Saad, A., & Wolf, J. (2017). Ramsey multiplicity of linear patterns in certain finite abelian groups. Quarterly Journal of Mathematics, 68(1), 125-140. DOI: 10.1093/qmath/haw011

RAMSEY MULTIPLICITY OF LINEAR PATTERNS IN CERTAIN FINITE ABELIAN GROUPS

A. SAAD AND J. WOLF

Abstract. In this article we explore an arithmetic analogue of a long-standing open problem in graph theory: what can be said about the number of monochromatic additive configurations in 2-colourings of finite abelian groups? We are able to answer several instances of this question using techniques from additive combinatorics and quadratic Fourier analysis. However, the main purpose of this paper is to advertise this sphere of problems and to put forward a number of concrete questions and conjectures.

1. Introduction

Using an elementary argument, Goodman [12] proved in the 1950s that a random 2-colouring of the edges of a large complete graph K_n contains asymptotically (in n) the minimum number of monochromatic triangles. Subsequently Erdős [11] conjectured that for all $s > 3$, it was indeed the random colouring that minimised the number of monochromatic copies of K_s amongst all 2-colourings of the edges of a large K_n. This conjecture was disproved by Thomason [35] in 1989, who exhibited an infinite family of 2-colourings of K_n which contained asymptotically strictly fewer than $2^{1-6}\binom{n}{4}$ monochromatic K_4s. A few years prior Burr and Rosta [2] had in fact optimistically generalised Erdős's conjecture from K_s to every fixed graph H. Clearly, since Erdős's conjecture is false, the Burr-Rosta conjecture does not hold for general graphs H. This was shown, independently of Thomason, by Sidorenko [29], who defined a sequence of edge-colourings of K_n that contained too few monochromatic copies of a graph H that consisted of a triangle with one additional edge attached to one of the vertices. However, the Burr-Rosta conjecture has been verified for several classes of graphs, which we shall call *common* in keeping with the usual terminology. Graphs known to be common include trees, cycles, even-spoked wheels, triangular edge- and vertex-trees [29, 30, 22, 36], to name just a few. Despite much research over the past two decades, the question of determining for which small graphs H the Burr-Rosta conjecture holds remains wide open. For a concise overview of the state-of-the-art, including a weaker version of the original Burr-Rosta conjecture that may hold for all graphs H, see [6, Section 2.6].

A particularly significant class of graphs that are known to be common stems from a closely related and arguably more acclaimed conjecture in graph theory, namely *Sidorenko's conjecture* [30], which has received a significant amount of attention lately. It states that the minimum number of copies of H in G essentially occurs when G is a random graph (and also appeared in a slightly weaker form in [31]). Clearly if Sidorenko's conjecture holds for a graph H, then H is common. It is easily seen that trees are Sidorenko, and as a result of recent work by Li and Szegedy [23], Szegedy [33] and Conlon et al. [5, 7], Sidorenko's conjecture is known for other large classes of bipartite graphs H. This includes bipartite graphs H which have at least

The authors are indebted to the School of Mathematics at the University of Bristol for sponsoring the first author's summer research project, of which this article is a result.

one vertex connected to all vertices in the other part [5]. The most general known examples are not straightforward to describe, and we refer the reader to [7].

In this article we shall explore the following arithmetic analogue of the above sphere of problems. Throughout we let Z be a finite abelian group. In some of our examples we shall take Z to be the cyclic group $\mathbb{Z}/N\mathbb{Z}$ for a sequence of primes $N \to \infty$, but even more frequently we shall consider the case where $Z = \mathbb{F}_p^n$ is a vector space of dimension n over a finite field of prime characteristic p. In the 'finite-field model' as it is generally understood (see [17, 38]), it is important that p be thought of as small and fixed (we shall see $p = 3$ and 5 most frequently in the sequel), while the dimension n is to be thought of as tending to ∞. Asymptotic results for this group are thus asymptotic in n.

Let L be a system of m homogeneous linear equations in d variables with integer coefficients, and let A be a subset of Z. We define the *arithmetic multiplicity* of L in A, denoted by $t_{Z,L}(A)$, to be the number of solutions to L in A divided by $|Z|^{\deg(L)}$, where $\deg(L)$ is the number of degrees of freedom of the linear system L (or the dimension of the solution space). In other words, $t_{Z,L}(A)$ denotes the probability that a randomly chosen solution to L in Z forms a solution to L in A. We also define the *arithmetic Ramsey multiplicity* of L with respect to the colouring induced by A and its complement A^C by $m_{Z,L}(A, A^C) := t_{Z,L}(A) + t_{Z,L}(A^C)$, where $A^C := Z \setminus A$.

For translation-invariant systems L, a generalised version of Szemerédi's theorem in finite abelian groups tells us that $m_{Z,L}(A, A^C) = \Omega(1)$ as $|Z| \to \infty$, so we shall almost exclusively focus our attention on translation-invariant systems.[1] Unlike the case of graphs, it seems to not be at all straightforward to show that the limit as $|Z| \to \infty$ of the minimum over $A \subseteq Z$ of $t_{Z,L}(A)$ (and hence $m_{Z,L}(A, A^C)$) exists, even for simple configurations L and specific families of finite abelian groups Z, see [8, 4]. We shall not pursue this matter here.

In analogy with the graph-theoretic problem, we shall call a linear patterns defined by L *Sidorenko* in Z if it occurs with frequency at least α^d in every subset of Z of density α, i.e. if $t_{Z,L}(A) \geq \alpha^d$ for all $A \subseteq Z$ of density α. We shall call a configuration L *common* in Z if it occurs asymptotically with frequency at least $2(1/2)^d$ in any 2-colouring of the elements of the group Z, that is, if for any $A \subseteq Z$, $m_{Z,L}(A, A^C) \geq 2(1/2)^d + o(1)$ (as $|Z| \to \infty$). As before, it is easy to see that if a linear system L is Sidorenko, then it is common.

Let us discuss a first (and important) example to clarify these definitions.

Example 1.1 (Additive quadruples)**.** Let Z be any finite abelian group. Let L be an additive quadruple, that is, a solution to the single equation $x + y = z + w$ with 3 degrees of freedom, which we shall denote by AQ. We claim that AQ is Sidorenko in Z and hence common. To see this, observe that the number of solutions to $x + y = z + w$ in a subset $A \subseteq Z$ can be written as

$$\sum_{x,a,b \in Z} 1_A(x)1_A(x+a)1_A(x+b)1_A(x+a+b) = \sum_{a \in Z} \left(\sum_{x \in Z} 1_A(x)1_A(x+a) \right)^2,$$

where 1_A denotes the indicator function of A, which takes the value $1_A(x) = 1$ if $x \in A$, and 0 otherwise. A simple application of the Cauchy-Schwarz inequality shows that the number of solutions to L in A is at least $|A|^4/|Z|$. In other words, $t_{Z,AQ}(A) \geq \alpha^4$, and hence $m_{Z,AQ} \geq 2 \cdot (1/2)^4$, as required.

[1] For some non-translation invariant linear systems such as Schur triples (see Example 2.2 below), a non-trivial lower bound on $m_{Z,L}(A, A^C)$ in certain families of finite abelian groups follows from a result of Deuber [10].

This example allows us to explain in more detail why we call the set-up above the arithmetic analogue of the graph-theoretic question formulated in the introductory paragraphs of the present paper. Additive quadruples in subsets of finite abelian groups are well known to be in direct correspondence with 4-cycles in graphs. Specifically, given a subset $A \subseteq Z$, which we shall assume to be symmetric in the sense that $A = -A := \{-a : a \in A\}$, we can construct the so-called *Cayley graph* $\Gamma = \Gamma(Z; A)$ as follows. Let $V(\Gamma) := Z$, and let uv be an edge in Γ if and only if $u - v \in A$. It is easily verified that the number of 4-cycles in Γ corresponds precisely to the number of additive quadruples in A.

More generally, such a correspondence can be set up for other linear patterns but the limitations are twofold. First, not every graph is (isomorphic to) a Cayley graph generated by a symmetric subset of Z (in fact, such Cayley graphs arise with vanishingly small frequently). Secondly, not all linear configurations can be made to arise from graphs. For example, it is well known through attempts of proving Szemerédi's theorem for arithmetic progressions of length 4 that the latter linear configuration cannot be represented in the graph-theoretic universe alone (and in fact, a 3-uniform hypergraph is needed to represent its solutions).

In the next section we ease into the problem with some relatively straightforward examples of additive configurations, including hypercubes and 3-term arithmetic progressions. In Section 3 we recall (and extend) the more complicated example of 4-term arithmetic progressions, studied by the second author in [37]. We examine to what extent adding free variables makes a given linear pattern uncommon (Section 4), and subsequently present some evidence to support a new conjecture concerning the (perhaps surprisingly) difficult case of linear patterns defined by one equation in an even number of variables (Section 5).

What we hope will spark the reader's interest when presented with this broad range of examples is the analysis of the underlying reason for which a linear pattern turns out to be common or uncommon. Unlike the case of graphs, where a multitude of cases has been studied but any attempt at classification appears extremely difficult, there is a strong structural theory for the arithmetic instance of the problem which allows us to be more systematic in our approach. However, we shall see that this theory turns up a range of reasons for the (un)commonality of a given linear pattern, which we believe shows that the problem is a difficult one even in the arithmetic setting. We present a summary of the emerging picture in Section 6.

2. Some straightforward patterns

We start by discussing a simple generalisation of Example 1.1. For a definition of the Fourier transform, the uniformity norms and other standard notation, we refer the reader to [18].

Example 2.1 (Additive k-tuples)**.** Let Z be any finite abelian group. Let $k \geq 2$ be an integer and consider solutions to the equation $x_1 + x_2 + \cdots + x_k = x_{k+1} + x_{k+2} + \cdots + x_{2k}$, denoted by AQ_k. This linear pattern is easily seen to be Sidorenko (and hence common). The Cauchy-Schwarz argument from Example 1.1 generalises, but we may also use the Fourier transform to see that

$$t_{Z,AQ_k}(A) = \mathbb{E}_x 1_A * 1_A * \cdots * 1_A(x)^2 = \sum_\gamma |\widehat{1_A}(\gamma)|^{2k},$$

where the convolution is k-fold, which by positivity of the summand is bounded below by $|\widehat{1_A}(0)|^{2k} = \alpha^{2k}$ as claimed.

What can we say about other linear patterns defined by one equation? We shall examine the first non-trivial case, namely that of an equation in three variables.

Example 2.2 (Schur triples)**.** Let Z be any finite abelian group. A *Schur triple*, denoted by ST, is a solution to the equation $x+y=z$. It is well known that there are dense sets containing no Schur triples at all, implying that this pattern is not Sidorenko. For example, in $\mathbb{F}_p^n$ we may take any non-trivial coset of a subspace of codimension 1, which has density $1/p$ in the group. In $\mathbb{Z}/N\mathbb{Z}$ it is easy to see that the "middle-thirds" set $\{x \in \mathbb{Z}/N\mathbb{Z} : N/3 \leq x < 2N/3\}$ contains no solutions to $x+y=z$. A comprehensive analysis of the quantity $t_{Z,ST}(A)$ for various groups Z has recently been undertaken by Samotij and Sudakov in [27].

On the other hand, it is elementary to see that Schur triples are common, see for example Theorem 1 in [9] and Corollary 3.1 in [3]. One way is to simply expand

$$m_{Z,ST}(A, A^C) = \mathbb{E}_{x,y} 1_A(x)1_A(y)1_A(x+y) + \mathbb{E}_{x,y} 1_{A^C}(x)1_{A^C}(y)1_{A^C}(x+y),$$

replace 1_{A^C} by $1-1_A$, and evaluate the corresponding sums. Note that the terms containing a triple product of indicator functions disappear, due to a sign change. Another way of reaching the same conclusion is to rewrite $t_{Z,ST}(A)$ in terms of the Fourier transform of 1_A and to use the fact that $\widehat{1_A}(\gamma) = -\widehat{1_{A^C}}(\gamma)$ for $\gamma \neq 0$.

It is remarkable that we not only obtain a lower bound on the number of monochromatic Schur triples, but in fact an exact formula for $m_{Z,ST}(A, A^C)$ whose value only depends on the density of the colour classes. An identical phenomenon occurs in the following example.

Example 2.3 (3-term arithmetic progressions)**.** Let $Z = \mathbb{Z}/N\mathbb{Z}$ or $Z = \mathbb{F}_p^n$ with $p > 2$. A *3-term arithmetic progression*, or $3-AP$, is defined by the equation $x+y=2z$. Giving a lower bound for $t_{3-AP}(A)$ in terms of the density of A corresponds to the infamously difficult problem of obtaining upper bounds in Roth's theorem (see for example Chapter 10 of [34]). In particular, it is known through recent work of Bloom [1] (building on prior work of Sanders [28]) that in $\mathbb{Z}/N\mathbb{Z}$,

$$t_{\mathbb{Z}/N\mathbb{Z},3-AP}(A) \geq \alpha^{O(\alpha^{-1}\log^3(\alpha^{-1}))},$$

a far cry from the expected α^3 in a random set. On the other hand, Green and Sisask [19] exhibited a subset A of $\mathbb{Z}/N\mathbb{Z}$ of density $1/2$ for which $t_{\mathbb{Z}/N\mathbb{Z},3-AP}(A) = 5/48 < 1/8$, using ideas we shall return to in Section 4. It follows that 3-term arithmetic progressions are not Sidorenko in $\mathbb{Z}/N\mathbb{Z}$. However, both proofs in Example 2.2 adapt without difficulty to show that 3-APs are common.

It was observed, by Cameron, Cilleruelo and Serra [3], for example, that the same argument can be extended to the analysis of any linear pattern defined by one equation in an odd number of variables. We leave the proof as an easy exercise to the reader.

Example 2.4 (One equation in an odd number of variables)**.** Let Z be any finite abelian group, and let OE be a linear pattern defined by one equation in an odd number of variables, i.e. $\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k = 0$ for some odd integer k and integers $\beta_1, \beta_2, \ldots, \beta_k$. Let A be a subset of Z of density α. Then

$$m_{Z,OE}(A, A^C) = \alpha^k + (1-\alpha)^k \geq \left(\frac{1}{2}\right)^{k-1},$$

from which it follows that OE is common.

Both the Fourier- and the Cauchy-Schwarz approach clearly fail, as a result of lack of cancellation, for even values of k. We shall return to the unexpectedly difficult case of one equation in an even number of variables in Section 5. We conclude this section by generalising

Example 1.1 in another direction, which belongs to a different "complexity class" from the above and which will naturally lead us on to Section 3.

Example 2.5 (Hypercubes)**.** Let Z be any finite abelian group, and let $d \geq 2$ be an integer. We call the linear configuration C_d given by the 2^d linear forms $(x_0 + \sum_{i=1}^{d} \epsilon_i x_i)_{\epsilon \in \{0,1\}^d}$ a *hypercube of dimension* d. When $d = 2$ this definition again reduces to that of an additive quadruple. Just like an additive quadruple, the hypercube of dimension d is easily seen to be Sidorenko (and hence common) for $d > 2$.

3. Square dependent and independent configurations

The reader may care to verify that the preceding example (Example 2.5) does not have a Fourier-based proof of commonality as for $d > 2$ there is no useful expression for the number of hypercubes in terms of the Fourier transform. The latter fact is also true of 4-term arithmetic progressions, and it is for this reason that Gowers introduced the uniformity norms (see, for example, Definition 2.12 in [38]).

Arithmetic progressions of length 4, denoted by $4 - AP$ and defined by the equations $x+y = 2z$ and $y+z = 2w$, are clearly not Sidorenko, and finding lower bounds on $t_{Z,4-AP}(A)$ in terms of the density of A corresponds to finding good bounds in Szemerédi's theorem in the group Z.

Motivated by Thomason's proof [35] that the graph K_4 is uncommon, the second author showed in [37] that 4-term arithmetic progressions are uncommon in $\mathbb{Z}/N\mathbb{Z}$. Specifically, it was shown that there exists a set $A \subseteq \mathbb{Z}/N\mathbb{Z}$ for which

$$m_{\mathbb{Z}/N\mathbb{Z},4-AP}(A, A^C) < \left(1 - \frac{1}{259200}\right) \times \left(\frac{1}{2}\right)^3 \approx 0.12499952,$$

where $(1/2)^3$ is of course the proportion of 4-APs expected in a random set. In subsequent work Lu and Peng [24] improved the right-hand side to the much more reasonable $68/75 \times (1/2)^3 \approx 0.113333$.[2]

Given that 4-term arithmetic progressions are uncommon in $\mathbb{Z}/N\mathbb{Z}$, it is natural to ask the following question.

Question 3.1. For a given finite abelian group Z, what is $\min_{A \subseteq Z} m_{Z,4-AP}(A, A^C)$?

In [37] it was shown that

$$\min_{A \subseteq \mathbb{Z}/N\mathbb{Z}} m_{\mathbb{Z}/N\mathbb{Z},4-AP}(A, A^C) \geq \left(\frac{1}{2}\right)^4,$$

and the right-hand side was improved by Lu and Peng [24] to 7/96. In graph theory the corresponding minimisation problem has been studied extensively. For the best known upper and lower bounds on the minimum number of monochromatic copies of K_4 in any 2-colouring of the edges of K_n see [36] and [32], respectively.

We use this section to add three further observations to the existing body of work on 4-term arithmetic progressions. First, we shall give a finite-field version of the construction in [37], which has the benefit of being significantly easier to understand, and requiring no strenuous computation whatsoever. Secondly, we examine what can be said about the structure of those colourings that show 4-APs to be uncommon in this setting. Thirdly, we analyse to what

[2] Here and elsewhere in the paper, numerical results are given to eight significant figures.

extent these methods can be used to obtain results about configurations containing 4-term progressions.

The first colouring containing fewer than the expected number of monochromatic 4-term progressions in [37] was based on an unpublished construction from quadratic Fourier analysis due to Gowers [14], who had constructed a subset of $\mathbb{Z}/N\mathbb{Z}$ of density $1/2$ which was uniform in the sense that the non-trivial Fourier coefficients of its indicator function were small, but which contained significantly fewer than the expected number of 4-APs.[3] In $\mathbb{F}_5^n$, many of the technical details simplify and we give them in full here.

Example 3.1 (4-term arithmetic progressions)**.** There exists a set $A \subseteq \mathbb{F}_5^n$ such that

$$m_{\mathbb{F}_5^n, 4-AP}(A, A^C) \leq \frac{1}{8} - \frac{7}{2^{10}5^2} \approx 0.12472656.$$

We shall break up the construction of A into a number of claims.

Claim 3.2. Let $f : \mathbb{F}_5 \to \{-1, 1\}$ be defined by $f(x) = -1$ if $x = 0$ and $f(x) = 1$ otherwise. Then

$$\mathbb{E}_{x,d \in \mathbb{F}_5} f(x)f(x+d)f(x+2d)f(x+3d) = -\frac{7}{25}.$$

Moreover, if we let $V := \mathbb{F}_5^{n-1}$ so that $\mathbb{F}_5^n = V \oplus V^\perp$, and define $F : \mathbb{F}_5^n \to \{-1, 1\}$ by setting $F(x) := f(y)$ when $x \in V + y$ for $y \in V^\perp$, then

$$\mathbb{E}_{x,d \in \mathbb{F}_5^n} F(x)F(x+d)F(x+2d)F(x+3d) = -\frac{7}{25}.$$

Proof. There are 25 4-term progressions in $\mathbb{F}_5$, including 5 trivial ones which contribute $5/25$ to the expectation in f. Each non-trivial progression is counted 4 times, and all besides $1, 2, 3, 4$ contribute $-4/25$ to the expectation, amounting to a total contribution of $-16/25$. The progression $1, 2, 3, 4$ contributes $4/25$, for a total of $(5 - 16 + 4)/25 = -7/25$. The second part of the claim is immediate by splitting $\mathbb{E}_{x \in \mathbb{F}_5^n} = \mathbb{E}_{y \in V^\perp} \mathbb{E}_{v \in V}$, and similarly for d. □

Claim 3.3. Let $G : \mathbb{F}_5^n \to [-4, 4]$ be defined by $G(x) := F(x)(\omega^{x \cdot x} + \omega^{-3x \cdot x} + \omega^{3x \cdot x} + \omega^{-x \cdot x})$. Then $|\widehat{G}(t)| = o(1)$ for all $t \in \mathbb{F}_5^n$. Moreover,

$$|\mathbb{E}_{x,d \in \mathbb{F}_5^n} G(x)G(x+d)G(x+2d)G(x+3d) - 4\mathbb{E}_{x,d \in \mathbb{F}_5^n} F(x)F(x+d)F(x+2d)F(x+3d)| = o(1).$$

Proof. It is easily computed that for any affine subspace $W = w + W_0$, where $W_0 \leqslant \mathbb{F}_5^n$, $|\widehat{1_W}(t)| = 1/|W_0^\perp|$ if $t \in W_0^\perp$, and 0 otherwise. Recall also that from standard Gauss sum estimates, $|\widehat{\omega^q}(t)| = o(1)$ for any $t \in \mathbb{F}_5^n$ and any quadratic form q of rank tending to infinity with n. Thus for any affine subspace W, the function $G'(x) = 1_W(x)\omega^{q(x)}$ has Fourier transform of size

$$|\widehat{G'}(t)| = |\widehat{1_W} * \widehat{\omega^q}(t)| = |\sum_s \widehat{1_W}(t-s)\widehat{\omega^q}(s)| \leq \sup_s |\widehat{\omega^q}(s)| \sum_s |\widehat{1_W}(s)| = \sup_s |\widehat{\omega^q}(s)| = o(1).$$

In order to obtain the result for G, it remains to observe that F is a plus/minus one combination of indicator functions of 5 affine subspaces.

To see why the second part of the claim is true we have to work a tiny bit harder. We start by expanding the product

$$\mathbb{E}_{x,d \in \mathbb{F}_5^n} G(x)G(x+d)G(x+2d)G(x+3d)$$

[3]It is significantly easier to obtain an example of a uniform set containing significantly *more* than the expected number of 4-APs, see for example Section 2.3 of [38].

into $4^4 = 256$ terms, each of which is of the form

$$\mathbb{E}_{x,d\in\mathbb{F}_5^n} F(x)F(x+d)F(x+2d)F(x+3d)\omega^{ax\cdot x+b(x+d)\cdot(x+d)+j(x+2d)\cdot(x+2d)+k(x+3d)\cdot(x+3d)},$$

where each of a, b, j, k takes one of the values $+1$,-3,3 or -1. It can easily be checked that the only assignments (a, b, j, k) that leave the exponent equal to zero are $(1, -3, 3, 1)$, $(-1, 3, -3, 1)$, $(3, 1, -1, -3)$ and $(-3, -1, 1, 3)$. Together these four contributions give rise to a term of the form

$$4\mathbb{E}_{x,d\in\mathbb{F}_5^n} F(x)F(x+d)F(x+2d)F(x+3d).$$

All remaining terms, which involve a product of copies of F with a quadratic exponential or a non-trivial bilinear phase in x and d, are negligible by a variant of the argument made for G' at the start of the proof. □

Claim 3.4. Let $h : \mathbb{F}_5^n \to [0, 1]$ be defined by $h(x) = \frac{1}{8}(G(x) + 4)$. Then $\mathbb{E}_x h(x) = \frac{1}{2} + o(1)$ and $|\widehat{h}(t)| = o(1)$ for all $t \neq 0$. Moreover,

$$\mathbb{E}_{x,d\in\mathbb{F}_5^n} h(x)h(x+d)h(x+2d)h(x+3d) \leq \frac{1}{16} - \frac{7}{2^{10}5^2} + o(1).$$

Proof. The claims concerning the average and the Fourier coefficients of h are easy to verify. To see the final inequality, expand the product

$$\mathbb{E}_{x,d\in\mathbb{F}_5^n} h(x)h(x+d)h(x+2d)h(x+3d)$$

into 16 terms. The term arising from having chosen 4 from each bracket gives the main contribution of $(1/2)^4 = 1/16$. Combining Claims 3.2 and 3.3, we see that the term

$$\frac{1}{2^{12}}\mathbb{E}_{x,d\in\mathbb{F}_5^n} G(x)G(x+d)G(x+2d)G(x+3d)$$

arising from having chosen G from each bracket contributes

$$\leq -\frac{4\times 7}{25}\frac{1}{2^{12}} + o(1).$$

All other terms are negligible as they define configurations in G consisting of at most 3-terms, all of which are controlled by the Fourier coefficients of G. □

The function h can now be converted into a subset of $\mathbb{F}_5^n$ by a standard probabilistic argument, namely by letting $x \in \mathbb{F}_5^n$ lie in the desired set with probability $h(x)$. We leave the details to the reader.

This concludes the proof of the example. In contrast to this analytic way of proceeding, Lu and Peng [24] used a brute-force computational approach in $\mathbb{Z}/N\mathbb{Z}$ which turned out to be quantitatively superior. However, the following simple lemma shows that, at least in a weak sense, any colouring which contains fewer than the expected number of monochromatic 4-term progressions (which we shall refer to as *bad* in the sequel) must arise from a quadratically structured example as in [37]. The proof uses a deep result, namely the inverse theorem for the U^3 norm (see Theorem 2.3 in [20], which is based on prior work of Gowers [13]), but is otherwise routine and therefore omitted.

Lemma 3.5 (Structure of bad colourings for 4-APs)**.** Let $0 < \delta < 1/8$ and suppose that $A \subseteq \mathbb{F}_5^n$ is such that

$$m_{\mathbb{F}_5^n,4-AP}(A, A^C) < \left(\frac{1}{2}\right)^3 - \delta.$$

Then there exists a quadratic form q on $\mathbb{F}_5^n$ such that $|\mathbb{E}_x 1_A(x)\omega^{q(x)}| \geq c(\delta)$, where c is a function that tends to zero as δ tends to zero.

The dependence of $c(\delta)$ on δ is relatively weak, meaning that Lemma 3.5 is primarily of qualitative interest.[4] Having said this, we believe it provides the first insight into the structure of bad colourings. In particular, to our knowledge no such result is known for the original graph-theoretic problem (although interestingly, Thomason's first construction in [35] was also based on a quadratic form).

Finally, we note that any example of the kind constructed above also gives rise to a bad colouring for 5-APs. Similarly to [37], we can write

$$m_{\mathbb{F}_5^n,5-AP}(A, A^C) = \mathbb{E}_{x,d} \prod_{j=0}^{4} 1_A(x + jd) + \mathbb{E}_{x,d} \prod_{j=0}^{4} 1_{A^C}(x + jd)$$

as

$$-\sum_{i=0}^{4} t_{\mathbb{F}_5^n,5-AP(i)}(A) + \sum_{\{i,k\}\in\{0,1,2,3,4\}^{(2)}} t_{\mathbb{F}_5^n,5-AP(i,k)}(A) - \binom{5}{2}\alpha^2 + 5\alpha - 1,$$

where we have written $t_{\mathbb{F}_5^n,5-AP(i)}(A)$ for the expression $\mathbb{E}_{x,d}\prod_{j=0}^4 f_j(x+jd)$ where $f_j = 1$ when $j = i$, and $f_j = 1_A$ otherwise, and $t_{\mathbb{F}_5^n,5-AP(i,k)}(A)$ for the expression $\mathbb{E}_{x,d}\prod_{j=0}^4 f_j(x + jd)$ where $f_j = 1$ when $j = i$ or $j = k$, and $f_j = 1_A$ otherwise. Notice that since we are working over $\mathbb{F}_5$, the configurations defined by $5 - AP(i)$ are still 4-term progressions, while the configurations defined by $5 - AP(i, k)$ are all 3-term progressions.[5]

Writing further $d_4 = t_{\mathbb{F}_5^n,4-AP}(A) - \alpha^4$ and $d_3 = t_{\mathbb{F}_5^n,3-AP}(A) - \alpha^3$ for the deviation from the expected number, we find after some rearranging that

$$m_{\mathbb{F}_5^n,5-AP}(A, A^C) = \alpha^5 + (1 - \alpha)^5 - 5d_4 + 10d_3.$$

It follows as in [37] that a set A which is uniform (implying that $d_3 = o(1)$) and which contains fewer than the expected number of 4-APs (meaning $d_4 \leq -c$ for some positive constant c) gives rise to a colouring that contains fewer than the expected number of 5-term arithmetic progressions. Again, cancellation has come to the rescue.

While perhaps not entirely unexpected, the fact that the bad colouring for 4-APs is also bad for 5-APs bears emphasising as we had no a priori information about the number of 5-APs in A. It also illustrates the power of the quadratic Fourier analysis approach, as the purely computational one would have required us to start our calculations again from zero. The following question naturally arises.

Question 3.6. Is it true that every linear configuration containing a 4-AP is uncommon?

Even though we expect the answer to be positive, it does not seem to follow immediately from the above observations. In graph theory the analogous result is known [22]: any graph containing a K_4 is uncommon.

We may dig deeper yet. To start with we recall a definition from [15].

Definition 3.7 (Square independence)**.** A linear configuration defined by m linear forms $L_1, L_2, \ldots, L_m$ in d variables with integer coefficients is said to be *square independent over* $\mathbb{F}_p$ if the quadratic forms $L_i^T L_i$, $i = 1, 2, \ldots, m$, are linearly independent over $\mathbb{F}_p$.

[4]A similar argument can be made in the case of $\mathbb{Z}/N\mathbb{Z}$, but since the statement of the U^3 inverse theorem is less clean in that setting we omit the details.

[5]This is not true in $\mathbb{Z}/N\mathbb{Z}$, but the argument that follows can be adapted.

In [15, 16] Gowers and the second author proved that a linear configuration defined by linear forms $L_1, L_2 \ldots, L_m$ is square independent if and only if it is controlled by Fourier analysis in the sense that for any $\epsilon > 0$, if $f : \mathbb{F}_p^n \to \mathbb{C}$ is any function satisfying $\|f\|_\infty \leq 1$ and $\|f\|_{U^2} \leq \epsilon$, then

$$|\mathbb{E}_{x_1,x_2,\ldots,x_d} \prod_{i=1}^{m} f(L_i(x_1, x_2, \ldots, x_m))| < c(\epsilon)$$

for some function $c(\epsilon)$ which tends to zero as ϵ tends to zero.

This suggests that *if* a square-independent configuration is uncommon, it suffices to look for bad colourings that exhibit a linear bias (see the discussion following Example 4.1 below). To illustrate this phenomenon we give just one of numerous possible examples here.

Example 3.2 (A square-independent configuration)**.** Let $Z = \mathbb{F}_5^n$. Consider the configuration SI given by the system of equations

$$\begin{array}{rcl} x + y & = & 2u \\ x + y + z & = & 3v \end{array},$$

which can easily be verified to be square independent over $\mathbb{F}_5$. Let $V \leqslant \mathbb{F}_5^n$ be a subspace of codimension 1. Let A be the union of the $+1$ and -1 cosets of V in $\mathbb{F}_5^n$, together with half the elements of V chosen at random. Then with high probability the density of A is 1/2, and it is not difficult to compute that

$$m_{\mathbb{F}_5^n, SI}(A, A^C) = 0.0525 < 2 \cdot \left(\frac{1}{2}\right)^5.$$

This is in stark contrast with the situation for square-dependent patterns (such as 4-term arithmetic progressions) which require the construction of bad colourings with genuinely quadratic structure.

4. Free variables skew densities

We continue our exploratory journey through the arithmetic forest, returning to a much simpler configuration. Again, the authors were inspired by an example in graph theory when considering the slightly odd-looking set-up below, in which we have an unconstrained variable.

Example 4.1 (3-AP with a free variable)**.** Let $Z = \mathbb{F}_3^n$, and let TP be the set of solutions (x, y, z, w) to the equation $x + y = 2z$. Then there exists $A \subseteq \mathbb{F}_3^n$ such that

$$m_{\mathbb{F}_3^n, TP}(A, A^C) \approx 0.12463884 < 2\left(\frac{1}{2}\right)^4.$$

In other words, TP is uncommon in $\mathbb{F}_3^n$.

In order to construct A, take eight linearly independent vectors u, v_1, v_2, w_1, w_2, y_1, y_2, y_3, and let U, V, W and Y be subspaces of codimension 1, 2, 2 and 3, respectively, whose orthogonal complements are spanned by the correspondingly labelled vectors. Let A be the union of U, V, W and Y. It is not too difficult to compute, using inclusion-exclusion, that

$$|A| \approx 0.49276024|\mathbb{F}_3^n|$$

and with considerably more effort that

$$m_{\mathbb{F}_3^n, TP}(A, A^C) \approx 0.12463884,$$

as desired. The latter calculation can be carried out in any number of ways: by brute-force computation; using the Fourier-coefficients of the function 1_A, which are reasonably straightforward to write down; or carefully counting the number of 3-APs by hand, using inclusion-exclusion. We leave the details to the energetic reader.

Much more interesting than the construction itself is the sequence of observations that led to it, which we shall briefly record here. It follows from the definitions that for any set $A \subseteq \mathbb{F}_3^n$

$$m_{\mathbb{F}_3^n,TP}(A,A^C) = \alpha t_{\mathbb{F}_3^n,3-AP}(A) + (1-\alpha) t_{\mathbb{F}_3^n,3-AP}(A^C),$$

which, since $t_{\mathbb{F}_3^n,3-AP}(A) + t_{\mathbb{F}_3^n,3-AP}(A^C) = \alpha^3 + (1-\alpha)^3$, means that

$$m_{\mathbb{F}_3^n,TP}(A,A^C) = \alpha^4 + (1-\alpha)^4 + (2\alpha-1)(t_{\mathbb{F}_3^n,3-AP}(A) - \alpha^3). \tag{4.1}$$

This immediately tells us that we are guaranteed to get the count of configurations expected in the random case whenever the set A is of density $1/2$. Therefore, if we wish to show that the configuration TP is uncommon we need to look for sets whose density differs from $1/2$ very slightly, as any large deviation will ensure that the term $\alpha^4 + (1-\alpha)^4$ takes control, undermining any hope of obtaining a lower than expected count. In addition, note that $m_{\mathbb{F}_3^n,TP}(A,A^C)$ depends on the deviation of the 3-AP count in A from the expected value. We are therefore led to searching for a set whose density is slightly below $1/2$ but which contains many more than the expected number of 3-APs. An obvious candidate for the latter is a subspace, and in Example 4.1 the codimensions were simply chosen so as to bring the density of the union as close to $1/2$ as possible.

These remarks also immediately lead us to an analogue of Example 4.1 in $\mathbb{Z}/N\mathbb{Z}$. Indeed, Green and Sisask [19] constructed for any $1/3 < \alpha < 2/3$ a set $A \subseteq \mathbb{Z}/N\mathbb{Z}$ (consisting of a union of arithmetic progressions, the $\mathbb{Z}/N\mathbb{Z}$-analogue of subspaces) of density α with the property that

$$t_{\mathbb{Z}/N\mathbb{Z},3-AP}(A) \leq \frac{2 - 12\alpha + 21\alpha^2}{12}.$$

An optimisation yields set $A \subseteq \mathbb{Z}/N\mathbb{Z}$ of density $\alpha \approx 0.50693243$ such that

$$m_{\mathbb{Z}/N\mathbb{Z},TP}(A,A^C) \approx 0.12485549 < 2\left(\frac{1}{2}\right)^4.$$

More can be said on the basis of Equation (4.1). Notice that if $m_{\mathbb{F}_3^n,TP}(A,A^C) < 2(1/2)^4 - \delta$ for some constant $0 < \delta < 1/8$, then the 3-AP count of A must deviate significantly from its expectation. Using standard Fourier-analytic arguments from the proof of Meshulam's theorem [25], for example, it can be shown that in this case there exists an element $t \in \mathbb{F}_3^n$, $t \neq 0$, such that $|\mathbb{E}_x 1_A(x)\omega^{x\cdot t}| \geq \delta/\alpha(1-2\alpha)$. From this it follows, again via a routine argument, that the set A must have linear structure in the sense that it is strongly biased towards a very large (potentially affine) subspace. Indeed, it is precisely a set with this property which gave us the construction in Example 4.1 in the first place. Our conclusions here are again merely of a qualitative nature, but should be compared with the structural information in Lemma 3.5, which stated that a bad colouring for 4-APs must be quadratically (and not just linearly) structured. This means that configurations can be uncommon for at least two genuinely distinct reasons, a point which we shall return to in Section 6.

It is of course possible to add unconstrained variables to other configurations. We record a natural question posed to us by Noga Alon.

Question 4.1. Is it true that adding sufficiently many free variables makes any linear configuration uncommon?

As before, a suitable version of such a statement is true for graphs (see Theorem 4 in [22]).

We already noted in the introduction that the Cayley graph construction can be used to transfer a bad colouring from the arithmetic to the graph setting. For this to work, however, the bad colouring found in the arithmetic case must be symmetric (that is, if x is coloured red then so is $-x$), and the coefficients in the linear system must all equal plus or minus 1. Below we give an example of one situation in which such a transfer can be carried out successfully.

Example 4.2 (Triangle with a pendant edge)**.** The construction in Example 4.1 gives a new proof that the graph T' consisting of a triangle with a pendant edge is uncommon, a fact originally proved by Sidorenko [29].

Note that since the set $A \subseteq \mathbb{F}_3^n$ in Example 4.1 is defined as a union of subspaces, it is symmetric. It follows that we can define a Cayley graph $\Gamma = \Gamma(Z; A)$ on vertex set $V(\Gamma) = \mathbb{F}_3^n$ with uv being an edge if and only if $u - v \in A$. Any quadruple $(x, y, z, w) \in A^4$ satisfying $x + y + z = 0$ thus corresponds to a triangle uv, vs, us with a pendant edge ut in Γ, which can be seen by setting $x = u - v$, $y = v - s$, $z = s - u$ and $w = u - t$.

In fact, it is interesting to compare the structure of our colouring with that of Sidorenko's (who obtained a better constant).

In the other direction, it is not difficult to convince oneself that any proof showing that a given graph H is common can be adapted to show that an associated linear configuration is common. The reason is that essentially all known such proofs are based on the Cauchy-Schwarz inequality. A relatively large number of linear systems can be shown to be common in this way. It is impossible to give an exhaustive list, but systems of equations associated with triangular edge- or vertex-trees, or those associated with square wheels and other regular grid structures (see Chapter II of [21]), fall in this category.

5. One equation in an even number of variables

We return to the case of configurations defined by a single equation. In Section 2 we established that for an odd number of variables, the configuration is not only always common, but that in fact an exact formula for the number of monochromatic solutions can be given in terms of the density of the colour classes. An identical argument exploiting cancellation, written down by Cameron, Cilleruelo and Serra in [3], shows that in the case of an even number of variables, the *difference* between the number of red solutions and the number of blue solutions is given by an exact formula in terms of the respective colour densities. When the colours are exactly balanced, one concludes that the number of red and the number of blue configurations is in fact the same. However, the formula for the difference rather than the sum of monochromatic solutions does not address the question of whether such a configuration is common, and indeed the case of one equation in an even number of variables remains one of the most mysterious.

For simplicity we shall initially restrict our attention to translation-invariant equations over $\mathbb{F}_5$ in 4 variables. It does not take long to check that there are only four genuinely distinct configurations.

(1) The *additive quadruple*, denoted by AQ (see Example 1.1). It is given by the equation $x + y = z + w$.
(2) The *heavy quadruple*, denoted by HQ. It is given by the equation $x + 2y = z + 2w$.

(3) The *heavy cycle*, denoted by HC. It is given by the equation $x + y + z = 3w$.
(4) The *skew quadruple*, denoted by SQ. It is given by the equation $2x + 2y = 3w + z$.

We already saw in the introduction that AQ is common. By an identical argument, so is the heavy quadruple HQ. What about the remaining two configurations?

Example 5.1 (The heavy cycle and the skew quadruple)**.** Both HC and SQ are uncommon in $\mathbb{F}_5^n$, and in fact it turns out that the same construction works for both configurations. As in Example 3.2, let $V \leqslant \mathbb{F}_5^n$ be a subspace of codimension 1, and let A be the union of the $+1$ and -1 cosets of V in $\mathbb{F}_5^n$, together with half the elements of V chosen at random. Then with high probability the density of A is 1/2, and it is verified (in a rather tedious manner) that

$$m_{HC}(A, A^C) = 0.105 < 2 \cdot \left(\frac{1}{2}\right)^4 .$$

Indeed, the calculation reduces to showing that there are comparatively few solutions in $\{+1, -1\}$ to $x + y + z = 3w$ and $2x + 2y = 3w + z$, respectively. Since it is imperative that the density of A be (at least close to) 1/2, we add half of the trivial coset at random, which keeps the total solution count below the expected value. Moreover, the same holds for A^C by the remarks in the introductory paragraph of this section.

By an argument similar to that outlined at the end of Section 4, it can be shown that a bad colouring for either of the above two configurations must have a large Fourier coefficient. This means that the colour distribution must exhibit a strong linear bias, i.e. be concentrated on a coset (or several) of a low-codimensional subspace. So again, in a qualitative sense the construction in Theorem 5.1 incorporates the true reason for HC and SQ being uncommon.

It is crucial at this point to try and understand what distinguishes AQ and HQ from HC and SQ. Clearly the former two equations exhibit some symmetry that is absent in the latter two. To make this precise we turn to a definition due to Ruzsa [26].[6]

Definition 5.1 (Genus)**.** Let $Z = \mathbb{F}_p^n$, and let $m \geq 2$ and $g \geq 1$ be integers. A translation-invariant equation of the form

$$b_1x_1 + b_2x_2 + \cdots + b_mx_m = 0$$

with variables $x_i \in \mathbb{F}_p^n$ and integer coefficients b_i is said to have *genus* g *over* $\mathbb{F}_p$ if g is the largest integer such that there is a partition of $\{1, 2, \ldots, m\}$ into g disjoint non-empty subsets $I_1, I_2, \ldots, I_g$ with the property that

$$\sum_{i \in I_j} b_i = 0$$

for every $j = 1, 2, \ldots, g$.

We immediately point out that of course this definition of genus is, just like the notion of translation invariance itself, dependent on the characteristic. It is easy to see that AQ and HQ have genus 2 over $\mathbb{F}_5$ (and any field of larger characteristic), while HC and SQ have genus 1. After a little thought this observation leads to the following conjecture.

Conjecture 5.2. Let $k \geq 2$ be an integer. A linear configuration given by a single equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k = a_{k+1}x_{k+1} + a_{k+2}x_{k+2} + \cdots + a_{2k}x_{2k} \tag{5.1}$$

is common in $\mathbb{F}_p^n$ if and only if it has genus k over $\mathbb{F}_p$.

[6]Ruzsa made this definition over the integers, but we use it over a finite field $\mathbb{F}_p$ here.

In one direction Conjecture 5.2 is easily seen to be true. Indeed, both the Fourier- and the Cauchy-Schwarz argument in Example 2.1 generalise to yield the result that any equation of genus k of the form (5.1) defines a common configuration. To test the reverse direction, we numerically investigated the case of translation-invariant equations in six variables over $\mathbb{F}_5$, of which only five are genuinely distinct.

$$x+y+z = u+v+w \tag{5.2}$$

$$x+y+2z = u+v+2w \tag{5.3}$$

$$x+y+z+w = 2u+2v \tag{5.4}$$

$$x+y+z+w+2u = v \tag{5.5}$$

$$x+y+z+2w+2u = 2v \tag{5.6}$$

The first two equations, (5.2) and (5.3), have genus 3 over $\mathbb{F}_5$, the remaining three equations have genus 2.[7] By the discussion above, configurations defined by either of the first two equations are therefore clearly common.

Example 5.2 (Genus < 3 in six variables over $\mathbb{F}_5$)**.** The configurations defined by any of the equations (5.4), (5.5) and (5.6) above are uncommon over $\mathbb{F}_5$. To see this for (5.5) and (5.6), use the cosets $1+V$ and $-1+V$ of a subspace $V \leqslant \mathbb{F}_5^n$ of codimension 1 as before, and add half the elements of V independently at random. For (5.4) we use $1+V$ and $-2+V$ instead.

Since there are no equations of genus 1 in six variables over $\mathbb{F}_5$, we also tested all fourteen translation-invariant equations in six variables over $\mathbb{F}_7$, of which five are of genus 3, eight are of genus 2 and one is of genus 1. All our results are consistent with Conjecture 5.2 above.

6. Concluding remarks

Having examined a reasonable number of examples, let us conclude by summarising the different behaviours we have encountered. We have witnessed that a linear configuration can be

(1) common because of cancellation (3-APs);
(2) common because of symmetry (AQ, HQ);
(3) uncommon because of skewed density (TP);
(4) uncommon because of pure linear bias (HC, SQ, SI);
(5) uncommon for quadratic reasons (4-APs, 5-APs).

Any reasonable conjecture concerning the classification of linear patterns as common or uncommon must take into account all of these possibilities.

Question 6.1. Formulate an arithmetic analogue of the Burr-Rosta conjecture.

Even though the analogy between the arithmetic and the graph setting is not perfect, this diverse array of underlying reasons may go some way towards explaining why the analogous graph-theoretic problem remains wide open despite having been studied for many years.

References

[1] Thomas Bloom. A quantitative improvement for Roth's theorem on arithmetic progressions. *arXiv*, 1405.5800, 2014.

[7] Note that only equation (5.4) has genus 2 over the integers.

[2] Stefan A Burr and Vera Rosta. On the Ramsey multiplicities of graphs—problems and recent results. *Journal of Graph Theory*, 4(4):347–361, 1980.

[3] Peter J Cameron, Javier Cilleruelo, and Oriol Serra. On monochromatic solutions of equations in groups. *Rev. Mat. Iberoam.*, 23:385–395, 2007.

[4] Pablo Candela and Olof Sisask. Convergence results for systems of linear forms on cyclic groups and periodic nilsequences. *SIAM J. Discrete Math.*, 28(2):786–810, 2014.

[5] David Conlon, Jacob Fox, and Benny Sudakov. An approximate version of Sidorenko's conjecture. *Geom. Funct. Anal.*, 20(6):1354–1366, 2010.

[6] David Conlon, Jacob Fox, and Benny Sudakov. Recent developments in graph Ramsey theory. *arXiv*, 1501.02474, 2015.

[7] David Conlon, Jeong Han Kim, Choongbum Lee, and Joonkyung Lee. Some advances on Sidorenko's conjecture. *arXiv*, 1510.06533, 2015.

[8] Ernie Croot. The minimal number of three-term arithmetic progressions modulo a prime converges to a limit. *Canad. Math. Bull.*, 51(1):47–56, 2008.

[9] Boris A Datskovsky. On the number of monochromatic Schur triples. *Advances in Applied Mathematics*, 31(1):193–198, 2003.

[10] Walter Deuber. Partition theorems for Abelian groups. *J. Combin. Theory Ser. A*, 19(1):95–108, July 1975.

[11] Paul Erdös. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 1962.

[12] A W Goodman. On Sets of Acquaintances and Strangers at any Party. *The American Mathematical Monthly*, 66(9):778, 1959.

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

[14] W. T. Gowers. Two examples in additive combinatorics. *Unpublished*, 2006.

[15] W. T. Gowers and Julia Wolf. The true complexity of a system of linear equations. *Proc. Lond. Math. Soc. (3)*, 100(1):155–176, 2010.

[16] W. T. Gowers and Julia Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, 21(1):36–69, 2011.

[17] Ben J. Green. Finite field models in additive combinatorics. In *Surveys in combinatorics 2005*, pages 1–27. Cambridge Univ. Press, Cambridge, Cambridge, 2005.

[18] Ben J. Green. Montréal notes on quadratic Fourier analysis. In *Additive combinatorics*, pages 69–102. Amer. Math. Soc., Providence, RI, 2007.

[19] Ben J. Green and Olof Sisask. On the maximal number of 3-term arithmetic progressions in subsets of $\mathbb{Z}/p\mathbb{Z}$. *Bull. Lond. Math. Soc.*, 40(6):945–955, 2008.

[20] Ben J. Green and Terence Tao. An inverse theorem for the Gowers $U^3(G)$ norm. *Proc. Edinb. Math. Soc. (2)*, 51(1):73–153, 2008.

[21] Chris Jagger. *Multiplicities of subgraphs*. PhD thesis, 1996.

[22] Chris Jagger, Pavel Stovicek, and Andrew Thomason. Multiplicities of subgraphs. *Combinatorica*, 16(1):123–141, 1996.

[23] J L Xiang Li and Balazs Szegedy. On the logarithmic calculus and Sidorenko's conjecture. *arXiv*, 1107.1153, 2011.

[24] Linyuan Lu and Xing Peng. Monochromatic 4-term arithmetic progressions in 2-colorings of $\mathbb{Z}_n$. *arXiv*, 1107.2888, 2011.

[25] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Combin. Theory Ser. A*, 71(1):168–172, 1995.

[26] Imre Z. Ruzsa. Solving a linear equation in a set of integers. I. *Acta Arith.*, 65(3):259–282, 1993.

[27] Wojciech Samotij and Benny Sudakov. The number of additive triples in subsets of abelian groups. *arXiv*, 1507.03764, 2015.

[28] Tom Sanders. On Roth's theorem on progressions. *Ann. of Math. (2)*, 174(1):619–636, 2011.

[29] Alexander Sidorenko. Cycles in graphs and functional inequalities. *Mat. Zametki*, 46(5):72–79, 104, 1989.

[30] Alexander Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.*, 9(2):201–204, 1993.

[31] Miklós Simonovits. Extremal graph problems, degenerate extremal problems, and supersaturated graphs. *Progress in graph theory*, 1984.

[32] Konrad Sperfeld. On the minimal monochromatic K_4-density. *arXiv*, 1106.1030, 2011.

[33] Balazs Szegedy. An information theoretic approach to Sidorenko's conjecture. *arXiv*, 1406.6738, 2014.
[34] Terence Tao and Van Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
[35] Andrew Thomason. A disproof of a conjecture of Erdős in Ramsey theory. *J. Lond. Math. Soc. (2)*, 39(2):246–255, 1989.
[36] Andrew Thomason. Graph products and monochromatic multiplicities. *Combinatorica*, 17(1):125–134, 1997.
[37] Julia Wolf. The minimum number of monochromatic 4-term progressions in $\mathbb{Z}_p$. *J. Comb.*, 1(1):53–68, 2010.
[38] Julia Wolf. Finite field models in arithmetic combinatorics – ten years on. *Finite Fields Appl.*, 32:233–274, 2015.

School of Mathematics, University of Bristol, Bristol BS8 1TW