_p[A] \leq \dim_{\mathbb{F}_p} \mathbb{F}_p[A^2]$$

due to the obvious containment. Now note that if $g(A^2) = 0$ then $Q_A(t)|g(t^2)$ in $\mathbb{F}_p[t]$. By hypothesis, we have $\gcd(Q_A(t), Q_A(-t)) = 1$, hence $Q_A(t)Q_A(-t)|g(t^2)$. Thus

$$2 \dim_{\mathbb{F}_p} \mathbb{F}_p[A] = \deg(Q_A(t)Q_A(-t)) \leq 2 \deg g.$$

Since this holds for all such g , it in particular holds for $g = Q_{A^2}$, which implies the result. \square

²These linear dependencies appear for the same reason that $(x^2, (x + d)^2, (x + 2d)^2, (x + 3d)^2)$ satisfies a linear equation.

Next we need the following abstract matrix equation which is used to derive the desired equidistribution statement.

Lemma 4.7. *Suppose $J \in \mathbb{F}_p^{k \times k}$ is such that $J, I - J, I + J$ are invertible and J has no pair of eigenvalues that are negatives of each other.*

- *Let $M = M^\top$ be nonzero. Then $(A_1, A_2, A_3, A_4) \in (\mathcal{S}_k)^4$ and*

$$\begin{aligned} & \text{tr}(A_1^\top X M X^\top + A_2^\top (X + D) M (X + D)^\top + A_3^\top (X + J D) M (X + J D)^\top \\ & \quad + A_4^\top (X + (I + J) D) M (X + (I + J) D)^\top) = 0 \end{aligned}$$

for all $X, D \in (\mathbb{F}_p^n)^k$ if and only if $(A_1, A_2, A_3, A_4) \in \Lambda_J$.

- *Let $M = -M^\top$ be nonzero. Then $(A_1, A_2, A_3, A_4) \in (\mathcal{S}'_k)^4$ and*

$$\begin{aligned} & \text{tr}(A_1^\top X M X^\top + A_2^\top (X + D) M (X + D)^\top + A_3^\top (X + J D) M (X + J D)^\top \\ & \quad + A_4^\top (X + (I + J) D) M (X + (I + J) D)^\top) = 0 \end{aligned}$$

for all $X, D \in (\mathbb{F}_p^n)^k$ if and only if $(A_1, A_2, A_3, A_4) \in \Lambda'_J$.

Proof. We prove the first claim as the second is analogous. Note that

$$\begin{aligned} & \text{tr}(A_1^\top X M X^\top + A_2^\top (X + D) M (X + D)^\top + A_3^\top (X + J D) M (X + J D)^\top \\ & \quad + A_4^\top (X + (I + J) D) M (X + (I + J) D)^\top) = 0 \end{aligned}$$

implies that

$$\begin{aligned} & \text{tr}((A_1^\top + A_2^\top + A_3^\top + A_4^\top) X M X^\top) = 0, \\ & \text{tr}(A_2^\top (D) M (D)^\top + A_3^\top (J D) M (J D)^\top + A_4^\top ((I + J) D) M ((I + J) D)^\top) = 0, \end{aligned}$$

taking $D = 0$ and $X = 0$ respectively. Using that $\text{tr}(\cdot)$ is additive along with the initial condition, we derive

$$\begin{aligned} & \text{tr}((A_1^\top + A_2^\top + A_3^\top + A_4^\top) X M X^\top) = 0, \\ & \text{tr}(A_2^\top (D) M (D)^\top + A_3^\top (J D) M (J D)^\top + A_4^\top ((I + J) D) M ((I + J) D)^\top) = 0, \\ & \text{tr}(A_2^\top (D M X^\top + X M D^\top) + A_3^\top (J D M X^\top + X M (J D)^\top) \\ & \quad + A_4^\top ((J + I) D M X^\top + X M ((J + I) D)^\top)) = 0. \end{aligned}$$

Using that trace $\text{tr}(A^\top B^\top) = \text{tr}(BA) = \text{tr}(AB)$ the above conditions are equivalent to

$$\begin{aligned} & \text{tr}((A_1^\top + A_2^\top + A_3^\top + A_4^\top) X M X^\top) = 0 \\ & \text{tr}((A_2^\top + J^\top A_3^\top J + (I + J)^\top A_4^\top (I + J)) (D M D)^\top) = 0 \\ & \text{tr}((2A_2^\top + 2A_3^\top J + 2A_4^\top (I + J)) (D M X^\top)) = 0. \end{aligned}$$

To derive the last, we used both $A_i^\top = A_i$ and $M^\top = M$. (In fact, one obtains identical equations in the skew-symmetric case.)

Since M is nonzero and symmetric we have that $\{X M X^\top\}$, $\{D M D^\top\}$ each span the space of all $k \times k$ symmetric matrices while $\{D M X^\top\}$ spans the space of all $k \times k$ matrices. This implies that

$$\begin{aligned} & A_1^\top + A_2^\top + A_3^\top + A_4^\top = 0 \\ & A_2^\top + J^\top A_3^\top J + (I + J)^\top A_4^\top (I + J) = 0 \\ & A_2^\top + A_3^\top J + A_4^\top (I + J) = 0 \end{aligned}$$

since the A_i are symmetric.

Note that the second and third equations imply that

$$(J^\top - I) A_2^\top = A_4^\top (I + J).$$

Noting that A_2^\top, A_4^\top are symmetric we find that

$$A_2^\top = (J^\top - I)^{-1}A_4^\top(I + J) = (I + J^\top)A_4^\top(J - I)^{-1}.$$

This is equivalent to

$$(J^\top)^2 A_4^\top = A_4^\top J^2.$$

It follows that for every polynomial $Q(t) \in \mathbb{F}_p[t]$ we have

$$Q((J^\top)^2)A_4^\top = A_4^\top Q(J^2).$$

Now $J \in \mathbb{F}_p[J^2]$ by [Lemma 4.6](#), so it follows that

$$J^\top A_4^\top = A_4^\top J.$$

A similar combination of the second and third equations implies that

$$J^\top A_2^\top = -A_3^\top J$$

and therefore

$$(J^\top)^{-1}A_3^\top J = (J^\top)A_3^\top J^{-1}.$$

This similarly implies that $J^\top A_3^\top = A_3^\top J$. Therefore we deduce

$$\begin{aligned} A_1^\top + A_2^\top + A_3^\top + A_4^\top &= 0 \\ A_2^\top + A_3^\top J^2 + A_4^\top (I + J)^2 &= 0 \\ A_2^\top + A_3^\top J + A_4^\top (I + J) &= 0. \end{aligned}$$

Subtracting the last two equations gives that

$$A_3^\top = A_4^\top (I + J)(I - J)^{-1}.$$

Substituting into the last equation gives

$$A_2^\top = A_4^\top (I + J)(-I - J(I - J)^{-1}) = -A_3^\top.$$

Finally using the first equation this implies that $A_1^\top = -A_4^\top$. Therefore we have proven that $(A_1, A_2, A_3, A_4) \in \Lambda_J$. The reverse implication is a straightforward calculation which we omit. \square

We are now ready to prove [Theorem 4.4](#).

Proof of Theorem 4.4. This is almost immediate from [Proposition 4.5](#) and [Lemmas 4.3](#) and [4.7](#). The noted fact that the desired function is the image of the function in [Proposition 4.5](#) under a linear mapping along with [Lemma 4.3](#) demonstrates that there is equidistribution over some subspace (with the multiplicative error term preserved).

This subspace is precisely the span of all possible vectors $(B(X), B(X + D), B(X + JD), B(X + (I + J)D))$. Due to the independence demonstrated in [Proposition 4.5](#), we see that the subspace factors as a direct sum.

[Lemma 4.7](#) characterizes the resulting vector spaces for $B_{2,i}, B_{3,i}$, since it demonstrates the form of all orthogonal vectors in corresponding host spaces (either tuples of symmetric or skew-symmetric matrices). For B_1 , it is easy to check that vectors of the form $(Xr, (X + D)r, (X + JD)r, (X + (I + J)D)r)$ span the space Ψ_J when $r \in \mathbb{F}_p^n \setminus 0$. Indeed, the orthogonal vectors of the form $(\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_3)$ are precisely those that satisfy

$$\vec{a}_1 + \vec{a}_2 + \vec{a}_3 + \vec{a}_4 = \vec{a}_2 + \vec{a}_3 J + \vec{a}_4 (I + J) = 0.$$

All such vectors are spans of $(\vec{t}, -\vec{t}, -\vec{t}, \vec{t})$ and $(J\vec{t}, -(I + J)\vec{t}, 0, \vec{t})$, which matches the orthogonal space of Ψ_J . This completes the proof. \square

5. POPULAR DIFFERENCES FOR 4-POINT PATTERNS

We now have the tools to prove the following popular difference result for four-point patterns. This theorem immediately implies the result stated in the introduction, [Theorem 1.2](#).

Theorem 5.1. *Fix $k \geq 1$ and an odd prime p . Let M_1, M_2 be $k \times k$ matrices with coefficients in \mathbb{F}_p such that $M_1, M_2, M_1 - M_2$, and $M_1 + M_2$ are invertible and no pair of eigenvalues of $M_1 M_2^{-1}$ (viewed over $\overline{\mathbb{F}}_p$) are negatives of each other. For any $\alpha, \epsilon > 0$, letting $G = \mathbb{F}_p^n$ for $n > n_0(\alpha, \epsilon, p)$, if $f: G^k \rightarrow [0, 1]$ satisfies $\mathbb{E}_{X \in G^k} f(X) \geq \alpha$ then there are $\Omega_{\alpha, \epsilon, p}(p^{kn})$ values $D \in G^k$ such that*

$$\mathbb{E}_{X \in G^k} f(X) f(X + M_1 D) f(X + M_2 D) f(X + (M_1 + M_2) D) \geq \alpha^4 - \epsilon.$$

Given the previous developments the proof is similar to that of [[10](#), Theorem 4.1] modulo deriving the necessary positivity.

Proof of Theorem 5.1. By replacing $M_1 D$ by D (note that M_1 is invertible) we can reduce the case of (M_1, M_2) to the case of $(I, M_2 M_1^{-1})$. Hence from now on we assume that $M_1 = I$ and $M_2 = J$. Applying [Theorem 3.2](#), we decompose

$$f = f_{\text{str}} + f_{\text{sml}} + f_{\text{psr}}$$

where $f_{\text{str}} = \mathbb{E}[f|\mathfrak{B}]$ for a k -symmetrized quadratic factor \mathfrak{B} of complexity (d_1, d_2, d_3) and rank at least $\omega_1(d_1 + d_2 + d_3)$, where $\|f_{\text{sml}}\|_2 \leq \epsilon/250$, and where $\|f_{\text{psr}}\|_{U^3(G^k)} \leq 1/\omega_2(d_1 + d_2 + d_3)$. The complexity (d_1, d_2, d_3) is bounded in terms of the given parameters and growth functions. We define H to be the subspace of $(\mathbb{F}_p^n)^k$ such that the linear factors of \mathfrak{B} are zero. We will prove that

$$\mathbb{E}_{X, D \in G^k} [f(X) f(X + D) f(X + JD) f(X + (J + I)D) \mathbb{1}_H(D)] \geq p^{-kd_1} (\alpha^4 - \epsilon). \quad (5.1)$$

This suffices to prove the result as we are summing over a density p^{-kd_1} subset of the differences D and hence at least one difference achieves the necessary bound. Furthermore, by Markov's inequality, a fraction of at least $\Omega_{\alpha, \epsilon}(1)$ of the differences in H satisfy the weaker bound of $\alpha^4 - 2\epsilon$. Adjusting ϵ appropriately will prove the desired upon noting that a positive fraction of G^k lies in H due to the bounded complexity of \mathfrak{B} .

Now we focus attention on [\(5.1\)](#). Expanding $f = f_{\text{str}} + f_{\text{sml}} + f_{\text{psr}}$ allows us to turn the left side into 81 terms of the form

$$\mathbb{E}_{X, D \in G^k} [f_1(X) f_2(X + D) f_3(X + JD) f_4(X + (J + I)D) \mathbb{1}_H(D)]$$

where $f_1, f_2, f_3, f_4 \in \{f_{\text{str}}, f_{\text{sml}}, f_{\text{psr}}\}$. If f_{sml} appears in the expression, we have that

$$\begin{aligned} |\mathbb{E}_{X, D} [f_1(X) f_2(X + D) f_3(X + JD) f_4(X + (J + I)D) \mathbb{1}_H(D)]| \\ \leq \mathbb{E}_X [|f_{\text{sml}}(X)| \mathbb{1}_H(D)] \leq p^{-kd_1} (\mathbb{E}_X [f_{\text{sml}}(X)^2])^{1/2} \leq p^{-kd_1} \epsilon/250. \end{aligned}$$

This bounds the 65 terms that include f_{sml} .

Next we bound the terms that include f_{psr} . Say $f_3 = f_{\text{psr}}$ (the other cases are analogous). Note that we have

$$\mathbb{1}_H(D) = \sum_{T \in H^\perp} \mathbb{1}_{T+H}(X + D) \mathbb{1}_{T+H}(X)$$

for all X . Therefore

$$\begin{aligned} & |\mathbb{E}[f_1(X) f_2(X + D) f_{\text{psr}}(X + JD) f_4(X + (I + J)D) \mathbb{1}_H(D)]| \\ &= \left| \sum_{T \in H^\perp} \mathbb{E}[f_1(X) f_2(X + D) f_{\text{psr}}(X + JD) f_4(X + (I + J)D) \mathbb{1}_{T+H}(X + D) \mathbb{1}_{T+H}(X)] \right| \\ &\leq |H^\perp| \|f_{\text{psr}}\|_{U^3(G^k)} \end{aligned}$$

where in the final line we have used [Lemma 2.2](#) (since $I, J, I + J, I - J$ are invertible). Taking the growth function ω_2 to be a sufficiently fast growing exponential (depending on p, k), we can ensure that $\|f_{\text{psr}}\|_{U^3(G^k)} \leq p^{-2kd_1} \epsilon / 250$.

Thus it suffices to prove

$$\mathbb{E}[f_{\text{str}}(X)f_{\text{str}}(X + D)f_{\text{str}}(X + JD)f_{\text{str}}(X + (J + I)D)\mathbb{1}_H(D)] \geq p^{-kd_1}(\alpha^4 - \epsilon/2). \quad (5.2)$$

Since $f_{\text{str}} = \mathbb{E}[f|\mathfrak{B}]$, we see that f_{str} is \mathfrak{B} -measurable and $[0, 1]$ -valued. Recall that a factor \mathfrak{B} defines a $\mathbf{B}: G^k = (\mathbb{F}_p^n)^k \rightarrow (\mathbb{F}_p^k)^{d_1} \times (\mathcal{S}_k)^{d_2} \times (\mathcal{S}'_k)^{d_3}$. Hence there is an associated function $\mathbf{f}: (\mathbb{F}_p^k)^{d_1} \times (\mathcal{S}_k)^{d_2} \times (\mathcal{S}'_k)^{d_3} \rightarrow [0, 1]$ such that $f_{\text{str}}(X) = \mathbf{f}(\mathbf{B}(X))$.

Claim 5.2.

$$\mathbb{E}[f_{\text{str}}(X)f_{\text{str}}(X + D)f_{\text{str}}(X + JD)f_{\text{str}}(X + (J + I)D)\mathbb{1}_H(D)]$$

and

$$p^{-kd_1} \mathbb{E}_{\vec{v} \in \mathbb{F}_p^k, \vec{M}_2 \in (\Lambda_J^\perp)^{d_2}, \vec{M}_3 \in (\Lambda'_J)^{d_3}} [\mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})\mathbf{f}(\vec{v}, M_2^{(2)}, M_3^{(2)})\mathbf{f}(\vec{v}, M_2^{(3)}, M_3^{(3)})\mathbf{f}(\vec{v}, M_2^{(4)}, M_3^{(4)})]$$

are equal to up to a multiplicative factor of $1 + O(p^{-r/2+2kd_1+(2\binom{k+1}{2}+k^2)d_2+(2\binom{k}{2}+k^2)d_3})$ multiplicative factor, where the rank r is the rank of \mathfrak{B} .

Proof. Restricting D to lie in H we need to understand the equidistribution of $(\mathbf{B}(X), \mathbf{B}(X + D), \mathbf{B}(X + JD), \mathbf{B}(X + (I + J)D))$. We apply [Theorem 4.4](#). We have that $(\mathbf{B}_1(X), \mathbf{B}_1(X + D), \mathbf{B}_1(X + JD), \mathbf{B}_1(X + (I + J)D))$ equidistributes over $(\vec{v}, \vec{v}, \vec{v}, \vec{v})$ with $\vec{v} \in (\mathbb{F}_p^k)^{d_1}$, since looking at $D \in H$ corresponds precisely to looking at atoms for which linear factors in \mathfrak{B} for $\mathbf{B}(X), \mathbf{B}(X + D)$ are equal. (We are also implicitly using that $D \in H$ implies each row of D is in the orthogonal space of the defining vectors $\{r_1, \dots, r_{d_1}\}$ of the linear factors of \mathfrak{B} , which implies $JD \in H$ as well.)

Furthermore

$$(M_2^{(1)}, M_2^{(2)}, M_2^{(3)}, M_2^{(4)}) = (\mathbf{B}_2(X), \mathbf{B}_2(X + D), \mathbf{B}_2(X + JD), \mathbf{B}_2(X + (I + J)D))$$

equidistributes over $(\Lambda_J^\perp)^{d_2}$ and finally

$$(M_3^{(1)}, M_3^{(2)}, M_3^{(3)}, M_3^{(4)}) = (\mathbf{B}_3(X), \mathbf{B}_3(X + D), \mathbf{B}_3(X + JD), \mathbf{B}_3(X + (I + J)D))$$

equidistributes over $(\Lambda'_J)^{d_3}$ with each of the components distributing independently. Therefore, even upon restricting the image of \mathbf{B}_1 to be zero, we have equidistribution over the remaining space. \square

It remains to study the expectation in [Claim 5.2](#). Recalling the definitions of $\Xi_J, \Lambda_J, \Lambda'_J$ from [Section 4](#), we define the following subspaces:

$$\begin{aligned} \Omega_J &= \{(-A, -A(I + J)(I - J)^{-1}): A^\top = +A, A \in \Xi_J\}, \\ \Omega'_J &= \{(-A, -A(I + J)(I - J)^{-1}): A^\top = -A, A \in \Xi_J\}. \end{aligned}$$

Claim 5.3. *Given $(M^{(1)}, M^{(2)}, M^{(3)}, M^{(4)}) \in \mathcal{S}_k^4$, we have $(M^{(1)}, M^{(2)}, M^{(3)}, M^{(4)}) \in \Lambda_J^\perp$ if and only if we have the equality of cosets*

$$(M^{(1)}, M^{(2)}) + \Omega_J^\perp = (M^{(4)}, M^{(3)}) + \Omega_J^\perp.$$

Similarly given $(M^{(1)}, M^{(2)}, M^{(3)}, M^{(4)}) \in \mathcal{S}_k^4$, we have $(M^{(1)}, M^{(2)}, M^{(3)}, M^{(4)}) \in \Lambda'_J^\perp$ if and only if we have the equality of cosets

$$(M^{(1)}, M^{(2)}) + \Omega'_J^\perp = (M^{(4)}, M^{(3)}) + \Omega'_J^\perp.$$

This claim follows by inspection of the definitions of Λ_J and Ω_J . From this we deduce

$$\begin{aligned}
& \mathbb{E}[\mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})\mathbf{f}(\vec{v}, M_2^{(2)}, M_3^{(2)})\mathbf{f}(\vec{v}, M_2^{(3)}, M_3^{(3)})\mathbf{f}(\vec{v}, M_2^{(4)}, M_3^{(4)})] \\
&= \mathbb{E}_{\vec{v}}\mathbb{E}_{\tau} \left(\mathbb{E}_{((M_2^{(1)}, M_2^{(2)})+\Omega_J^\perp, (M_3^{(1)}, M_3^{(2)})+\Omega_J'^\perp)}_{=\tau} \mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})\mathbf{f}(\vec{v}, M_2^{(2)}, M_3^{(2)}) \right)^2 \\
&\geq \left(\mathbb{E}_{\vec{v}}\mathbb{E}_{\tau} \mathbb{E}_{((M_2^{(1)}, M_2^{(2)})+\Omega_J^\perp, (M_3^{(1)}, M_3^{(2)})+\Omega_J'^\perp)}_{=\tau} \mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})\mathbf{f}(\vec{v}, M_2^{(2)}, M_3^{(2)}) \right)^2 \\
&= \left(\mathbb{E}_{\vec{v}}\mathbb{E}_{(M_2^{(1)}, M_2^{(2)}, M_3^{(1)}, M_3^{(2)})} \mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})\mathbf{f}(\vec{v}, M_2^{(2)}, M_3^{(2)}) \right)^2 \\
&= \mathbb{E}_{\vec{v}} \left(\mathbb{E}_{(M_2^{(1)}, M_3^{(1)})} \mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})^2 \right)^2 \\
&\geq \left(\mathbb{E}_{\vec{v}, (M_2^{(1)}, M_3^{(1)})} \mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)}) \right)^4
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality twice. To finish note that by the equidistribution derived in [Proposition 4.2](#) we have that $\mathbb{E}_{\vec{v}, (M_2^{(1)}, M_3^{(1)})} \mathbf{f}(\vec{v}, M_2^{(1)}, M_3^{(1)})$ is $\mathbb{E}[f_{\text{str}}]$ up to a multiplicative factor of $1 + O(p^{-r/2+kd_1+\binom{k+1}{2}d_2+\binom{k}{2}d_3})$. Finally since $\mathbb{E}[f_{\text{str}}] = \mathbb{E}[\mathbb{E}[f|\mathfrak{B}]] = \alpha$ the desired result follows upon taking the growth function ω_2 (and thus r , the rank of \mathfrak{B}) to be large enough. \square

Remark. The key reason that this argument works is a “positivity” result in the final expectation. This “positivity” occurs from the symmetry of Λ_J^\perp which ultimately is derived from the spectral condition on J . In the next section we consider what occurs when the spectral condition on J is no longer satisfied. In essence, Λ_J^\perp will be one dimension larger than otherwise in such a way that it no longer has this symmetry property.

6. COUNTEREXAMPLE TO POPULAR DIFFERENCES FOR ROTATED SQUARES IN \mathbb{F}_5^n

Given the results of the last section it is natural to ask whether the popular difference result holds for all matrix patterns which are controlled by the U^3 -norm. We now prove that this is false and demonstrate [Theorem 1.3](#).

As is standard, it suffices to give a construction for a set of fixed positive density, as smaller sets will follow from subsampling and simple concentration facts. In fact, we will merely give a function in $[0, 1]$ (which one can scale down appropriately and sample from). The counterexample proceeds in stages. In the first stage we give a construction which rules out all sufficiently generic differences (i.e., (a, b) such that $\text{span}_{\mathbb{F}_5}\{a, b\}$ is 2-dimensional), using the failure of the key positivity in [Section 4](#). We then modify the function to rule out the non-generic directions using techniques from earlier work of the second and third authors with Zhao in [[18](#), Sections 2, 3] (which deals with, for instance, axis-aligned squares) with [[18](#), Section 3] itself building on a construction of Mandache [[15](#)]. Finally, we expect that the construction here can be used to disprove the ergodic analogue [[1](#), Question 1.11] but we do not pursue this direction here.

First, if Γ is the linear automorphism of $(\mathbb{F}_5^n)^2$ defined by $(x, y) \mapsto (x - 2y, x + 2y)$ then by replacing A by $\Gamma^{-1}A$ and reparametrizing the pattern we are counting, it suffices to consider the “diagonalized” pattern

$$(x, y), (x + a, y + b), (x + 2a, y - 2b), (x + 3a, y - b).$$

(Note that this is two arithmetic progressions in the coordinates with a twist, though we will not make use of this fact)

6.1. Initial construction. Let $f_1: (\mathbb{F}_5^n)^2 \rightarrow [0, 1]$ be defined by

$$f_1(x, y) = g_1(x \cdot x, x \cdot y),$$

where $g_1: \mathbb{F}_5^2 \rightarrow [0, 1]$ is to be chosen later. With this choice, the pattern count for f depends only on g_1 and the distribution of 8-tuples of the form

$$(x \cdot x, x \cdot y, \dots, (x + 3a) \cdot (x + 3a), (x + 3a) \cdot (y - b)).$$

Let $\Lambda'_2 \leq \mathbb{F}_5^8$ be the space orthogonal to the vectors

$$(1, 0, -1, 0, -1, 0, 1, 0), \quad (0, 1, 0, -1, 0, -1, 0, 1), \quad (1, 0, -3, 0, 3, 0, -1, 0)$$

We claim that these 8-tuples equidistribute over Λ'_2 .

Proposition 6.1. *Fix nonzero $a, b \in \mathbb{F}_5^n$ that are not multiples of each other. Then*

$$(x \cdot x, x \cdot y, (x + a) \cdot (x + a), (x + a) \cdot (y + b), (x + 2a) \cdot (x + 2a), \\ (x + 2a) \cdot (y - 2b), (x + 3a) \cdot (x + 3a), (x + 3a) \cdot (y - b))$$

obtains every value in $(0, 0, a \cdot a, a \cdot b, 4a \cdot a, -4a \cdot b, 9a \cdot a, -3a \cdot b) + \Lambda'_2$ with probability

$$5^{-5} + O(5^{-n/2})$$

Remark. Note that $(0, 0, 0, 1, 0, -4, 0, -3) \in \Lambda'_2$, so it actually equidistributes in $(0, 0, a \cdot a, 0, 4a \cdot a, 0, 9a \cdot a, 0) + \Lambda'_2$.

Proof. This is a hands-on computation of equidistribution. Apply [Proposition 4.1](#) to the concatenation $x' = (x, y)$ with linear forms $x \cdot a, x \cdot b, a \cdot y$ and quadratic forms $x \cdot x, x \cdot y$ (the last of which can be written as a symmetric matrix when $p \neq 2$). We conclude that the image of (x, y) under this 5-tuple of forms obtains each point in \mathbb{F}_5^5 with probability $5^{-5} + O(5^{-n/2})$. Given these five values one can solve for all values in the 8-tuple, and solving we obtain equidistribution of the 8-tuple over the appropriate 5-dimensional subspace Λ'_2 as desired. \square

Consequently, if $a, b \in \mathbb{F}_5^n$ are nonzero and not multiples of each other,

$$\begin{aligned} \beta_1(a, b) &:= \mathbb{E} f_1(x, y) f_1(x + a, y + b) f_1(x + 2a, y - 2b) f_1(x + 3a, y - b) \\ &= \mathbb{E} g_1(x \cdot x, x \cdot y) g_1((x + a) \cdot (x + a), (x + a) \cdot (y + b)) g_1((x + 2a) \cdot (x + 2a), (x + 2a) \cdot (y - 2b)) \\ &\quad g_1((x + 3a) \cdot (x + 3a), (x + 3a) \cdot (y - b)) \\ &= \mathbb{E}_{v \in \Lambda'_2} g_1(v_1, v_2) g_1(v_3 + a \cdot a, v_4) g_1(v_5 + 4a \cdot a, v_6) g_1(v_7 + 9a \cdot a, v_8) + O(5^{-n/2}), \end{aligned} \quad (6.1)$$

using the remark after [Proposition 6.1](#) in the last line.

Let $g_1(x, y) = \mathbb{1}_S(x, y)$ be the indicator of the set

$$S = \{(0, 2), (0, 3), (0, 4), (1, 0), (1, 3), (1, 4), (2, 1), (2, 2), (3, 0), (3, 1)\}.$$

Then a direct verification proves that

$$\sup_{a \cdot a \in \mathbb{F}_5} \mathbb{E}_{v \in \Lambda'_2} g_1(v_1, v_2) g_1(v_3 + a \cdot a, v_4) g_1(v_5 + 4a \cdot a, v_6) g_1(v_7 + 9a \cdot a, v_8) = \frac{73}{5^5}$$

while clearly

$$\alpha_1 = \mathbb{E} f_1(x, y) = \mathbb{E}_{v_1, v_2 \in \mathbb{F}_5} g_1(v_1, v_2) + O(5^{-n/2}) = \frac{2}{5} + O(5^{-n/2})$$

by [Proposition 6.1](#). (Code verifying this explicit finite computation is attached on the arXiv.)

If we merely wish to establish [Theorem 1.3](#) for all but $O(\sqrt{|G|})$ directions, where $G = (\mathbb{F}_5^n)^2$, then we are done due to $73/5^5 < (2/5)^4$. Now we introduce further constructions which allow us to fix the directions where a, b are nontrivially related.

6.2. Fixing most special directions. We next proceed with a modification of a construction which appears in [15] and was extended in [18]. Let $X_x, Y_y, Z_z, X'_x, Y'_y, Z'_z$ for $x, y, z \in \mathbb{F}_5^n$ be independent random variables uniform on $[0, 1]$. Let $F_2, F_3: (\mathbb{F}_5^n)^2 \rightarrow [0, 1]$ be defined by

$$F_2(x, y) = g_2(X_{-x-y}, Y_{-2x+2y}, Z_{2x+y}), \quad F_3(x, y) = g_2(X'_{-x-2y}, Y'_{-2x-y}, Z'_{2x+2y})$$

where $g_2: [0, 1]^3 \rightarrow [0, 1]$ is a function to be chosen later. This is a random function. Let $h(x, y) = f_1(x, y)F_2(x, y)F_3(x, y)$, which is also a random function. Let

$$\alpha_h = \mathbb{E}_{x, y \in \mathbb{F}_5^n} h(x, y)$$

and let

$$\beta_h(a, b) = \mathbb{E}_{x, y \in \mathbb{F}_5^n} h(x, y)h(x+a, y+b)h(x+2a, y-2b)h(x+3a, y-b),$$

which are both random in the $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{X}', \mathbf{Y}', \mathbf{Z}'$ variables. We have

$$\begin{aligned} \alpha_2 &= \mathbb{E}_{\mathbf{X}, \mathbf{Y}, \mathbf{Z}} \alpha_h = \mathbb{E}_{x, y \in \mathbb{F}_5^n} \mathbb{E}_{\substack{\mathbf{X}, \mathbf{Y}, \mathbf{Z} \\ \mathbf{X}', \mathbf{Y}', \mathbf{Z}'}} f_1(x, y) g_2(X_{-x-y}, Y_{-2x+2y}, Z_{2x+y}) g_2(X'_{-x-y}, Y'_{-2x+2y}, Z'_{2x+y}) \\ &= \mathbb{E}_{x, y \in \mathbb{F}_5^n} f_1(x, y) \cdot (\mathbb{E}_{u, v, w \in [0, 1]} g_2(u, v, w))^2 \end{aligned}$$

since $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ are independent, etc. We also have

$$\begin{aligned} \beta_2(a, b) &= \mathbb{E}_{\mathbf{X}, \mathbf{Y}, \mathbf{Z}} \beta_h(a, b) \\ &= \mathbb{E}_{x, y \in \mathbb{F}_5^n} \mathbb{E}_{\substack{\mathbf{X}, \mathbf{Y}, \mathbf{Z} \\ \mathbf{X}', \mathbf{Y}', \mathbf{Z}'}} f_1(x, y) f_1(x+a, y+b) f_1(x+2a, y-2b) f_1(x+3a, y-b) \\ &\quad g_2(X_{-x-y}, Y_{-2x+2y}, Z_{2x+y}) g_2(X_{-x-y-a-b}, Y_{-2x+2y-2a+2b}, Z_{2x+y+2a+b}) \\ &\quad g_2(X_{-x-y-2a+2b}, Y_{-2x+2y+a+b}, Z_{2x+y-a-2b}) g_2(X_{-x-y+2a+b}, Y_{-2x+2y-a-2b}, Z_{2x+y+a-b}) \\ &\quad g_2(X'_{-x-2y}, Y'_{-2x-y}, Z'_{2x+2y}) g_2(X'_{-x-2y-a-2b}, Y'_{-2x-y-2a-b}, Z'_{2x+2y+2a+2b}) \\ &\quad g_2(X'_{-x-2y-2a-b}, Y'_{-2x-y+a+2b}, Z'_{2x+2y-a+b}) g_2(X'_{-x-2y+2a+2b}, Y'_{-2x-y-a+b}, Z'_{2x+2y+a-2b}). \end{aligned}$$

If $a, b \in \mathbb{F}_5^n$ are not linearly dependent, then we easily see that all the terms in the product are independent, and linearity of expectation demonstrates

$$\begin{aligned} \beta(a, b) &= \mathbb{E}_{x, y \in \mathbb{F}_5^n} f_1(x, y) f_1(x+a, y+b) f_1(x+2a, y-2b) f_1(x+3a, y-b) (\mathbb{E} g_2(u, v, w))^8 \\ &= \beta_1(a, b) (\mathbb{E} g_2(u, v, w))^8. \end{aligned}$$

Otherwise we have cases depending on the 6 possible linear dependencies. The key point is that, for example,

$$\mathbb{E}_{\mathbf{X}, \mathbf{Y}, \mathbf{Z}} g_2(X_0, Y_0, Z_0) g_2(X_0, Y_1, Z_1) = \mathbb{E}_{u_0, v_0, w_0, v_1, w_1 \in [0, 1]} g_2(u_0, v_0, w_0) g_2(u_0, v_1, w_1),$$

so in each of these 6 cases we can reduce to some sort of expectation of g_2 , or rather, the product of two such terms coming from the independent sets of variables $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ and $(\mathbf{X}', \mathbf{Y}', \mathbf{Z}')$.

Explicit computation yields

$$\begin{aligned} \beta_2(a, 0) &= \beta_1(a, 0) (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_1, v_1, w_1) g_2(u_2, v_2, w_2) g_2(u_3, v_3, w_3))^2 \\ &= \beta_1(a, 0) (\mathbb{E} g_2(u, v, w))^8, \\ \beta_2(a, a) &= \beta_1(a, a) (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_1, v_0, w_1) g_2(u_0, v_2, w_2) g_2(u_1, v_2, w_0)) \\ &\quad (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_1, v_1, w_1) g_2(u_1, v_2, w_0) g_2(u_3, v_0, w_1)), \\ \beta_2(a, -a) &= \beta_1(a, -a) (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_0, v_1, w_1) g_2(u_2, v_0, w_1) g_2(u_2, v_1, w_3)) \\ &\quad (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_1, v_1, w_0) g_2(u_2, v_1, w_2) g_2(u_0, v_3, w_2)), \\ \beta_2(a, 2a) &= \beta_1(a, 2a) (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_1, v_1, w_1) g_2(u_1, v_2, w_0) g_2(u_3, v_0, w_1)) \\ &\quad (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_0, v_1, w_1) g_2(u_2, v_0, w_1) g_2(u_2, v_1, w_3)), \\ \beta_2(a, -2a) &= \beta_1(a, -2a) (\mathbb{E} g_2(u_0, v_0, w_0) g_2(u_1, v_1, w_0) g_2(u_2, v_1, w_2) g_2(u_0, v_3, w_2)) \end{aligned}$$

$$\begin{aligned}
& (\mathbb{E}g_2(u_0, v_0, w_0)g_2(u_1, v_0, w_1)g_2(u_0, v_2, w_2)g_2(u_1, v_2, w_0)), \\
\beta_2(0, b) &= \beta_1(0, b)(\mathbb{E}g_2(u_0, v_0, w_0)g_2(u_1, v_1, w_1)g_2(u_2, v_2, w_2)g_2(u_3, v_3, w_3))^2 \\
&= \beta_1(0, b)(\mathbb{E}g_2(u, v, w))^8.
\end{aligned}$$

Here all variables other than $a, b \in \mathbb{F}_5^n$ are uniform on $[0, 1]$. It is worth noting every complicated product of expectations for the middle four terms are very similar. In particular, they are all equal to the product of the densities of two tripartite 3-uniform hypergraphs in the symmetric tripartite hypergraphon g_2 (with embeddings respecting the tripartition). Furthermore, the two hypergraphs attained for each term is the same pair of hypergraphs if we disregard the data of the tripartition. For the function g_2 we will choose, these considerations will not affect the bounds we prove, so we will focus on a single term.

We first define g_2 . Let $L \geq 1$ and let Λ be a subset of $\mathbb{Z}/L\mathbb{Z}$ avoiding arithmetic progressions of length 3 of size $L \exp(-C\sqrt{\log L})$ for some absolute constant $C > 0$. Let $U = V = W = \mathbb{Z}/L\mathbb{Z}$ and H be a tripartite graph on $U \times V \times W$ with $(s, s+t) \in U \times V$ an edge when $t \in \Lambda$, $(s, s+t) \in V \times W$ an edge when $t \in \Lambda$, and $(s, s+2t) \in U \times W$ an edge when $t \in \Lambda$. Note that the triangles in H are of the form $(s, s+t, s+2t)$ for $t \in \Lambda$ precisely since Λ has no nontrivial arithmetic progressions of length 3. This has the special property that every edge is in a unique triangle of H (this is the Ruzsa-Szemerédi graph).

Now let $g_2(u, v, w) = 1$ if $(\lfloor Lu \rfloor, \lfloor Lv \rfloor, \lfloor Lw \rfloor) \pmod L$ encodes one of these triangles. We have

$$\mathbb{E}g_2(u, v, w) = \frac{L|\Lambda|}{L^3} = L^{-1} \exp(-C\sqrt{\log L}).$$

Now we find

$$\mathbb{E}g_2(u_0, v_0, w_0)g_2(u_1, v_0, w_1)g_2(u_0, v_2, w_2)g_2(u_1, v_2, w_0).$$

When we refer to $u \in [0, 1]$ in the following discussion, we mean $\lfloor Lu \rfloor \pmod L$. In order for a term to be 1, we must have $u_0v_0w_0, u_1v_0w_1, u_0v_2w_2, u_1v_2w_0$ be triangles. But then u_0w_0, u_0v_2, v_2w_0 are all edges in H , so $u_0v_2w_0$ is also a triangle. This forces $v_2 = v_0$ (as edges are in unique triangles). Then $u_1v_0w_0$ is a triangle, so $u_1 = u_0$. Finally, this means $u_0v_0w_1$ is a triangle so $w_0 = w_1$. Thus the expectation equals

$$L^{-4}\mathbb{E}g_2(u, v, w) = \frac{L|\Lambda|}{L^7} = L^{-5} \exp(-C\sqrt{\log L}).$$

Next, we find

$$\mathbb{E}g_2(u_0, v_0, w_0)g_2(u_1, v_1, w_1)g_2(u_1, v_2, w_0)g_2(u_3, v_0, w_1) \leq \frac{L^4}{L^8} = L^{-4}$$

since there are at most L^4 ways to choose w_0, u_1, w_1, v_0 . After that, the variables u_0, v_1, v_2, u_3 are forced in order to guarantee a nonzero term, since edges are in unique triangles. This demonstrates that

$$\beta_2(a, \lambda a) \leq \beta_1(a, \lambda a)L^{-9} \exp(-C\sqrt{\log L}) \leq L^{-9+o(1)}$$

for $\lambda \in \mathbb{F}_5^*$. Now we take L sufficiently large. First, if $a, b \in \mathbb{F}_5^n$ are linearly independent, then by [Section 6.1](#) and the above we have

$$\beta_2(a, b) = \beta_1(a, b)(\mathbb{E}g_2(u, v, w))^8 \leq \left(\frac{73}{80} + O(5^{-n/2}) \right) \alpha_2^4.$$

On the other hand, if $b = \lambda a$ for $\lambda \in \mathbb{F}_5^*$ then

$$\beta_2(a, b) \leq L^{-9+o(1)} \leq \frac{1}{2}\alpha_2^4$$

since α_2 grows as $L^{-2+o(1)}$ times $\mathbb{E}f_1(x, y) = 2/5 + O(5^{-n/2})$.

Therefore, in expectation our random function satisfies the conclusion of [Theorem 1.3](#) for $(a, b) \in (\mathbb{F}_5^n \setminus 0)^2$. To obtain a single function which satisfies all these inequalities (of which there are $O(|G|)$), we use a concentration argument. A straightforward extension of [[18](#), Lemma 3.2] suffices. Roughly note that each of the above quantities can be controlled using Azuma–Hoeffding inequality on the Doob-martingale where one successively conditions on each of the random variables $X_x, Y_y, Z_z, X'_x, Y'_y, Z'_z$ for $x, y, z \in \mathbb{F}_5^n$ as each random variable only participates in a small number of terms. The quality of concentration is at least $\exp(-|G|^c)$ for deviations on the scale of $1/|G|^{1/4}$, which is more than sufficient. Let $h(x, y)$ now denote a specific instantiation of the above defined random function which satisfies

$$\sup_{(a,b) \in (\mathbb{F}_5^n \setminus 0)^2} \mathbb{E}[h(x, y)h(x + a, y + b)h(x + 2a, y - 2b)h(x + 3a, y - b)] \leq (1 - \delta)\mathbb{E}[h(x, y)]^4$$

for an absolute constant $\delta > 0$.

6.3. Finishing the construction. Finally we are in position to fix the directions (a, b) where $a = 0$ or $b = 0$. We now use a randomized version of the construction from [[18](#), Section 2] in order to eliminate these final special differences. Define $T = \{0, 1, 2\}^\gamma \times (\mathbb{F}_5)^{n-\gamma}$, where $\gamma \geq 1$ is an integer to be chosen later. Note that this set has density $\beta = (3/5)^\gamma$ but a four term arithmetic progression density of $(3/25)^\gamma \leq \beta^{4.15}$ noting that the set $\{0, 1, 2\}$ has no nontrivial four term arithmetic progressions. Now for each $g \in \mathbb{F}_5^n$ choose a two uniformly random bijective affine transformations $\phi(g), \phi'(g)$ of \mathbb{F}_5^n . Then let

$$f(x, y) = h(x, y)\mathbb{1}_{x \in \phi(y)S}\mathbb{1}_{y \in \phi'(x)S}.$$

First, we have

$$\mathbb{E}_{\phi, \phi'}[\mathbb{E}_{x, y} f(x, y)] = \mathbb{E}_{x, y}[h(x, y)\mathbb{E}_{\phi, \phi'}[\mathbb{1}_{x \in \phi(y)S}\mathbb{1}_{y \in \phi'(x)S}]] = \beta^2\mathbb{E}[h(x, y)].$$

Similarly for (a, b) with nonzero coordinates we find that

$$\begin{aligned} & \mathbb{E}_{\phi, \phi'}[\mathbb{E}_{x, y} f(x, y)f(x + a, y + b)f(x + 2a, y - 2b)f(x + 3a, y - b)] \\ &= \beta^8\mathbb{E}_{x, y}[h(x, y)h(x + a, y + b)h(x + 2a, y - 2b)h(x + 3a, y - b)] \\ &\leq (1 - \delta)\beta^8(\mathbb{E}h(x, y))^4. \end{aligned}$$

Finally assume that exactly one of a or b is zero. We handle the case when a is zero as the other is analogous. We find

$$\begin{aligned} & \mathbb{E}_{\phi, \phi'}[\mathbb{E}_{x, y} f(x, y)f(x, y + b)f(x, y - 2b)f(x, y - b)] \\ &\leq \mathbb{E}_{\phi, \phi', x, y}[\mathbb{1}_{x \in \phi(y)S}\mathbb{1}_{y \in \phi'(x)S}\mathbb{1}_{x \in \phi(y+b)S}\mathbb{1}_{y+b \in \phi'(x)S}\mathbb{1}_{x \in \phi(y-2b)S}\mathbb{1}_{y-2b \in \phi'(x)S}\mathbb{1}_{x \in \phi(y-b)S}\mathbb{1}_{y-b \in \phi'(x)S}] \\ &= \beta^4\mathbb{E}_{\phi', x, y}[\mathbb{1}_{y \in \phi'(x)S}\mathbb{1}_{y+b \in \phi'(x)S}\mathbb{1}_{y-2b \in \phi'(x)S}\mathbb{1}_{y-b \in \phi'(x)S}] \\ &\leq \beta^{8.15}. \end{aligned}$$

The last line follows since for every random map $\phi'(x)$, there is at most a density of $\beta^{4.15}$ of four term arithmetic progressions in $\phi'(x)S$. Therefore taking γ to be a sufficiently large multiple of $\log(1/\mathbb{E}[h(x, y)])$ (which is of constant order due to [Section 6.2](#)) guarantees that each of these expectations is at most $(1 - \delta)\beta^8(\mathbb{E}h(x, y))^4$. Furthermore, each of the above densities concentrates (with respect to the randomness of ϕ, ϕ') with high probability by Azuma–Hoeffding, so [Theorem 1.3](#) follows.

7. POPULAR DIFFERENCES FOR 3-POINT PATTERNS

We end by proving popularity for all admissible three-point patterns, namely, [Theorem 1.1](#). As is standard, we actually prove an analogous result over all compact abelian groups. We deduce the version over \mathbb{Z} from results (mod N) using a trick of Green [\[9\]](#).

These patterns can be handled in a direct Fourier-analytic manner. The proof here is closely modeled after the proof for 3-term arithmetic progressions. For a compact abelian group G with Haar measure μ , finite $S \subseteq \widehat{G}$, and $\delta > 0$, define the Bohr set

$$B(S, \delta) = \{x \in G : \max_{\xi \in S} (\|\xi x\|_{\mathbb{R}/\mathbb{Z}}) < \delta\}.$$

We recall the standard fact that $\mu(B(S, \delta)) = \Omega_{|S|, \delta}(1)$ (see [\[20, Lemma 4.20\]](#)). We will further write for a Bohr set B that μ_B is the uniform measure on it obtained by restricting the Haar measure appropriately.

Theorem 7.1. *For any $\alpha, \epsilon > 0$ there exists $C(\epsilon) > 0$ so that the following holds. Let G be a compact abelian group with Haar probability measure μ . Let M_1, M_2 be continuous automorphisms of G such that $M_1 - M_2$ is an automorphism. Then for any function $f: G \rightarrow [0, 1]$ with $\int_{g \in G} f(g) \mu(g) \geq \alpha$, there is a measure ν_D with $\|\nu_D\|_{L^\infty} \leq C$ such that*

$$\int_{x \in G, d \in G} f(x) f(x + M_1 d) f(x + M_2 d) \nu_D(d) \mu(x) \geq \alpha^3 - \epsilon. \quad (7.1)$$

Moreover, given a Bohr set $B_0 = B(S_0, \rho_0)$, one may take $\nu_D = \mu_B * \mu_B$, where $B = B(S, \rho)$ with $S \supseteq S_0$ and $\rho \leq \rho_0$. In this case, the absolute constant C will also depend on $|S_0|, \rho_0$.

From this general statement one can easily obtain popular difference results in their standard forms for finite abelian groups ([Theorem 7.3](#)) and for \mathbb{Z} ([Theorem 1.1](#)). The latter answers (as a special case) [\[1, Question 1.16\]](#) and is the combinatorial analog of [\[1, Theorem 1.10\]](#). As much of the proof is identical to that of the three-term arithmetic progression case in [\[19\]](#), we collect the necessary results in the following proposition. Essentially we are extracting a statement of the strong regularity lemma from [\[19\]](#). Closely related statements appear in [\[5, 10\]](#).

Proposition 7.2 ([\[19\]](#)). *Fix a compact abelian group G , parameters $\delta, \epsilon > 0$, and a set $S_0 \subseteq \widehat{G}$, as well as growth functions $\omega_1, \omega_2: \mathbb{R}^+ \rightarrow \mathbb{R}^+$. Given a function $f: G \rightarrow [0, 1]$, there are $\gamma_1, \gamma_2 > 0$ and a finite set T satisfying $S_0 \subseteq T \subseteq \widehat{G}$ with the following properties.*

1. $|T| = O_{\delta, \epsilon, |S_0|, \omega_1, \omega_2}(1)$.
2. $\gamma_1 \leq 1/\omega_1(|T| + \delta^{-1} + \epsilon^{-1})$ and $\gamma_2 \leq 1/\omega_2(\gamma_1^{-1})$, whereas γ_2 is bounded away from 0 independent of f .
3. We have the decomposition $f = f_1 + f_2 + f_3$, where
 - (a) f_1, f_2, f_3 are 1-bounded.
 - (b) f_1 is nonnegative, has mean $\int f_1 d\mu = \int f d\mu$, and obeys the bound

$$f_1(x + r) = f_1(x) + O(\epsilon)$$

whenever $x \in G$ and $r \in B(T, \gamma_1)$.

(c) $\|f_2\|_{L^2(G)} \leq \epsilon$.

(d) $\|\widehat{f_3}\|_{\ell^\infty(\widehat{G})} \leq \gamma_2$.

At this point, Tao [\[19\]](#) studies progressions with common difference in $B(T, \gamma_1)$. We require a small modification to handle more general matrix patterns. Define

$$B' = \{r \in G : M_1 r, M_2 r \in B(T, \gamma_1)\}$$

Observe that B' is the Bohr set $B(T', \gamma_1)$, where

$$T' = \{\xi \circ M_1 : \xi \in T\} \cup \{\xi \circ M_2 : \xi \in T\}$$

The fact that $\xi \circ M_1$ and $\xi \circ M_2$ are elements of \widehat{G} is immediate by the identification $\widehat{G} = \text{Hom}(G, \mathbb{R}/\mathbb{Z})$ and the fact that M_1 and M_2 are automorphisms. As an immediate consequence, we have for any $x \in G, d \in B'$ that

$$f_1(x + M_1 d), f_1(x + M_2 d) = f_1(x) + O(\epsilon). \quad (7.2)$$

With this additional Lipschitz condition in the directions of M_1 and M_2 within B' , the remainder of the proof follows in the standard manner. We will now evaluate

$$\int f(x) f(x + M_1 d) f(x + M_2 d) \mu_{B'} * \mu_{B'}(d) \mu(x).$$

Decompose each occurrence of f into $f_1 + f_2 + f_3$, so that the integral has 27 terms. We will show that all terms other than the f_1, f_1, f_1 term are bounded in magnitude by $O(\epsilon)$.

We first bound terms containing f_2 ; we consider the representative case where the second term is f_2 . In this case, take absolute values, and use $|f_a|, |f_b| \leq 1$. We have

$$\int f_a(x) f_2(x + M_1 d) f_b(x + M_2 d) \mu_{B'} * \mu_{B'}(d) \mu(x) \leq \int \mu_{B'} * \mu_{B'}(d) \int |f_2(x + M_1 d)| \mu(x).$$

The inner integral is bounded by $\|f_2\|_{L^1} \leq \|f_2\|_{L^2} \leq \epsilon$. Thus we are left with a bound of $O(\epsilon)$ as desired.

We next bound the terms containing f_3 . For the sake of simplicity we consider the case where the second term is f_3 . Standard Fourier analysis allows us to obtain that

$$\begin{aligned} & \int f_a(x) f_3(x + M_1 d) f_b(x + M_2 d) \mu_{B'} * \mu_{B'}(d) \mu(x) \\ &= \sum_{\xi_4} |\widehat{\mu}_{B'}|^2(\xi_4) \sum_{\xi_1 M_1 + \xi_3 (M_1 - M_2) = \xi_4} \widehat{f}_a(\xi_1) \widehat{f}_3(-\xi_1 - \xi_3) \widehat{f}_b(\xi_3) \end{aligned}$$

We take absolute values and use $|\widehat{f}_3| \leq \gamma_2$. When M_1 and $M_1 - M_2$ are invertible, Cauchy–Schwarz lets us bound the inner sum by

$$\gamma_2 \sum_{\xi_1} |\widehat{f}_a(M_1^{-1} \xi_1)|^2 \sum_{\xi_3} |\widehat{f}_b((M_1 - M_2)^{-1} \xi_3)|^2 \leq \gamma_2$$

where the final inequality follows by Plancherel. Plancherel also implies $\sum_{\xi_4} |\widehat{\mu}_{B'}|^2 = \|\mu_{B'}\|_{L^2}^2 \leq \|\mu_{B'}\|_{L^\infty} = O_{|T'|, \gamma_1^{-1}}(1)$. Recalling $|T'| \leq 2|T|$ and that γ_1 is small with respect to $|T|$, we find that

$$\sum_{\xi_4} |\widehat{\mu}_{B'}|^2 = O_{\gamma_1^{-1}}(1).$$

The total contribution is therefore bounded by $\gamma_2 \cdot O_{\gamma_1^{-1}}(1) \leq \epsilon$ as long as ω_1, ω_2 grow fast enough.

Finally we are left with the term

$$\int f_1(x) f_1(x + M_1 d) f_1(x + M_2 d) \mu_{B'} * \mu_{B'}(d) \mu(x).$$

Since $\mu_{B'} * \mu_{B'}(d)$ is supported on $B' + B'$, for any d in this support we have

$$f_1(x + M_i d) = f_1(x) + O(2\epsilon)$$

for $i \in \{1, 2\}$ by triangle inequality and two applications of (7.2) each. Rewriting, and recalling f_1 is $[0, 1]$ -valued, we have

$$\int f_1(x) f_1(x + M_1 d) f_1(x + M_2 d) \mu_{B'} * \mu_{B'}(d) \mu(x)$$

$$\begin{aligned}
&= \int f_1^3(x) \mu_{B'} * \mu_{B'}(d) \mu(x) + O(\epsilon) \\
&\geq \left(\int f_1(x) \mu(x) \right)^3 + O(\epsilon) && \text{(Hölder's inequality)} \\
&\geq \alpha^3 + O(\epsilon).
\end{aligned}$$

This completes the proof. \square

We now deduce a pair of corollaries of this result.

Theorem 7.3. *For any $\alpha, \epsilon > 0$ there exists N_0 so that the following holds. Let G be a finite abelian group of order $N \geq N_0$. Let M_1, M_2 be automorphisms of G so that $M_1 - M_2$ is an automorphism. Then for any $A \subseteq G$ with $|A| \geq \alpha N$, there is a popular difference $d \neq 0$ so that*

$$\#\{x \in G : x, x + M_1 d, x + M_2 d \in A\} \geq (\alpha^3 - \epsilon)N.$$

Proof. Choose $N_0 = C/\epsilon$, where $C = C(\epsilon)$ is the constant from [Theorem 7.1](#). Given this choice, the contribution from $d = 0$ in [\(7.1\)](#) is at most ϵN , and therefore by the pigeonhole principle, some nonzero d in the support of μ_D must have at least $\alpha^3 - O(\epsilon)$ density of this three-point pattern. This gives the desired result with an adjusted value of ϵ . \square

Finally we deduce the statement over the interval $[N]$ which was stated in the introduction.

Proof of Theorem 1.1. Choose $N_0 = C/\epsilon$, where $C = C(\epsilon, |S_0| = 2k, \delta_0 = \epsilon/(2k))$ is the constant from [Theorem 7.1](#). Embed $S \subseteq [p]^k \hookrightarrow (\mathbb{Z}/p\mathbb{Z})^k$ naturally, where p is a prime with $N < p < (1 + \epsilon/k)N$. Initialize $\delta_0 = \epsilon/(2k)$, and

$$S_0 = \left\{ x \mapsto \frac{(M_1 x)_i}{p} \right\}_{i \in [k]} \cup \left\{ x \mapsto \frac{(M_2 x)_i}{p} \right\}_{i \in [k]}.$$

Given this setup, we apply [Theorem 7.1](#). Our condition ensures that our matrices have nonzero determinant (and are therefore invertible) modulo p . The contribution from $d = 0$ is at most ϵ as before, and we can find at least $(\alpha^3 - O(\epsilon))N^k$ patterns with common difference $d \neq 0$ for some $d \in B(S, \delta) + B(S, \delta) \subseteq B(S, 2\delta)$, where $S \supseteq S_0$ and $\delta \leq \delta_0$. (Note that the density may decrease by $O(\epsilon)$ under this embedding.) Plugging in the elements of S_0 and using $\delta \leq \epsilon/(2k)$ gives us the constraints

$$(M_1 x)_i, (M_2 x)_i \in [-\epsilon p/k, \epsilon p/k], \text{ for all } i \in [k], x = (x_i)_{i=1}^k \in B(S, 2\delta).$$

Now we attempt to lift these $(\alpha^3 - \epsilon)N^k$ patterns into \mathbb{Z} by viewing d and each choice of x as integer vectors. Throw out at most ϵN^k choices of x for which $x_i \notin [(\epsilon/k)p, (1 - \epsilon/k)p]$ for some i . For the remainder, we see by triangle inequality in each coordinate that $x, x + M_1 d, x + M_2 d \in [p]^k$. By assumption, this triple must map to a three-point pattern under our embedding $[p]^k \hookrightarrow (\mathbb{Z}/p\mathbb{Z})^k$, so each of these points must have been an element of A originally, and the triple forms a three-point pattern over \mathbb{Z} . We have therefore found $(\alpha^3 - O(\epsilon))N^k$ of the necessary 3-point patterns in $[N]^k$ as desired, which completes the proof upon adjusting ϵ . \square

Remark. As usual, one can actually find $\Omega_{\alpha, \epsilon}(N^k)$ differences by using Markov's inequality and adjusting ϵ .

REFERENCES

- [1] E. Ackelsberg, V. Bergelson, and A. Best, *Multiple recurrence and large intersections for abelian group actions*, arXiv:2101.02811v1.
- [2] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332.

- [3] Vitaly Bergelson, Bernard Host, and Bryna Kra, *Multiple recurrence and nilsequences*, Invent. Math. **160** (2005), 261–303, with an appendix by Imre Ruzsa.
- [4] Aaron Berger, *Popular differences for corners in abelian groups*, Math. Proc. Cambridge Philos. Soc. (2020).
- [5] J. Bourgain, *A Szemerédi type theorem for sets of positive density in \mathbf{R}^k* , Israel J. Math. **54** (1986), 307–316.
- [6] Qing Chu, *Multiple recurrence for two commuting transformations*, Ergodic Theory Dynam. Systems **31** (2011), 771–792.
- [7] Sebastián Donoso and Wenbo Sun, *A pointwise cubic average for two commuting transformations*, Israel J. Math. **216** (2016), 657–678.
- [8] Jacob Fox, Ashwin Sah, Mehtaab Sawhney, David Stoner, and Yufei Zhao, *Triforce and corners*, Math. Proc. Cambridge Philos. Soc. **169** (2020), 209–223.
- [9] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), 340–376.
- [10] Ben Green, *Montreal lecture notes on quadratic fourier analysis*, arXiv:math/0604089.
- [11] Ben Green, *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27.
- [12] Ben Green and Terence Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinb. Math. Soc. (2) **51** (2008), 73–153.
- [13] Ben Green and Terence Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 261–334.
- [14] V. Kovač, *Popular differences for right isosceles triangles*, arXiv:2101.12714.
- [15] Matei Mandache, *A variant of the corners theorem*, arXiv:1804.03972.
- [16] Sean Prendiville, *Matrix progressions in multidimensional sets of integers*, Mathematika **61** (2015), 14–48.
- [17] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.
- [18] Ashwin Sah, Mehtaab Sawhney, and Yufei Zhao, *Patterns without a popular difference*, arXiv:2004.07722.
- [19] Terence Tao, *A proof of Roth’s theorem*, 2014, <https://terrytao.wordpress.com/2014/04/24/a-proof-of-roths-theorem/>.
- [20] Terence Tao and Van Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006.
- [21] P. Varnavides, *On certain sets of positive density*, J. London Math. Soc. **34** (1959), 358–360.

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA
 Email address: {bergera, asah, msawhney, jtidor}@mit.edu