Proceedings of the Edinburgh Mathematical Society (2020) **63**, 630–653
doi:10.1017/S0013091520000048

BOOTSTRAPPING PARTITION REGULARITY OF LINEAR SYSTEMS

TOM SANDERS

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK
(tom.sanders@maths.ox.ac.uk)

(Received 19 December 2019; first published online 9 March 2020)

Abstract Suppose that A is a $k \times d$ matrix of integers and write $\mathfrak{R}_A : \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ for the function taking r to the largest N such that there is an r-colouring $\mathcal{C}$ of $[N]$ with $\bigcup_{C\in\mathcal{C}} C^d \cap \ker A = \emptyset$. We show that if $\mathfrak{R}_A(r) < \infty$ for all $r \in \mathbb{N}$ then $\mathfrak{R}_A(r) \leqslant \exp(\exp(r^{O_A(1)}))$ for all $r \geqslant 2$. When the kernel of A consists only of Brauer configurations – that is, vectors of the form $(y, x, x + y, \dots, x + (d - 2)y)$ – the above statement has been proved by Chapman and Prendiville with good bounds on the $O_A(1)$ term.

Keywords: colouring; Rado numbers; Gowers uniformity

2010 *Mathematics subject classification:* Primary 05D10

1. Introduction

Our work concerns colourings. For a set X and natural r we say that $\mathcal{C}$ is an *r-colouring of X* if $\mathcal{C}$ is a cover of X, that is, $X \subset \bigcup_{C\in\mathcal{C}} C$, and $\mathcal{C}$ has size r. In particular, we shall not need our colours to be disjoint, though such colourings are included.

Suppose that A is a $k \times d$ matrix of integers. We write $\mathfrak{R}_A : \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ for the function taking r to the largest N such that there is an r-colouring $\mathcal{C}$ of $[N] := \{1, \dots, N\}$ with $\bigcup_{C\in\mathcal{C}} C^d \cap \ker A = \emptyset$ – in words, such that there are no monochromatic solutions to $Ax = 0$. Note that the function $\mathfrak{R}_A$ is monotonically increasing.

Not all matrices A have $\mathfrak{R}_A(r) < \infty$ for all $r \in \mathbb{N}$ (e.g., if all the non-zero terms in A are positive), but those that do we call *partition regular*. There are matrices A such that van der Waerden's theorem [**39**, Exercise 6.3.7] (first proved in [**40**]) is implied by the partition regularity of A (see [**31**, Satz I]), and similarly for Schur's theorem [**39**, 6.12] (first proved in [**36**]). Schur's theorem actually gives the stronger fact that* $\mathfrak{R}_A(r) \leqslant \lfloor er! \rfloor$, and since the celebrated work of Gowers [**12**, **13**] we know that van der Waerden's theorem also has reasonable bounds in terms of the number of colours.

It is the purpose of this paper to use Gowers' work to show the following theorem.

* It is a result of Abbott and Moser [**1**] that we cannot do much better.

© 2020 The Edinburgh Mathematical Society

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press
CrossMark

Theorem 1.1. *Suppose that A is a $k \times d$ integer-valued partition regular matrix and $r \geqslant 2$ is natural. Then there is some $N \leqslant \exp(\exp(r^{O_A(1)}))$ such that any r-colouring of $[N]$ contains a colour class C and some $x \in C^d$ such that $Ax = 0$.*

The basic method is expounded in the model setting[*] of $\mathbb{F}_2^n$ by Shkredov in [**37**, Theorem 24] for the purpose of illustrating how analytic techniques can be applied to colouring results. Chapman and Prendiville in [**6**] independently discovered the argument given in [**37**, Theorem 24] (though with some technical differences around expansion vs large Fourier coefficients) and importantly showed how it could be applied to provide good bounds in colouring problems in the integers where none were previously known. Specifically, in [**6**, Theorem 1.1] they prove Theorem 1.1 for Brauer configurations, meaning for a matrix A whose kernel is the set of vectors of the form $(y, x, x+y, \dots, x+(d-2)y)$ for some fixed $d \geqslant 3$, with a doubly exponential bound on d in place of the $O_A(1)$ term. (They also show in [**6**, Theorem 1.2] that one may replace the $O_A(1)$ term by $1 + o(1)$ for Brauer configurations with $d = 4$.)

It is the purpose of this note to extend the arguments of Chapman and Prendiville to partition regular linear systems. This entails a large notational burden and as a result, while they are able to give rather good estimates for the $O_A(1)$-term when A is a matrix corresponding to a Brauer configuration, we give no meaningful estimates.[†]

The above work comes on the back of a wave of investigations using analytic techniques for colouring problems. This really took off with the paper [**8**] by Cwalina and Schoen, and was followed by the work of Green and collaborators [**18**, **19**], then Chow *et al.* [**7**], and most recently Chapman [**5**] which inspired this particular paper.

One would often like to insist that the x found in Theorem 1.1 is in a certain sense non-degenerate. The extent to which this is possible varies, but the question has been dealt with comprehensively by Hindman and Leader in [**25**]. See also [**10**] for a related supersaturated formulation.

1.2. Existing bounds on the Rado numbers $\mathfrak{R}_A(r)$

Other than the aforementioned [**6**, Theorems 1.1 and 1.2], most work has focused on the case where A has one row, that is, systems with one equation which for clarity we write in the comma-delimited form $A = (a_1, \dots, a_k)$. In this case Rado's theorem [**26**, Theorem 9.5] tells us that if (and only if) A is partition regular then there is $\emptyset \neq I \subset [k]$ such that $\sum_{i \in I} a_i = 0$.

Schur's theorem itself gives rather good bounds on $\mathfrak{R}_A(r)$ when $A = (1, 1, -1)$, and more generally [**8**, Theorem 1.3] gives singly exponential bounds when A is a partition regular row. Stronger results when the equation satisfies additional properties are given in [**8**, Theorems 1.4 and 1.5] and [**11**, Theorem 4.7].

When $A = (1, \dots, 1, -1)$ the numbers $\mathfrak{R}_A(r)$ are sometimes called the generalized diagonal Schur numbers (although they are just called Schur numbers in [**3**]). These have been computed for many values of r, being completely known for $r = 2$ [**3**, Theorem 1.3], and for $r \geqslant 3$ the reader is directed to [**2**, Table 1] for recent calculations. Note that the

[*] The model setting has proved very fruitful for distilling the important aspects of arguments in additive combinatorics. See the paper [**17**] and the sequel [**42**].

[†] Though see the remark after the proof of Theorem 2.3.

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

bounds in Theorem 1.1 as ineffective so, for example, when $r = 2$ our result says nothing more than $\mathfrak{R}_A(2) < \infty$.

When A has just one row there is a large body of work computing the exact value of $\mathfrak{R}_A(2)$ using arguments which are much more combinatorial than those in the present paper. This work has many extensions covering things such as Rado numbers for inhomogeneous equations [**26**, p. 259]; off-diagonal Rado numbers [**26**, p. 280]; and Rado numbers for disjunctive equations [**26**, p. 293]. We restrict ourselves to recording those results which ask for bounds on $\mathfrak{R}_A(2)$ under the same hypotheses as Theorem 1.1.

When $A = (a_1, a_2, -a_2)$ for $a_1, a_2 \in \mathbb{N}$ the value of $\mathfrak{R}_A(2)$ is computed in [**26**, Theorem 9.17]; when $A = (a_1, a_2, -(a_1 + a_2))$ for $a_1, a_2 \in \mathbb{N}$ the value of $\mathfrak{R}_A(2)$ is computed in [**23**, Theorem 1.1]; when $A = (1, 1, a_3, -a_4)$ for $a_3, a_4 \in \mathbb{N}$ (where partition regularity of A ensures that $a_4 \in \{1, 2, a_3, a_3 + 1, a_3 + 2\}$) the value of $\mathfrak{R}_A(2)$ is computed in [**33**, Theorems 3, 4 and 8] for $a_4 = a_3$, $a_4 = a_3 + 1$ and $a_4 = 2$ respectively, with the case $a_4 = a_3 + 2$ being trivial; and when

$$A = (\overbrace{1, \dots, 1}^{n \text{ times}}, -a_{n+1}, \dots, -a_k) \tag{1.1}$$

with $a_{n+1}, \dots, a_k \in \mathbb{N}$ and $n \geqslant a_{n+1} + \cdots + a_k$, the value of $\mathfrak{R}_A(2)$ is computed in [**35**, Theorem 3]. Note that the work of [**26**], [**33**] and [**35**] goes further and computes $\mathfrak{R}_A(2)$ for some A which are not partition regular. (This makes sense since we may have $\mathfrak{R}_A(2) < \infty$ without $\mathfrak{R}_A(r) < \infty$ for all $r \in \mathbb{N}$. See [**26**, Theorem 9.2] for conditions on a single row A such that $\mathfrak{R}_A(2) < \infty$.)

Finally, [**8**, Theorem 1.8] shows that $\mathfrak{R}_A(r) = o_{r\to\infty}(r!)$ when A is as in (1.1) with $n = 3$, $l = 2$ and $a_1, a_2 = 1$, beating the bound following from Schur's argument.

1.3. Variable conventions

There are some conflicts between standard uses for certain symbols in different areas. m, p and c are the parameters of an (m, p, c)-set in Deuber's sense (as defined in § 2.1), so that c need not be an absolute constant, and p need not be prime. $\mathcal{C}$ usually denotes a colouring and $\mathfrak{C}$ the conjugation operator (see § 2.6). C then typically denotes a colour class in $\mathcal{C}$, rather than an absolute constant.

1.4. Big-O notation

We use big-O notation in the usual way; see, for example, [**39**, p. 11]. The constants behind the big-O and Ω expressions may depend in peculiar ways on other parameters, and we shall sometimes need some control. We capture this in the same way as [**19**, p. 17]. Some big-O expressions will be replaced by 'universal functions' of the form $f : D_1 \times \cdots \times D_k \to D_0$ where each D_i is one of the sets $(0, 1]$, $\mathbb{N}_0$, or $\mathbb{N}$. If $D_i = (0, 1]$ then we write $x \preceq_i y$ if and only if $y \leqslant x$; otherwise we write $x \preceq_i y$ if and only if $x \leqslant y$. We say that f is *monotone* if $f(x) \preceq_0 f(y)$ whenever $x_i \preceq_i y_i$ for all $1 \leqslant i \leqslant k$.

Note that the above is the usual order on $\mathbb{N}$ and $\mathbb{N}_0$ and the opposite of the usual order on $(0, 1]$. This reflects the fact that we shall want bounds on, say, the size of an interval which do not get too much worse as, say, the number of colours grows (that would be a natural number parameter) and also as the density of some related set does not get too

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

small (that would be a $(0,1]$ parameter). Our notation for monotone aligns these different notions of large and small to point in the same direction.

It is useful to note that if $f(x) = O_a(g(x))$ where $a \in \mathbb{N}_0^d$ then there is a monotone function $F : \mathbb{N}_0^d \to \mathbb{N}$ such that

$$|f(x)| \leqslant F(a)g(x), \quad \text{for all } x.$$

This can be shown by letting $F(a)$ be the max of the constants behind the $O_{a'}$ term as $a' \preceq a$ – a finite set.

The universal functions mapping into $\mathbb{N}$ or $\mathbb{N}_0$ will usually be denoted by Fs with various decorations (e.g., subscripts and superscripts), while those mapping into $(0,1]$ will usually be denoted by ηs with various decorations. To avoid too many different functions, we shall often use the same functions in situations where the optimal functions are almost certainly different but where there is little cost to doing so.

2. Set-up and tools

In this section we record the tools we need. First, in § 2.1, we explain Deuber's framework [**9**] for understanding colouring problems. This will reduce the problem to proving Theorem 2.3. The key tools to prove this are recorded in § 2.6. Finally, we gather a few more technical facts in § 2.9.

2.1. Deuber's theorem

In [**9**, Satz 3.1] Deuber proved a conjecture of Rado, and we shall use Deuber's ideas here too. We follow the exposition and definitions of [**22**]: as in [**22**, Definition 2.5][*], given $m, p, c \in \mathbb{N}$, a set $S \subset \mathbb{N}$ is an (m,p,c)-*set* if there is some $s = (s_0, \dots, s_{m+1}) \in \mathbb{N}^{m+1}$ such that

$$S = \bigcup_{j=0}^{m} \{cs_{m-j} + i_{m-j+1}s_{m-j+1} + \cdots + i_m s_m : -p \leqslant i_{m-j+1}, \dots, i_m \leqslant p\}.$$

For example, if $m = 2$ then

$$S = \{cs_2\} \cup \{cs_1 + i_2 s_2 : -p \leqslant i_2 \leqslant p\} \cup \{cs_0 + i_1 s_1 + i_2 s_2 : -p \leqslant i_1, i_2 \leqslant p\},$$

and even more concretely, the set $\{s_1\} \cup \{s_0 - s_1, s_0, s_0 + s_1\}$ – which is a three-term arithmetic progression along with its common difference – is a $(1,1,1)$-set.

We shall give a little more motivation for these sets in a moment, but first we state Deuber's theorem.

Theorem 2.2 (Deuber's theorem [22, Theorem 2.8]). *Suppose that $m, p, c, r \in \mathbb{N}$. Then there are $M, P, C \in \mathbb{N}$ such that any r-colouring of an (M,P,C)-set contains a monochromatic (m,p,c)-set.*

The main result of this paper is the following theorem.

[*] Which Gunderson notes is slightly different from Deuber's original.

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

Theorem 2.3. *Suppose that $m, p, c, r \in \mathbb{N}$. Then there is $N \leqslant \exp(\exp(r^{O_{m,p,c}(1)}))$ such that any r-colouring of $[N]$ contains a monochromatic (m, p, c)-set.*

Qualitatively this is a special case of Deuber's theorem since $[N]$ is a $(1, N, N+1)$-set.

One of the reasons why (m, p, c)-sets are important is their relationship with solutions of equations, which we now explain. In [31, Satz IV] Rado famously proved that partition regularity of a system is equivalent to something called the columns condition: we say that a $k \times d$ matrix A satisfies the *columns condition* if there exist a $d \times t$ matrix of rationals α and a partition $[d] = I_1 \sqcup \cdots \sqcup I_t$, such that, writing $a_1, \dots, a_d$ for the columns of A in their given order, we have

$$\sum_{i \in I_{j+1}} a_i = \sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{ij} a_i, \quad \text{for all } 0 \leqslant j < t,$$

with the usual convention that the empty sum is 0. When we need to refer to a specific α we shall call it a *witness* for the columns condition. It is natural to assume that $\alpha_{ij} = 0$ for all $i \in I_{j+1} \cup \cdots \cup I_t$, and we shall always do so without remark. In view of this we see that $t = 1 + \operatorname{rk} \alpha$.

Theorem 2.4 (Rado's theorem, [22, Theorem 2.3]). *Suppose that A is a $k \times d$ integer-valued matrix. Then A is partition regular if and only if A satisfies the columns condition.*

Deuber connected the columns condition to (m, p, c)-sets through the following result.

Theorem 2.5 ([22, Theorem 2.6(i)]). *Suppose that A is a $k \times d$ integer-valued matrix satisfying the columns condition as witnessed by α. Then, writing c for the least common multiple of the denominators of the rationals in α, every $(1 + \operatorname{rk} \alpha, \max_{i,j} |c\alpha_{ij}|, c)$-set S has some $x \in S^d$ such that $Ax = 0$.*

This is not [22, Theorem 2.6(i)] as stated, but a quick look at the proof shows that this is what is proved.

Proof of Theorem 1.1 given Theorem 2.3. By Theorems 2.4 and 2.5, we see that if A is partition regular then there are naturals $m, p, c = O_A(1)$ such that any (m, p, c)-set S contains some $x \in S^d$ with $Ax = 0$. By Theorem 2.3, we see that for $N \leqslant \exp(\exp(r^{O_{m,p,c}(1)})) = \exp(\exp(r^{O_A(1)}))$ any r-colouring of $[N]$ has a colour class C containing a set S that is an (m, p, c)-set, and hence there is some $x \in S^d \subset C^d$ with $Ax = 0$ as required. □

2.6. Gowers norms

The Gowers norms are defined in [13, Lemma 3.9], though they are not given that name, and while they can be defined more generally for finite abelian groups we shall restrict attention to cyclic groups of prime order (in line with [13]). We use [38] as our basic reference, though admittedly many of the results there are left as exercises. The material is developed in considerable generality in [21]; the generality we need is closer to that discussed in [14, §2]. (Other introductions may be found in many places,

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

including [20, § 4], [24, §§ 2 and 3], [41, Appendix A], and [29, § 1]. All these, including [14], ultimately refer to [21] for details, though the paper [41] does expand on the details somewhat in § 4.)

For $N \in \mathbb{N}$ (which will be prime, though need not be right now), $k \in \mathbb{N}$ and $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ we put

$$\|f\|_{U^k(\mathbb{Z}/N\mathbb{Z})} := \left(\mathbb{E}_{x,h_1,\dots,h_k \in \mathbb{Z}/N\mathbb{Z}} \prod_{\omega \in \{0,1\}^k} \mathfrak{C}^{|\omega|} f(x + \omega \cdot h)\right)^{2^{-k}},$$

where $\mathfrak{C}$ denotes the operation of complex conjugation.

The map $\|\cdot\|_{U^k(\mathbb{Z}/N\mathbb{Z})}$ defines a norm [38, Exercise 1.3.19] for $k \geqslant 2$, and enjoys the nesting property $\|\cdot\|_{U^k(\mathbb{Z}/N\mathbb{Z})} \leqslant \|\cdot\|_{U^{k+1}(\mathbb{Z}/N\mathbb{Z})}$ for $k \in \mathbb{N}$ [38, Exercise 1.3.19]. (Proofs of these two facts are given explicitly in [39, p. 466] and [39, (11.7)].)

One of the reasons why these norms are important is that they control counts of various linear configurations. Specifically, suppose that $\Psi : \mathbb{Z}^d \to \mathbb{Z}^l$ is a homomorphism and f is a vector of l functions $\mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$. We define

$$\Lambda_\Psi(f) := \mathbb{E}_{x \in [N]^d} \prod_{i=1}^{l} f_i(\Psi_i(x) + N\mathbb{Z}). \tag{2.1}$$

The following result is the 'generalized von Neumann theorem' we need. It is a special case of [20, Theorem 4.1] once the notation has been unpacked, and also of [38, Exercise 1.3.23] combined with [38, Exercise 1.3.14].

Theorem 2.7. *Suppose that $\Psi : \mathbb{Z}^d \to \mathbb{Z}^l$ is a homomorphism and, for every $i \neq j$, (Ψ_i, Ψ_j) is a pair of independent vectors (i.e., if $z\Psi_i + w\Psi_j \equiv 0$ for some $z, w \in \mathbb{Z}$ then $z = w = 0$). Then there are naturals $N_0(\Psi)$ and $k(\Psi)$ such that if $N \geqslant N_0(\Psi)$ is a prime and f is a vector of l functions $\mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ bounded by 1, then*

$$|\Lambda_\Psi(f)| \leqslant \inf_{1 \leqslant i \leqslant l} \|f_i\|_{U^k(\mathbb{Z}/N\mathbb{Z})}.$$

We shall use the above to count (m, p, c)-sets. This was already done by Lê in [27] for the purpose of transferring the partition regularity of Brauer configurations to the sets $\{p - 1 : p \text{ is prime}\}$ and $\{p + 1 : p \text{ is prime}\}$, itself answering a question of Li and Pan [28].

We also need Gowers' inverse theorem. The following result is what is proved in [13, Theorem 18.1], though it is not stated in precisely this way.

Theorem 2.8. *There is a monotone function $F_1 : \mathbb{N} \to \mathbb{N}$ such that the following statement holds. Suppose that N is prime, $\epsilon \leqslant \frac{1}{2}$ and $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is bounded in magnitude by 1, with $\|f\|_{U^k(\mathbb{Z}/N\mathbb{Z})} \geqslant \epsilon$. Then there is a partition of $[N]$ into arithmetic progressions $P_1, \dots, P_M$ of average size at least $N^{\epsilon^{F_1(k)}}$ such that*

$$\sum_{j=1}^{M} \left| \sum_{s \in P_j} f(s + N\mathbb{Z}) \right| \geqslant \epsilon^{F_1(k)} N.$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

2.9. Convolution, dilation and progressions

First we record notation for dilation and translation: given $x, y \in \mathbb{Z}$, we write $\lambda_x(y) := xy$; further, given $f : \mathbb{Z} \to \mathbb{C}$, we write $\tau_x(f)(y) := f(y + x)$.

Suppose that $P \subset \mathbb{Z}$ is an arithmetic progression of odd length. Then

$$P = x_P + d_P \cdot \{-N_P, \dots, N_P\}$$

for some $x_P \in \mathbb{Z}$ called the *centre*, $d_P \in \mathbb{N}_0$ called the *common difference*, and $N_P \in \mathbb{N}_0$ called the *radius*. Technically the common difference and radius need not be uniquely defined, but this only becomes a problem for arithmetic progressions of size 1 where the necessary adaptations of any argument are trivial and omitted for clarity.

We work with arithmetic progressions of odd length for convenience, not because of any important difference. We say that P is a *centred arithmetic progression* if $x_P = 0$, so in particular a centred arithmetic progression is of odd length.

For $\delta \geqslant 0$ we shall define

$$I_\delta(P) := d_P \cdot \{-\lfloor \delta N_P \rfloor, \dots, \lfloor \delta N_P \rfloor\},$$

and below record some basic properties of these 'fractional dilates' of progressions. In many cases these properties are special cases of properties of Bohr sets (see [**39**, §4.4]). We do not require the generality of Bohr sets here because the result of Gowers' inverse theorem (Theorem 2.8) is a decomposition in terms of progressions. This has the additional benefit of meaning that we do not need to deal with the problem of finding regular Bohr sets (see [**39**, Lemma 4.24]), since all progressions are regular in a suitable sense. This is captured in the last three properties below.

Lemma 2.10 (Basic properties). *Suppose that P and P' are arithmetic progressions of odd length, and that $c \in \mathbb{N}$, $x \in \mathbb{Z}$ and $\delta, \delta' \in (0, 1]$.*

(i) (Symmetry) *$I_\delta(P)$ is a centred progression of size at least $\frac{1}{3}\delta|P|$.*

(ii) (Monotonicity in radius) *$I_{\delta'}(P) \subset I_\delta(P)$ whenever $\delta' \leqslant \delta$.*

(iii) (Monotonicity in progression) *$I_\delta(P') \subset I_\delta(P)$ whenever $P' \subset P$.*

(iv) (Translation) *$x + P$ is a progression of odd length and $I_\delta(P) = I_\delta(x + P)$.*

(v) (Dilations) *$c \cdot P$ is an arithmetic progression of odd length and $I_\delta(c \cdot P) = c \cdot I_\delta(P)$.*

(vi) (Sub-additivity) *$I_\delta(P) + I_{\delta'}(P) \subset I_{\delta+\delta'}(P)$.*

(vii) (Composition) *$I_\delta(I_{\delta'}(P)) \subset I_{\delta\delta'}(P)$.*

(viii) (Interiors) *there is an arithmetic progression of odd length, $\mathrm{Int}_\delta(P)$, such that*

$$\mathrm{Int}_\delta(P) + I_\delta(P) \subset P \quad \text{and} \quad |\mathrm{Int}_\delta(P)| \geqslant (1 - \delta)|P|.$$

(ix) (Closures) *$P + I_\delta(P)$ is an arithmetic progression of odd length and*

$$|P + I_\delta(P)| \leqslant (1 + \delta)|P|.$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

(x) (Invariance) *For all* $f : \mathbb{Z} \to \mathbb{C}$ *and* $y \in I_\delta(P)$ *we have*

$$|\mathbb{E}_{x\in P}\tau_y(f)(x) - \mathbb{E}_{x\in P}f(x)| \leqslant 2\delta\|f\|_{L_\infty}.$$

Proof. (i) is trivial on noting that $|I_\delta(P)| = 2\lfloor\delta N_P\rfloor + 1 \geqslant \frac{1}{3}\delta(2N_P+1)$. (ii)–(v) are immediate. (vi) follows since $\lfloor\delta N_P\rfloor + \lfloor\delta' N_P\rfloor \leqslant \lfloor(\delta+\delta')N_P\rfloor$; and (vii) since $\lfloor\delta\lfloor\delta' N\rfloor\rfloor \leqslant \lfloor\delta\delta' N\rfloor$. For (viii) set

$$\mathrm{Int}_\delta(P) := x_P + d_P\cdot\{-(N_P - \lfloor\delta N_P\rfloor), \dots, N_P - \lfloor\delta N_P\rfloor\},$$

so that $\mathrm{Int}_\delta(P)$ is an arithmetic progression of odd length, $\mathrm{Int}_\delta(P) + I_\delta(P) \subset P$, and

$$|\mathrm{Int}_\delta(P)| \geqslant 2(N_P - \lfloor\delta N_P\rfloor) + 1 = 2N_P + 1 - 2\lfloor\delta N_P\rfloor \geqslant (1-\delta)|P|.$$

For (ix) we note that

$$P + I_\delta(P) = x_P + d_P\cdot\{-N_P, \dots, N_P\} + d_P\cdot\{-\lfloor\delta N_P\rfloor, \dots, \lfloor\delta N_P\rfloor\},$$

and so

$$|P + I_\delta(P)| = 2(N_P + \lfloor\delta N_P\rfloor) + 1 \leqslant (1+\delta)(2N_P+1).$$

For (x) note that if $s' \in I_\delta(P)$ then

$$\begin{aligned}
&|\mathbb{E}_{s\in P}\tau_{s'}(f)(s) - \mathbb{E}_{s\in P}f(s)|\\
&\quad= |\mathbb{E}_{s\in P}(1 - 1_{\mathrm{Int}_\delta(P)})(s+s')f(s+s') + \mathbb{E}_{s\in P}1_{\mathrm{Int}_\delta(P)}(s+s')f(s+s')\\
&\qquad - \mathbb{E}_{s\in P}1_{\mathrm{Int}_\delta(P)}(s)f(s) - \mathbb{E}_{s\in P}(1 - 1_{\mathrm{Int}_\delta(P)})(s)f(s)|\\
&\quad\leqslant |\mathbb{E}_{s\in P}(1 - 1_{\mathrm{Int}_\delta(P)})(s+s')f(s+s')| + |\mathbb{E}_{s\in P}(1 - 1_{\mathrm{Int}_\delta(P)})(s)f(s)| \leqslant 2\delta\|f\|_{L_\infty}.
\end{aligned}$$

The result is proved. □

3. An example

The notation in the final proof is quite heavy, so before turning to this we present an example case kindly suggested by one of the referees.

The arguments of [6] work to deal with $(1,p,c)$-sets (see [6, Theorem 5.1]) and are similar to ours in terms of how these sorts of sets are dealt with. The additional complexity we encounter is in dealing with (m,p,c)-sets for $m > 1$. Our approach is inductive on m, and our intention is that by motivating the $m = 2$ case the general argument will become clear.

We consider the problem of finding monochromatic septuples

$$(x; y, x+y; z, z+x, z+y, z+x+y) \tag{3.1}$$

in r-colourings of $\{1, \dots, N\}$. Such septuples do not correspond to an (m,p,c)-set, but for our purposes behave rather like a $(2,1,1)$-set. In fact looking for configurations of this type is a special case of Folkman's theorem [16, Theorem 11, § 3.4] (an explanation of

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

the name may also be found in that reference), which was discovered independently by Folkman,[*] Rado [**32**], and Sanders [**34**, Theorem 2].

We treat the three sets of terms in (3.1) separated by semicolons at three different scales. In particular, we shall find arithmetic progressions P_1, P_2 and P_3 iteratively by using the Gowers inverse theorem (Theorem 2.8) to give a density increment for a colour class on a certain progression. This increment translates to an increment to the sum over all colour classes of their maximum density on the translate of a progression. Importantly the translate may be different for different colour classes; and the process terminates since the sum of the maximal densities is bounded above by r. On termination we have control of some localized Gowers norms like

$$\|f\|_{U^3(P_i;P_3)} := \mathbb{E}_{z\in P_3}\|f1_{z+P_i}\|_{U^3}, \quad \text{for } i \in \{1,2\}.$$

Localized Gowers norms of this type are defined in [**30**, (2.12)] among other places and the additional discussion around that definition may be of interest.

The generalized von Neumann theorem (Theorem 2.7) ensures that control of these localized Gowers norms of a colour class C on the translate $a_C + P_3$ on which C has maximal density δ leads to

$$\begin{aligned}\mathbb{E}_{x\in P_1,y\in P_2,z\in a_C+P_3}&1_C(x)1_C(y)1_C(x+y)1_C(z)1_C(x+z)1_C(y+z)1_C(x+y+z)\\ &\approx \delta^4\mathbb{E}_{x\in P_1,y\in P_2}1_C(x)1_C(y)1_C(x+y).\end{aligned}$$

Inductively we can ensure that this second term is large for some colour class C, and the largeness of that colour class in turn ensures that δ is large. This gives a large count of septuples in one colour class as required.

4. The proof

We begin by recording some notation for linear forms associated with (m,p,c)-sets. For $p,c\in\mathbb{N}$ and $t\in\mathbb{N}_0$ put

$$\mathcal{D}_{p,c;t} := \left\{i\in\mathbb{Z}^{\mathbb{N}_0} : i_j\in\{-p,\dots,p\} \text{ for all } j<t; i_t = c; \text{ and } i_j = 0 \text{ for all } j>t\right\}.$$

The sets $\mathcal{D}_{p,c;t}$ as t ranges over $\mathbb{N}_0$ are disjoint. For $m\in\mathbb{N}_0$ put

$$\mathcal{U}_{m,p,c} := \bigcup_{t=0}^{m}\mathcal{D}_{p,c;t}.$$

Then for $i\in\mathcal{U}_{m,p,c}$ write

$$L_i : \mathbb{Z}^{\mathbb{N}_0}\to\mathbb{Z}; s\mapsto \sum_{j:i_j\neq 0} s_j i_j, \tag{4.1}$$

which is well defined since the set of j such that $i_j\neq 0$ has size at most $m+1$ (and, in particular, is finite) for $i\in\mathcal{U}_{m,p,c}$.

[*] Folkman's proof was unpublished, but a record of the fact he proved is found in [**15**, Corollary 4].

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

The unique element of $\mathcal{D}_{p,c;t}$ with support of size 1 is particularly important: we put

$$i^*(c,t) := (\overbrace{0,\dots,0}^{t \text{ times}},c,0,\dots) \quad \text{and} \quad \mathcal{D}^*_{p,c;t} := \mathcal{D}_{p,c;t} \setminus \{i^*(c,t)\}.$$

Suppose that $P_0,\dots,P_m$ are arithmetic progressions. We are interested in the count

$$Q_{m,p,c}(A;P_0,\dots,P_m) := \mathbb{E}_{s_0\in P_0,\dots,s_m\in P_m} \prod_{i\in\mathcal{U}_{m,p,c}} 1_A(L_i(s)), \tag{4.2}$$

since if $P_0,\dots,P_m \subset \mathbb{N}$ then this quantity is non-zero only if A contains an (m,p,c)-set.

The following is our key counting/density-increment dichotomy.

Lemma 4.1. *There are monotone functions $F_2 : \mathbb{N}^3 \to \mathbb{N}$ and $\eta_0 : \mathbb{N}^3 \to (0,1]$ such that the following result holds. Suppose that $m,p,c \in \mathbb{N}$, $\delta \in (0,1]$ and $P_0,\dots,P_m \subset \mathbb{Z}$ are arithmetic progressions of odd length with*

$$P_i \subset I_\delta(c\cdot P_m), \quad \text{for all } 0 \leqslant i \leqslant m-1; \tag{4.3}$$

$P'' \subset \mathbb{Z}$ is a centred arithmetic progression with

$$P'' \subset I_\delta(P_i), \quad \text{for all } 0 \leqslant i \leqslant m; \tag{4.4}$$

$P_m \subset \mathbb{N}$ and $A \subset \mathbb{Z}$ has $\alpha := \mathbb{E}_{x\in c\cdot P_m} 1_A(x) > 0$. Then at least one of the following statements holds:

(i) $|P''| \leqslant \exp(\delta^{-F_2(m,p,c)})$;

(ii) $\delta \geqslant \eta_0(m,p,c)$;

(iii) *there is an arithmetic progression of odd length, $P''' \subset \mathbb{N}$, with $I_1(c\cdot P''') \subset I_{F_2(m,p,c)\delta^2}(c\cdot P_m)$, such that*

$$|P'''| \geqslant |P''|^{\delta^{F_2(m,p,c)}} \quad \text{and} \quad \mathbb{E}_{x\in c\cdot P'''} 1_A(x) \geqslant \alpha+\delta;$$

(iv) *or*

$$\Big|Q_{m,p,c}(A;P_0,\dots,P_m) - \alpha^{|\mathcal{D}_{p,c;m}|} Q_{m-1,p,c}(A;P_0,\dots,P_{m-1})\Big| \leqslant F_2(m,p,c)\delta^{\eta_0(m,p,c)}.$$

The proof below is long but not at all conceptually difficult. The length is a result of taking care with technicalities and somewhat licentious notation. The basic idea is to use Theorem 2.7 to control the Qs by suitable uniformity norms and then Theorem 2.8 to show that if that error is not small then there is a density increment. There are two types of density increment. One is the expected increment resulting from large U^k norm in Theorem 2.8. The other results from ensuring that the density of A is the same on two progressions, one of which is a small dilate of the other. This second increment is common to arguments where groups are replaced by approximate groups – in this case

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

progressions – and they perhaps originate in the work of Bourgain [4]. (See [39, (10.16)] and the definition of the function G there.)

Proof. With L_i defined as in (4.1), let $\Psi := (L_i)_{i\in\mathcal{U}_{m,p,c}}$ so that $\Psi : \mathbb{Z}^{m+1} \to \mathbb{Z}^{\mathcal{U}_{m,p,c}}$ is a homomorphism. Every (L_i, L_j) with $i \neq j$ is a pair of independent vectors, so by Theorem 2.7 applied to Ψ there is $k = k(\Psi) = O_{m,p,c}(1)$ such that if $N \geqslant N_0(\Psi)$ is prime and h is a vector of functions $\mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ (indexed by $\mathcal{U}_{m,p,c}$) then

$$\left|\mathbb{E}_{x\in[N]^{m+1}} \prod_{i\in\mathcal{U}_{m,p,c}} h_i(L_i(x) + N\mathbb{Z})\right| \leqslant \inf_{i\in\mathcal{U}_{m,p,c}} \|h_i\|_{U^k(\mathbb{Z}/N\mathbb{Z})}. \tag{4.5}$$

Note that this will be applied with a vector h to be determined, but which will be a combination of translates of the set A suitably restricted, and also translates of the balanced function of A. The particular choice is made in (4.14) (which itself depends on (4.13) and (4.10)).

As remarked in § 1.4, since m, p and c are in $\mathbb{N}$ we see that there is a monotone function $F : \mathbb{N}^3 \to \mathbb{N}$ such that $F(m, p, c) \geqslant N_0(\Psi)$. (We shall take $F_2 \geqslant F$, but there are other functions later which will determine exactly what F_2 needs to be.)

Since P'' is a centred arithmetic progression there are natural numbers d'' and N'' such that $P'' = d'' \cdot \{-N'', \dots, N''\}$. By Bertrand's postulate, there is a prime N such that

$$\max\{(mp + c)N'', F(m, p, c)\} < N = O_{m,p,c}(N'').$$

The reason for this choice will become clear just before (4.15) below. Before that we record the following claim.

Claim. *There is $\delta' = O_{m,p,c}(\delta)$ such that if $s_0 \in P_0, \dots, s_{m-1} \in P_{m-1}$, $x \in d'' \cdot \{-N, \dots, N\}$ and $i \in \mathcal{D}_{p,c;m}$ then*

$$L_i(s) + x - cs_m \in c \cdot I_{\delta'}(P_m),$$

and if additionally $s_m \in \mathrm{Int}_{\delta'}(P_m)$ then

$$L_i(s) + x \in c \cdot P_m.$$

Proof. Write $l = \lceil N/N'' \rceil$ so that, by (4.3), (4.4) and Lemma 2.10, we have

$$\begin{aligned} d'' \cdot \{-N, \dots, N\} &\subset \overbrace{(d'' \cdot \{-N'', \dots, N''\}) + \cdots + (d'' \cdot \{-N'', \dots, N''\})}^{l \text{ times}} \\ &= lP'' \subset I_{l\delta}(P_0) \subset I_{l\delta}(I_\delta(c \cdot P_m)) \subset I_{l\delta^2}(c \cdot P_m). \end{aligned}$$

If follows that if $i \in \mathcal{D}_{p,c;m}$, $s_0 \in P_0, \dots, s_{m-1} \in P_{m-1}$, and $x \in d'' \cdot \{-N, \dots, N\}$, then, by (4.3) and Lemma 2.10, we have

$$i_0 s_0 + \cdots + i_{m-1} s_{m-1} + x \in mpI_\delta(c \cdot P_m) + I_{l\delta^2}(c \cdot P_m) \subset I_{(mp+l\delta)\delta}(c \cdot P_m). \tag{4.6}$$

Let $\delta' := (mp + l\delta)\delta = O_{m,p,c}(\delta)$. If $i \in \mathcal{D}_{p,c;m}$ we have $i_m = c$ and then, by (4.6) and Lemma 2.10, we have

$$L_i(s) + x - cs_m = i_0 s_0 + \cdots + i_{m-1} s_{m-1} + x \in I_{\delta'}(c \cdot P_m) = c \cdot I_{\delta'}(P_m),$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

giving the first conclusion. Finally, if $s_m \in \mathrm{Int}_{\delta'}(P_m)$ then, by Lemma 2.10 again,

$$L_i(s) + x \in cs_m + c \cdot I_{\delta'}(P_m) \subset c \cdot (\mathrm{Int}_{\delta'}(P_m) + I_{\delta'}(P_m)) \subset c \cdot P_m.$$

The claim is proved. □

Let $\epsilon > 0$ be a further constant to be optimized later and suppose (using τ for translation as defined in § 2.9) that for some $j \in \mathcal{D}_{p,c;m}$ we have

$$\mathbb{E}_{s_0 \in P_0, \dots, s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) |\mathbb{E}_{x \in d'' \cdot [N]} \tau_{L_j(s)}(1_A - \alpha 1_{c \cdot P_m})(x)| > \epsilon. \tag{4.7}$$

By interchanging order of summation, we have

$$\begin{aligned} &|\mathbb{E}_{s_0 \in P_0, \dots, s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \mathbb{E}_{x \in d'' \cdot [N]} \tau_{L_j(s)}(1_A - \alpha 1_{c \cdot P_m})(x)| \\ &\quad = |\mathbb{E}_{s_0 \in P_0, \dots, s_{m-1} \in P_{m-1}, x \in d'' \cdot [N]} \\ &\qquad (\mathbb{E}_{s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \tau_{L_j(s)+x-cs_m} 1_A(cs_m) \\ &\qquad\quad - \alpha \mathbb{E}_{s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \tau_{L_j(s)+x-cs_m}(1_{c \cdot P_m})(cs_m))|. \end{aligned} \tag{4.8}$$

Suppose that $s_0 \in P_0, \dots, s_{m-1} \in P_{m-1}$ and $x \in d'' \cdot [N]$. The claim and Lemma 2.10 tell us that

$$\begin{aligned} \mathbb{E}_{s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \tau_{L_j(s)+x-cs_m} 1_A(cs_m) &\leqslant \mathbb{E}_{s_m \in P_m} \tau_{L_j(s)+x-cs_m} 1_A(cs_m) \\ &\leqslant \mathbb{E}_{s_m \in P_m} 1_A(cs_m) + 2\delta' = \alpha + 2\delta' \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}_{s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \tau_{L_j(s)+x-cs_m} 1_A(cs_m) &\geqslant \mathbb{E}_{s_m \in P_m} \tau_{L_j(s)+x-cs_m} 1_A(cs_m) - \delta' \\ &\geqslant \mathbb{E}_{s_m \in P_m} 1_A(c \cdot s_m) - 3\delta' = \alpha - 3\delta'. \end{aligned}$$

We conclude that

$$|\mathbb{E}_{s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \tau_{L_j(s)+x-cs_m} 1_A(cs_m) - \alpha| = O(\delta') = O_{m,p,c}(\delta),$$

and similarly

$$|\mathbb{E}_{s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \tau_{L_j(s)+x-cs_m} 1_{c \cdot P_m}(cs_m) - 1| = O(\delta') = O_{m,p,c}(\delta)$$

for all $s_0 \in P_0, \dots, s_{m-1} \in P_{m-1}$ and $x \in d'' \cdot [N]$. It follows from these and (4.8) that

$$|\mathbb{E}_{s_0 \in P_0, \dots, s_m \in P_m} 1_{\mathrm{Int}_{\delta'}(P_m)}(s_m) \mathbb{E}_{x \in d'' \cdot [N]} \tau_{L_j(s)}(1_A - \alpha 1_{c \cdot P_m})(x)| = O_{m,p,c}(\delta).$$

Combining this with (4.7) and averaging, we see that there are elements $s_0 \in P_0, \dots, s_{m-1} \in P_{m-1}, s_m \in \mathrm{Int}_{\delta'}(P_m)$ such that

$$\mathbb{E}_{x \in d'' \cdot [N]} \tau_{L_j(s)}(1_A - \alpha 1_{c \cdot P_m})(x) > \tfrac{1}{2}\epsilon - O_{m,p,c}(\delta). \tag{4.9}$$

Since $s_m \in \mathrm{Int}_{\delta'}(P_m)$ the claim tells us that

$$\tau_{L_j(s)}(1_{c \cdot P_m})(x) = 1_{c \cdot P_m}(L_j(s) + x) = 1, \quad \text{for all } x \in d'' \cdot \{-N, \dots, N\},$$

and so $L_j(s) - d'' \cdot \{-N, \dots, N\} \subset c \cdot P_m \subset \mathbb{N}$. It follows from this that c divides $L_j(s)$ and of course c divides d'' (since $P'' \subset I_1(P_0) \subset I_1(c \cdot P_m) = c \cdot I_1(P_m)$ by Lemma 2.10).

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

From (4.9) we then have

$$\mathbb{E}_{x\in c\cdot(c^{-1}L_j(s)-(c^{-1}d'')\cdot[N])}1_A(x) > \alpha + \tfrac{1}{2}\epsilon - O_{m,p,c}(\delta).$$

We can take $\epsilon \in [\delta, O_{m,p,c}(\delta)]$ such that the right-hand side is at least $\alpha+\delta$, and we are in case (iii) of the lemma (with $P''' := c^{-1}L_j(s) - (c^{-1}d')\cdot[N]$ so $|P'''| = N \geqslant N'' = |P''|$ where $|P'''|$ is odd since it is prime, and

$$I_1(c\cdot P''') = I_1(d''\cdot[N]) \subset I_{O_{m,p,c}(1)}(P'') \subset I_{O_{m,p,c}(\delta^2)}(c\cdot P_m)$$

by Lemma 2.10. Again by Lemma 2.10 and the discussion in the section on 1.4 there is a monotone function $F' : \mathbb{N}^3 \to \mathbb{N}$ such that $I_1(c\cdot P'') \subset I_{F(m,p,c)\delta^2}(c\cdot P_m)$. And, again, we shall take $F_2 \geqslant F$.)*

In view of the above we may suppose that (4.7) does not happen for any $j \in \mathcal{D}_{p,c;m}$; we are in the main case.

We split the integrand in (4.2) into two factors:

$$\prod_{i\in\mathcal{U}_{m,p,c}} 1_A(L_i(s)) = \Bigg(\prod_{t=0}^{m-1}\prod_{i\in\mathcal{D}_{p,c;t}} 1_A(L_i(s))\Bigg)\cdot\Bigg(\prod_{i\in\mathcal{D}_{p,c;m}} 1_A(L_i(s))\Bigg).$$

The first term on the right-hand side is independent of s_m; we decompose the second through an arbitrary fixed total order on $\mathcal{D}^*_{p,c;m}$. For $i,j \in \mathcal{D}^*_{p,c;m}$ put

$$f_{j,i} := \begin{cases} 1_A, & \text{if } i<j,\\ 1_A - \alpha 1_{c\cdot P_m}, & \text{if } i=j,\\ \alpha 1_{c\cdot P_m}, & \text{if } i>j, \end{cases} \tag{4.10}$$

so for all $x \in \mathbb{Z}^{\mathcal{D}_{p,c;m}}$ we have

$$\begin{aligned}&\sum_{j\in\mathcal{D}^*_{p,c;m}} 1_A(x_{i^*(c,m)})\prod_{i\in\mathcal{D}^*_{p,c;m}} f_{j,i}(x_i)\\ &\quad= \prod_{i\in\mathcal{D}_{p,c;m}} 1_A(x_i) - \alpha^{|\mathcal{D}^*_{p,c;m}|}1_A(x_{i^*(c,m)})\prod_{i\in\mathcal{D}^*_{p,c;m}} 1_{c\cdot P_m}(x_i).\end{aligned}$$

It follows that

$$\begin{aligned}&Q_{m,p,c}(A;P_0,\dots,P_m)\\ &\quad- \alpha^{|\mathcal{D}^*_{p,c;m}|}\mathbb{E}_{s_0\in P_0,\dots,s_{m-1}\in P_{m-1}}\Bigg(\prod_{t=0}^{m-1}\prod_{i\in\mathcal{D}_{p,c;t}} 1_A(L_i(s))\Bigg)\\ &\quad\times\Bigg(\mathbb{E}_{s_m\in P_m}1_A(L_{i^*(c,m)}(s))\prod_{i\in\mathcal{D}^*_{p,c;m}} 1_{c\cdot P_m}(L_i(s))\Bigg)\end{aligned}$$

* This is much stronger than the conclusion offered in case (iii), but this is because this is the easy density increment mentioned at the end of the discussion before the proof of this lemma.

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

$$= \sum_{j\in\mathcal{D}^*_{p,c;m}} \mathbb{E}_{s_0\in P_0,\dots,s_m\in P_m}\Bigg(\prod_{t=0}^{m-1}\prod_{i\in\mathcal{D}_{p,c;t}} 1_A(L_i(s))\Bigg)$$
$$\times\Bigg(1_A(L_{i^*(c,m)}(s))\prod_{i\in\mathcal{D}^*_{p,c;m}} f_{j,i}(L_i(s))\Bigg). \tag{4.11}$$

On the other hand, for all $i\in\mathcal{D}_{p,c;m}$ and $s\in P_0\times\cdots\times P_{m-1}\times \mathrm{Int}_{\delta'}(P_m)$, we have from the claim (with $x=0$) that $L_i(s)\in c\cdot P_m$. Hence for all $s_0\in P_0,\dots,s_{m-1}\in P_{m-1}$ we have

$$\mathbb{E}_{s_m\in P_m} 1_A(L_{i^*(c,m)}(s))\prod_{i\in\mathcal{D}^*_{p,c;m}} 1_{c\cdot P_m}(L_i(s))$$
$$\geqslant \mathbb{E}_{s_m\in P_m} 1_A(cs_m)1_{c\cdot\mathrm{Int}_{\delta'}(P_m)}(cs_m)\prod_{i\in\mathcal{D}^*_{p,c;m}} 1_{c\cdot P_m}(L_i(s))$$
$$= \mathbb{E}_{s_m\in P_m} 1_A(cs_m)1_{c\cdot\mathrm{Int}_{\delta'}(P_m)}(cs_m) \geqslant \alpha - O_{m,p,c}(\delta).$$

Now,

$$\mathbb{E}_{s_m\in P_m} 1_A(L_{i^*(c,m)}(s))\prod_{i\in\mathcal{D}^*_{p,c;m}} 1_{c\cdot P_m}(L_i(s)) \leqslant \mathbb{E}_{s_m\in P_m} 1_A(L_{i^*(c,m)}(s)) = \alpha,$$

and so

$$\mathbb{E}_{s_m\in P_m} 1_A(L_{i^*(c,m)}(s))\prod_{i\in\mathcal{D}^*_{p,c;m}} 1_{c\cdot P_m}(L_i(s)) = \alpha + O_{m,p,c}(\delta).$$

Moreover,

$$Q_{m-1,p,c}(A;P_0,\dots,P_{m-1}) = \mathbb{E}_{s_0\in P_0,\dots,s_{m-1}\in P_{m-1}}\Bigg(\prod_{t=0}^{m-1}\prod_{i\in\mathcal{D}_{p,c;t}} 1_A(L_i(s))\Bigg)$$

since the sets $\mathcal{D}_{p,c;t}$ are disjoint over $0\leqslant t<m$. We conclude that the left-hand side of (4.11) is equal to

$$Q_{m,p,c}(A;P_0,\dots,P_m) - \alpha^{|\mathcal{D}_{p,c;m}|}Q_{m-1,p,c}(A;P_0,\dots,P_m) + O_{m,p,c}(\delta).$$

To estimate the right-hand side of (4.11), first note (by (4.4) and Lemma 2.10) that for $j\in\mathcal{D}^*_{p,c;m}$ the summand equals

$$\mathbb{E}_{s_0\in P_0,\dots,s_m\in P_m}\mathbb{E}_{s'_0,\dots,s'_m\in d''\cdot[N'']}\Bigg(\prod_{t=0}^{m-1}\prod_{i\in\mathcal{D}_{p,c;t}} 1_A(L_i(s+s'))\Bigg)$$
$$\times\Bigg(1_A(L_{i^*(c,m)}(s+s'))\prod_{i\in\mathcal{D}^*_{p,c;m}} f_{j,i}(L_i(s+s'))\Bigg)$$
$$+ O_m(\delta). \tag{4.12}$$

We shall look at the inner expectation of the first term here for which it will be useful to introduce some more notation. We shall use λ for dilation in the way defined in § 2.9,

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

and then for $s \in P_0 \times \cdots \times P_m$ and $y \in \mathbb{Z}^{m+1}$ put

$$g_i^{(s)}(y) = \begin{cases} \tau_{L_i(s)}(1_A) \circ \lambda_{d''}(y), & \text{if } i \in \mathcal{U}_{m-1,p,c} \text{ or } i = i^*(c,m), \\ \tau_{L_i(s)}(f_{j,i}) \circ \lambda_{d''}(y), & \text{if } i \in \mathcal{D}^*_{p,c;m}. \end{cases} \tag{4.13}$$

With this notation, the inner expectation in (4.12) equals

$$\begin{aligned} &\mathbb{E}_{y_0,\dots,y_m \in [N'']} \prod_{i \in \mathcal{U}_{m,p,c}} g_i^{(s)}(L_i(y)) \\ &\quad = \frac{1}{(N'')^{m+1}} \cdot \sum_{y \in \mathbb{Z}^{m+1}} \Big(\prod_{t=0}^m 1_{[N'']}(y_t) \Big) \Big(\prod_{t=0}^m \prod_{i \in \mathcal{D}^*_{p,c;t}} g_i^{(s)}(L_i(y)) \Big) \Big(\prod_{t=0}^m g^{(s)}_{i^*(c,t)}(L_{i^*(c,t)}(y)) \Big) \\ &\quad = \frac{1}{(N'')^{m+1}} \cdot \sum_{y \in \mathbb{Z}^{m+1}} \Big(\prod_{t=0}^m \prod_{i \in \mathcal{D}^*_{p,c;t}} g_i^{(s)}(L_i(y)) \Big) \Big(\prod_{t=0}^m g^{(s)}_{i^*(c,t)}|_{c\cdot[N'']}(L_{i^*(c,t)}(y)) \Big), \end{aligned}$$

since $L_{i^*(c,t)}(y) = cy_t$ for all $0 \leqslant t \leqslant m$. The notation is potentially a little confusing here: $g^{(s)}_{i^*(c,t)}|_{c\cdot[N'']}$ denotes the function $g^{(s)}_{i^*(c,t)}$ restricted to the set $c \cdot [N'']$.

For $x \in [N]$ write

$$h_i^{(s)}(x + N\mathbb{Z}) = \begin{cases} g_i^{(s)}(x), & \text{if } i \in \bigcup_{t=0}^m \mathcal{D}^*_{p,c;t}, \\ g_i^{(s)}|_{c\cdot[N'']}(x), & \text{if } i \in \{i^*(c,t) : 0 \leqslant t \leqslant m\}. \end{cases} \tag{4.14}$$

In view of this definition, for $x \in [N]^{m+1}$, the product

$$\prod_{i \in \mathcal{U}_{m,p,c}} h_i^{(s)}(L_i(x) + N\mathbb{Z})$$

is non-zero only if $x \in ((c \cdot [N''] + N\mathbb{Z})^{m+1}) \cap ([N]^{m+1})$. This set equals $(c \cdot [N''])^{m+1}$ since $N > cN''$. Now, if $x \in (c \cdot [N''])^{m+1}$ then $L_i(x) \in [N]$ since $N > (mp + c)N''$ and so $h_i^{(s)}(x + N\mathbb{Z}) = g_i^{(s)}(x)$. It follows that

$$\sum_{x \in [N]^{m+1}} \prod_{i \in \mathcal{U}_{m,p,c}} h_i^{(s)}(L_i(x) + N\mathbb{Z}) = (N'')^{m+1} \cdot \mathbb{E}_{x_0,\dots,x_m \in [N'']} \prod_{i \in \mathcal{U}_{m,p,c}} g_i^{(s)}(L_i(x)). \tag{4.15}$$

Apply (4.5) to the above and conclude that the right-hand side of (4.11) is at most

$$\sum_{j \in \mathcal{D}^*_{p,c;m}} \mathbb{E}_{s_0 \in P_0,\dots,s_m \in P_m} \left(\frac{N}{N''} \right)^{m+1} \|h_j^{(s)}\|_{U^k(\mathbb{Z}/N\mathbb{Z})} + O_{m,p,c}(\delta).$$

Let $\eta := (4\sqrt{\epsilon})^{1/F_1(k)} + \sqrt{\epsilon} + \delta'$ (where F_1 is as in Theorem 2.8) and suppose that

$$\sum_{j \in \mathcal{D}^*_{p,c;m}} \mathbb{E}_{s_0 \in P_0,\dots,s_m \in P_m} \|h_j^{(s)}\|_{U^k(\mathbb{Z}/N\mathbb{Z})} < \eta |\mathcal{D}^*_{p,c;m}|. \tag{4.16}$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

Then it follows that

$$|Q_{m,p,c}(A;P_0,\dots,P_m)-\alpha^{|\mathcal{D}_{p,c;m}|}Q_{m-1,p,c}(A;P_0,\dots,P_m)|=O_{m,p,c}(\delta)+O_{m,p,c}(\eta),$$

and we will find ourselves in case (iv) in view of the definition of η and choice of ϵ earlier. On the other hand, suppose that (4.16) does not hold, so that by averaging there is some $j\in\mathcal{D}^*_{p,c;m}$ such that

$$\mathbb{E}_{s_0\in P_0,\dots,s_m\in P_m}\|h_j^{(s)}\|_{U^k(\mathbb{Z}/N\mathbb{Z})}\geqslant\eta. \tag{4.17}$$

Since (4.7) does not happen, writing

$$\mathcal{S}:=\{s\in P_0\times\cdots\times P_m : s_m\in \mathrm{Int}_{\delta'}(P_m) \text{ and } |\mathbb{E}_{x\in d''\cdot[N]}\tau_{L_j(s)}(1_A-\alpha 1_{c\cdot P_m})(x)|>\sqrt{\epsilon}\},$$

we have

$$\mathbb{E}_{s_0\in P_0,\dots,s_m\in P_m}1_{\mathcal{S}}(s)\sqrt{\epsilon}\leqslant\epsilon.$$

By (4.17), Lemma 2.10, the value of η and the triangle inequality, we see that

$$\mathbb{E}_{s_0\in P_0,\dots,s_m\in P_m}1_{\mathrm{Int}_{\delta'}(P_m)}(s_m)1_{\mathcal{S}^c}(s)\|h_j^{(s)}\|_{U^k(\mathbb{Z}/N\mathbb{Z})}\geqslant\eta-\sqrt{\epsilon}-\delta'\geqslant(4\sqrt{\epsilon})^{1/F_1(k)}.$$

By averaging, there is some $s\in(P_0\times\cdots\times P_{m-1}\times\mathrm{Int}_{\delta'}(P_m))\setminus\mathcal{S}$ such that

$$\|h_j^{(s)}\|_{U^k(\mathbb{Z}/N\mathbb{Z})}\geqslant(4\sqrt{\epsilon})^{1/F_1(k)}.$$

By Theorem 2.8 (applicable since $4\sqrt{\epsilon}\leqslant 2^{-F_1(k)}$ as otherwise we are in case (ii) of the lemma since $\epsilon\geqslant\delta$), there is a partition of $[N]$ into arithmetic progressions $Q_1,\dots,Q_M$ of average size at least $N^{\epsilon^{F_1(k)}}$ such that

$$\sum_{l=1}^M\Big|\sum_{x\in Q_l}\tau_{L_j(s)}(f_{j,j})\circ\lambda_{d''}(x)\Big|=\sum_{l=1}^M\Big|\sum_{x\in Q_l}g_j^{(s)}(x)\Big|=\sum_{l=1}^M\Big|\sum_{x\in Q_l}h_j^{(s)}(x+N\mathbb{Z})\Big|\geqslant 4\sqrt{\epsilon}N. \tag{4.18}$$

Of course

$$\begin{aligned}\sum_{l=1}^M\sum_{x\in Q_l}\tau_{L_j(s)}(f_{j,j})\circ\lambda_{d''}(x)&=\sum_{x\in d''\cdot[N]}\tau_{L_j(s)}(f_{j,j})(x)\\&=\sum_{x\in d''\cdot[N]}\tau_{L_j(s)}(1_A-\alpha 1_{c\cdot P_m})(x),\end{aligned}$$

and so (since $s\notin\mathcal{S}$)

$$\Big|\sum_{l=1}^M\sum_{x\in Q_l}\tau_{L_j(s)}(f_{j,j})\circ\lambda_{d''}(x)\Big|\leqslant\sqrt{\epsilon}N.$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

Moreover, since the average size of Q_l is at least $N^{\epsilon^{F_1(k)}}$, we have

$$\sum_{\substack{1\leqslant l\leqslant M\\ |Q_l|\leqslant N^{\epsilon^{F_1(k)}/2}}}\left|\sum_{x\in Q_l}\tau_{L_j(s)}(f_{j,j})\circ\lambda_{d''}(x)\right|\leqslant N^{\epsilon^{F_1(k)/2}}M\leqslant N^{-\epsilon^{F_1(k)/2}}N\leqslant\sqrt{\epsilon}N,$$

since we may assume that $N^{-\epsilon^{F_1(k)/2}}\leqslant\sqrt{\epsilon}$ (or else we are in case (i) of the lemma since $\epsilon\geqslant\delta$ and $N\geqslant N''=|P''|$). We may assume that all the Q_l are of odd size by removing at most one point from each at a cost of at most M in (4.18). (Again $M\leqslant\sqrt{\epsilon}N$ or else we are in case (i) of the lemma.) It follows by the triangle inequality and averaging that there is some $1\leqslant l\leqslant M$ with

$$\sum_{x\in Q_l}\tau_{L_j(s)}(f_{j,j})\circ\lambda_{d''}(x)\geqslant\tfrac{1}{2}\sqrt{\epsilon}|Q_l|\text{ and }|Q_l|>N^{\epsilon^{F_1(k)}/2}.$$

Rewriting the first expression, we get that

$$\mathbb{E}_{x\in d''\cdot Q_l}1_A(L_j(s)+x)\geqslant\alpha\mathbb{E}_{x\in d''\cdot Q_l}1_{c\cdot P_m}(L_j(s)+x)+\tfrac{1}{2}\sqrt{\epsilon}=\alpha+\tfrac{1}{2}\sqrt{\epsilon}$$

by the claim. Firstly, $\frac{1}{2}\sqrt{\epsilon}\geqslant 2\delta$ (or else we are in case (ii) of the lemma in view of the fact that $\epsilon\geqslant\delta$). Secondly, as noted after (4.9), $L_j(s)-d''\cdot\{-N,\dots,N\}\subset c\cdot P_m\subset\mathbb{N}$, c divides $L_j(s)$ and c divides d'', and so

$$\mathbb{E}_{x\in c^{-1}L_j(s)-(c^{-1}d'')\cdot Q_l}1_A(x)\geqslant\alpha+\delta,$$

and, putting $P'''':=c^{-1}L_j(s)-(c^{-1}d'')\cdot Q_l\subset\mathbb{N}$, we are in case (iii) and hence are done. (Indeed, $|P''''|=|Q_l|\geqslant N^{\epsilon^{F_1(k)}/2}\geqslant N^{\delta^{F_1(k)}/2}$, and $|Q_l|$ is odd. As before,

$$I_1(c\cdot P'''')=I_1(d''\cdot[N])\subset I_{O_{m,p,c}(1)}(P'')\subset I_{O_{m,p,c}(\delta^2)}(c\cdot P_m)$$

by Lemma 2.10. Again by Lemma 2.10 and the discussion in the section on —1.4 there is a monotone function $F':\mathbb{N}^3\to\mathbb{N}$ such that $I_1(c\cdot P'''')\subset I_{F(m,p,c)\delta^2}(c\cdot P_m)$. And, again, we shall take $F_2\geqslant F$.)

The lemma is proved. □

Our main result is the following theorem.

Theorem 4.2. *Suppose that $m,p,c,r\in\mathbb{N}$, $\delta\in(0,1]$; $P\subset\mathbb{N}$ is an arithmetic progression of odd length N; and $\mathcal{C}$ is an r-colouring of $I_1(P)\cap\mathbb{N}$. Then at least one of the following statements holds.*

(i) $N\leqslant\exp(\exp(\delta^{-O_{m,p,c}(1)}))$.

(ii) $\delta\geqslant(2r)^{-O_{m,p,c}(1)}$.

(iii) *There are progressions $P_0,\dots,P_m\subset I_\delta(c\cdot P)\cap\mathbb{N}$ of odd length with*

$$P_i\subset I_1(P_{i+1}),\quad\text{for all }0\leqslant i\leqslant m-1\text{ and }|P_0|\geqslant N^{\exp(-\delta^{-O_{m,p,c}(1)})},$$

and some $C\in\mathcal{C}$ such that

$$Q_{m,p,c}(C;P_0,\dots,P_m)\geqslant(2r)^{-O_{m,p,c}(1)}.$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

We shall proceed by a double induction. The outer induction will be on m and the inner is a density increment argument.

Lemma 4.3 (Iteration lemma). *Suppose that Theorem 4.2 holds for some $m \in \mathbb{N}_0$, which we express as follows.*

There are monotone functions $F^{(m)} : \mathbb{N}^2 \times (0,1] \to \mathbb{N}$, $\eta_0^{(m)} : \mathbb{N}^3 \to (0,1]$ and $\eta_1^{(m)} : \mathbb{N}^2 \times (0,1] \to (0,1]$ such that, for any $p, c, r \in \mathbb{N}$, $\delta \in (0,1]$, $P \subset \mathbb{N}$ an arithmetic progressions of odd length N, and r-colouring $\mathcal{C}$ of $I_1(P) \cap \mathbb{N}$, at least one of the following holds.

(i) $N \leqslant F^{(m)}(p, c, \delta)$,

(ii) $\delta \geqslant \eta_0^{(m)}(p, c, r)$

(iii) *There are progressions $P_0, \dots, P_m \subset I_\delta(c \cdot P) \cap \mathbb{N}$ of odd length with*

$$P_i \subset I_1(P_{i+1}), \quad \text{for all } 0 \leqslant i \leqslant m-1, \quad \text{and} \quad |P_0| \geqslant N^{\eta_1^{(m)}(p,c,\delta)},$$

and some $C \in \mathcal{C}$ such that

$$Q_{m,p,c}(C; P_0, \dots, P_m) \geqslant \eta_0^{(m)}(p, c, r).$$

Then for any $p, c, r \in \mathbb{N}$, $\delta \in (0,1]$, $P \subset \mathbb{N}$ an arithmetic progressions of odd length N, and r-colouring $\mathcal{C}$ of $I_1(P) \cap \mathbb{N}$ at least one of the following holds.

(i)
$$N \leqslant \min\{F^{(m)}(p, c, \delta), \exp(\delta^{-O_{m,p,c}(1)} \eta_1^{(m)}(p, c, \delta)^{-1})\}.$$

(ii)
$$\delta \geqslant \left(\frac{\eta_0^{(m)}(p, c, r)}{2r}\right)^{O_{m,p,c}(1)}.$$

(iii) *There is a progression P''' of odd length with*

$$I_1(P''') \subset I_1(P) \quad \text{and} \quad |P'''| \geqslant N^{\delta^{O_{m,p,c}(1)} \eta_1^{(m)}(p,c,\delta)}$$

such that

$$\sum_{C \in \mathcal{C}} \max_{y : y + P''' \subset \mathbb{N}} \mathbb{E}_{x \in c \cdot (y + P''')} 1_C(x) \geqslant \sum_{C \in \mathcal{C}} \max_{y : y + P \subset \mathbb{N}} \mathbb{E}_{x \in c \cdot (y+P)} 1_C(x) + \tfrac{1}{2}\delta.$$

(iv) *There are progressions $P_0, \dots, P_{m+1} \subset I_\delta(c \cdot P) \cap \mathbb{N}$ of odd length with*

$$P_i \subset I_1(P_{i+1}), \quad \text{for all } 0 \leqslant i \leqslant m, \quad \text{and} \quad |P_0| \geqslant N^{\eta_1^{(m)}(p,c,\delta)},$$

and some $C \in \mathcal{C}$ such that

$$Q_{m+1,p,c}(C; P_0, \dots, P_{m+1}) \geqslant \left(\frac{\eta_0^{(m)}(p, c, r)}{2}\right)^{O_{m,p,c}(1)}.$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

Proof. Apply the content of the box to P to get that either we are in case (i) or (ii) of the hypothesis and so in case (i) and (ii) respectively of the present lemma, or else there are progressions $P_0, \dots, P_m \subset I_\delta(c \cdot P) \cap \mathbb{N}$ with $P_i \subset I_1(P_{i+1})$ for all $0 \leqslant i \leqslant m-1$, and some $C \in \mathcal{C}$ with

$$Q_{m,p,c}(C; P_0, \dots, P_m) \geqslant \eta_0^{(m)}(p,c,r) \quad \text{and} \quad |P_0| \geqslant N^{\eta_1^{(m)}(p,c,\delta)}.$$

Let $y_0 \in \mathbb{Z}$ be such that $y_0 + P \subset \mathbb{N}$ and

$$\mathbb{E}_{x \in c\cdot(y_0+P)} 1_C(x) = \max_{y : y+P \subset \mathbb{N}} \mathbb{E}_{x \in c\cdot(y+P)} 1_C(x), \tag{4.19}$$

and let $P_{m+1} = y_0 + P$. Then $P_0, \dots, P_{m+1}$ are arithmetic progressions of odd length. Furthermore, by Lemma 2.10 (and assuming we are not in case (ii)), we have

$$P_i \subset I_\delta(c \cdot P_m) \subset I_\delta(c \cdot I_\delta(c \cdot P)) \subset I_\delta(c \cdot P_{m+1}), \quad \text{for all } 0 \leqslant i \leqslant m-1,$$

and

$$P_m \subset I_\delta(c \cdot P) = I_\delta(c \cdot P_{m+1}).$$

It follows that we can apply Lemma 4.1 with parameters $m+1, p, c \in \mathbb{N}$ and $\delta \in (0,1]$, set C, and odd-length arithmetic progressions $P_0, \dots, P_{m+1} \subset \mathbb{Z}$, and

$$P'' := I_\delta(P_0) \subset I_\delta(P_i), \quad \text{for all } 0 \leqslant i \leqslant m+1.$$

We have four cases.

(i) Then

$$\delta N^{\eta_1^{(m)}(p,c,\delta)} \leqslant |I_\delta(P_0)| = |P''| \leqslant \exp(\delta^{-F_2(m+1,p,c)}),$$

and we are in case (i) of this lemma.

(ii) Then

$$\delta \geqslant \eta_0(m+1,p,c)$$

and we are in case (ii) of the lemma.

(iii) Then there is an arithmetic progression $P''' \subset \mathbb{N}$ of odd length with $I_1(c \cdot P''') \subset I_{F_2(m+1,p,c)\delta^2}(c \cdot P)$ such that

$$|P'''| \geqslant N^{\eta_1^{(m)}(p,c,\delta)\delta^{O_{m,p,c}(1)}} \quad \text{and} \quad \mathbb{E}_{x \in c\cdot P'''} 1_C(x) \geqslant \mathbb{E}_{x \in c\cdot P_{m+1}} 1_C(x) + \delta.$$

(Then either we are in case (ii) of the lemma or else $I_1(P''') \subset I_1(P)$.) In view of the choice of y_0 (4.19) and the definition of P_{m+1}, the second expression tells us

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

that

$$\max_{w:w+P'''\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P''')}1_C(x)\geqslant\mathbb{E}_{x\in c\cdot P'''}1_C(x)\geqslant\max_{y:y+P\subset\mathbb{N}}\mathbb{E}_{x\in c\cdot(y+P)}1_C(x)+\delta.$$

For (the other) $C'\in\mathcal{C}$ and $y\in\mathbb{Z}$ such that $y+P\subset\mathbb{N}$ we have, by Lemma 2.10, that

$$\begin{aligned}\max_{w:w+P'''\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P''')}1_{C'}(z)&\geqslant\mathbb{E}_{x\in c\cdot P}\mathbb{E}_{z\in c\cdot(x+y+P''')}1_{C'}(z)\\&=\mathbb{E}_{u\in c\cdot(y+P)}\mathbb{E}_{z\in c\cdot P'''}\tau_z(1_{C'})(u)\\&\geqslant\mathbb{E}_{x\in c\cdot(y+P)}1_{C'}(x)-F_2(m+1,p,c)\delta^2.\end{aligned}$$

Taking the maximum over y such that $y+P\subset\mathbb{N}$ and summing, it follows that

$$\begin{aligned}&\sum_{C'\in\mathcal{C}}\max_{w:w+P'''\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P''')}1_{C'}(z)\\&\quad\geqslant\sum_{C'\in\mathcal{C}}\max_{w:w+P\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P)}1_{C'}(z)+\delta-(r-1)F_2(m+1),p,c)\delta^2.\end{aligned}$$

So we are either in case (ii) of the lemma or in case (iii) of the lemma.

(iv) Then

$$\begin{aligned}&|Q_{m+1,p,c}(C;P_0,\dots,P_{m+1})\\&\quad-\alpha^{|\mathcal{D}_{m+1,p,c}|}Q_{m,p,c}(C;P_0,\dots,P_m)|\leqslant F_2(m+1,p,c)\delta^{\eta_0(m+1,p,c)},\end{aligned}\tag{4.20}$$

where $\alpha:=\mathbb{E}_{x\in c\cdot P_{m+1}}1_C(x)$. First, suppose that

$$\max_{y:y+P_m\subset\mathbb{N}}\mathbb{E}_{x\in c\cdot(y+P_m)}1_C(x)>\alpha+((r-1)F_2(m+1,p,c)+1)\delta\tag{4.21}$$

and so, by (4.19) and the definition of P_{m+1}, we have

$$\begin{aligned}\max_{y:y+P_m\subset\mathbb{N}}\mathbb{E}_{x\in c\cdot(y+P_m)}1_C(x)>&\max_{y:y+P\subset\mathbb{N}}\mathbb{E}_{x\in y+P}1_C(x)\\&+((r-1)F_2(m+1,p,c)+1)\delta.\end{aligned}$$

For the other $C'\in\mathcal{C}$, we use that $P_m\subset I_\delta(c\cdot P_{m+1})=I_\delta(c\cdot P)$ and Lemma 2.10 to give that for any $y+P\subset\mathbb{N}$ we have

$$\begin{aligned}\max_{w:w+P_m\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P_m)}1_{C'}(z)&\geqslant\mathbb{E}_{x\in c\cdot(y+P)}\mathbb{E}_{z\in c\cdot(x+y+P_m)}1_{C'}(z)\\&=\mathbb{E}_{u\in c\cdot(y+P)}\mathbb{E}_{z\in c\cdot P'''}\tau_z(1_{C'})(u)\\&\geqslant\mathbb{E}_{x\in c\cdot(y+P)}1_{C'}(x)-F_2(m+1,p,c)\delta.\end{aligned}$$

Taking the maximum over y such that $y+P\subset\mathbb{N}$ and summing, it follows that

$$\begin{aligned}&\sum_{C'\in\mathcal{C}}\max_{w:w+P_m\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P_m)}1_{C'}(z)\\&\quad\geqslant\sum_{C'\in\mathcal{C}}\max_{w:w+P\subset\mathbb{N}}\mathbb{E}_{z\in c\cdot(w+P)}1_{C'}(z)+\delta.\end{aligned}$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

We are in case (iii). (We should also note that we are in case (ii) of the lemma or else $I_1(P_m) \subset I_1(P)$.) We conclude that (4.21) does not hold and so

$$\begin{aligned}\alpha + ((r-1)F_2(m+1,p,c)+1)\delta &\geqslant \max_{y: y+P_m \subset \mathbb{N}} \mathbb{E}_{x \in c\cdot(y+P_m)} 1_C(x)\\ &\geqslant \mathbb{E}_{x\in c\cdot P_m} 1_C(x) \geqslant Q_{m,p,c}(C; P_0, \dots, P_m) \geqslant \eta_0^{(m)}(p,c,r).\end{aligned}$$

Either

$$\delta \geqslant \frac{\eta_0^{(m)}(p,c,r)}{2((r-1)F_2((m+1),p,c)+1)}$$

and we are in case (ii) of the lemma; or from (4.20) we have

$$F_2(m+1,p,c)\delta^{\eta_0(m+1,p,c)} \geqslant \tfrac{1}{2}(\tfrac{1}{2}\eta_0^{(m)}(p,c,r))^{|\mathcal{D}_{m+1,p,c}|+1}$$

and we are in case (ii) of the lemma; or we are in case (iv) of the lemma.

The lemma is proved. □

Proof of Theorem 4.2. We proceed by induction on m to show that the content of the box in Lemma 4.3 holds. This gives the theorem. The result holds for $m = 0$ and it is convenient to use that as the base case. To see this, take $P_0 := I_\delta(c\cdot P) \cap \mathbb{N}$. If $\delta N < 2$ then we are in case (i) of the box. If not then $|P_0| = \Omega(\delta N)$, and again either we are in case (i), or $|P_0|$ satisfies the required lower bound. Finally, we are in case (ii) or else $I_\delta(c^2\cdot P) \subset I_1(P)$ and

$$\sum_{C\in\mathcal{C}} Q_{0,p,c}(C;P_0) = \sum_{C\in\mathcal{C}} \mathbb{E}_{s_0\in P_0} 1_C(cs_0) \geqslant \mathbb{E}_{s_0\in P_0} 1_{I_1(P)\cap\mathbb{N}}(cs_0) = 1.$$

The result follows by averaging. Now, suppose we have proved that the content of the box holds for some m.

We proceed iteratively defining progressions $P^{(0)}, P^{(1)}, \dots$ with $I_1(P^{(j+1)}) \subset I_1(P^{(j)})$ for all $j \geqslant 0$. Begin with $P^{(0)} := I_1(P)\cap\mathbb{N}$ and define

$$\mu_j := \sum_{C\in\mathcal{C}} \max_{y: y+P^{(j)}\subset\mathbb{N}} \mathbb{E}_{z\in y+P^{(j)}} 1_C(z).$$

By hypothesis, we have $\mu_0 \geqslant 1$ and we also have $\mu_j \leqslant r$ for all j. At stage $j \in \mathbb{N}_0$ we apply Lemma 4.3 to $P^{(j)}$ and unless we are in case (iii) we terminate. If we are in case (iii) then we let $P^{(j+1)}$ be the progression given, which has

$$I_1(P^{(j+1)}) \subset I_1(P^{(j)}), |P^{(j+1)}| \geqslant |P^{(j)}|^{\delta^{O_{m,p,c}(1)}\eta_1^{(m)}(p,c,\delta)} \quad \text{and} \quad \mu_{j+1} \geqslant \mu_j + \tfrac{1}{2}\delta.$$

In view of the last fact this iteration can proceed for at most $2\delta^{-1}$ steps before terminating. When it terminates we have either

$$N^{\delta^{-O_{m,p,c}(\delta^{-1})}(\eta_1^{(m)}(p,c,\delta))^{2\delta^{-1}}} \leqslant \min\{F^{(m)}(p,c,\delta), \exp(\delta^{-O_{m,p,c}(1)}\eta_1^{(m)}(p,c,\delta)^{-1})\},$$

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

or

$$\delta \geqslant \left(\frac{\eta_0^{(m)}(p,c,r)}{2r}\right)^{O_{m,p,c}(1)},$$

or there is some $C \in \mathcal{C}$ such that

$$Q_{m+1,p,c}(C; P_0, \dots, P_{m+1}) \geqslant \left(\frac{\eta_0^{(m)}(p,c,r)}{2}\right)^{O_{m,p,c}(1)} \quad \text{and}$$

$$|P_0| \geqslant N^{\delta^{-O_{m,p,c}(\delta^{-1})}(\eta_1^{(m)}(p,c,\delta))^{2\delta^{-1}}}.$$

It follows that we can take

$$F^{(m+1)}(p,c,\delta) \leqslant (2F^{(m)}(p,c,\delta))^{\eta_1^{(m)}(p,c,\delta)^{-O_{m,p,c}(\delta^{-1})}},$$

$$\eta_1^{(m+1)}(p,c,\delta) \geqslant \left(\frac{\eta_1^{(m)}(p,c,\delta)}{2}\right)^{O_{m,p,c}(\delta^{-1})}$$

and

$$\eta_0^{(m+1)}(p,c,r) \geqslant \left(\frac{\eta_0^{(m)}(p,c,r)}{2r}\right)^{O_{m,p,c}(1)}.$$

These recursions give the claimed bounds. □

Proof of Theorem 2.3. We apply Theorem 4.2 with $P = [N]$ (or $P = [N-1]$ if N is even) with $1/2c \geqslant \delta = r^{-O_{m,p,c}(1)}$ such that case (ii) never holds. If the colouring contains no (m,p,c)-set then $Q_{m,p,c}(C; P_0, \dots, P_m) = 0$ for any $P_0, \dots, P_m$ of form described in case (iii) and so that does not happen. We conclude that N is bounded in a way that yields the result. □

As a final remark, although we have made no effort to track the m, p, and c dependencies they should also not be too bad, given the known bounds in Theorem 2.7 and Theorem 2.8.

Acknowledgements. The author should like to thank David Conlon and Julia Wolf for useful conversations, Jonathan Chapman and Sean Prendiville for providing an early copy of [6], and the referees for careful reading of the paper including the example of how the main argument works.

References

1. H. L. Abbott and L. Moser, Sum-free sets of integers, *Acta Arith.* **11** (1966), 393–396.
2. T. Ahmed and D. J. Schaal, On generalized Schur numbers, *Exp. Math.* **25**(2) (2016), 213–218.
3. A. Beutelspacher and W. Brestovansky, Generalized Schur numbers, in *Combinatorial theory (Schloss Rauischholzhausen, 1982)* (eds D. Jungnickel and K. Vedder), pp. 30–38, Lecture Notes in Mathematics, Volume 969 (Springer, Berlin, 1982).
4. J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.* **9**(5) (1999), 968–984.
5. J. Chapman, Partition regularity and multiplicatively syndetic sets, preprints (arXiv.org/abs/1902.01149, 2019).

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

6. J. Chapman and S. Prendiville, On the Ramsey number of the Brauer configuration, preprint (arXiv.org/abs/1904.07567v1, 2019).
7. S. Chow, S. Lindqvist and S. Prendiville, Rado's criterion over squares and higher powers, preprint (arXiv.org/abs/1806.05002, 2018).
8. K. Cwalina and T. Schoen, Tight bounds on additive Ramsey-type numbers, *J. Lond. Math. Soc.* **96**(3) (2017), 601–620.
9. W. Deuber, Partitionen und lineare Gleichungssysteme, *Math. Z.* **133** (1973), 109–123.
10. P. Frankl, R. L. Graham and V. Rödl, Quantitative theorems for regular systems of equations, *J. Combin. Theory Ser. A* **47**(2) (1988), 246–261.
11. W. Gasarch, R. Moriarty and N. Tumma, New upper and lower bounds on the Rado numbers, preprint (arXiv.org/abs/1206.4885, 2012).
12. W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, *Geom. Funct. Anal.* **8**(3) (1998), 529–551.
13. W. T. Gowers, A new proof of Szemerédi's theorem, *Geom. Funct. Anal.* **11**(3) (2001), 465–588.
14. W. T. Gowers and J. Wolf, The true complexity of a system of linear equations, *Proc. Lond. Math. Soc. (3)* **100**(1) (2010), 155–176.
15. R. L. Graham and B. L. Rothschild, Ramsey's theorem for n-parameter sets, *Trans. Amer. Math. Soc.* **159** (1971), 257–292.
16. R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey theory*, 2nd edn, Wiley-Interscience Series in Discrete Mathematics and Optimization (John Wiley & Sons, New York, 1990).
17. B. J. Green, Finite field models in additive combinatorics, in *Surveys in combinatorics 2005*, pp. 1–27, London Mathematical Society Lecture Note Series, Volume 327 (Cambridge University Press, Cambridge, 2005).
18. B. J. Green and S. Lindqvist, Monochromatic solutions to $x + y = z^2$, *Can. J. Math.* **71**(3) (2019), 579–605.
19. B. J. Green and T. Sanders, Monochromatic sums and products, *Discrete Anal.* (5) (arXiv:1510.08733, 2016). doi:10.19086/da.613.
20. B. J. Green and T. C. Tao, An arithmetic regularity lemma, an associated counting lemma, and applications, in *An irregular mind*, pp. 261–334, Bolyai Society Mathematical Studies, Volume 21 (János Bolyai Mathematical Society, Budapest, 2010).
21. B. J. Green and T. C. Tao, Linear equations in primes, *Ann. Math. (2)* **171**(3) (2010), 1753–1850.
22. D. S. Gunderson, On Deuber's partition theorem for (m, p, c)-sets, *Ars Combin.* **63** (2002), 15–31.
23. S. Gupta, J. Thulasi Rangan and A. Tripathi, The two-colour Rado number for the equation $ax + by = (a + b)z$, *Ann. Comb.* **19**(2) (2015), 269–291.
24. H. Hatami and S. Lovett, Higher-order Fourier analysis of $\mathbb{F}_p^n$ and the complexity of systems of linear forms, *Geom. Funct. Anal.* **21**(6) (2011), 1331–1357.
25. N. Hindman and I. B. Leader, Nonconstant monochromatic solutions to systems of linear equations, in *Topics in discrete mathematics*, pp. 145–154, Algorithms and Combinatorics, Volume 26 (Springer, Berlin, 2006).
26. B. M. Landman and A. Robertson, *Ramsey theory on the integers*, 2nd edn, Student Mathematical Library, Volume 73 (American Mathematical Society, Providence, RI, 2014).
27. T. H. Lê, Partition regularity and the primes, *C. R. Math.* **350**(9) (2012), 439–441.
28. H. Li and H. Pan, A Schur-type addition theorem for primes, *J. Number Theory* **132**(1) (2012), 117–126.
29. F. Manners, Good bounds in certain systems of true complexity one, *Discrete Anal.* (21) (arXiv:1705.06801, 2018). doi:10.19086/da.6814.

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press

30. S. Prendiville, Quantitative bounds in the polynomial Szemerédi theorem: the homogeneous case, *Discrete Anal.* **5** (2017), 1–34.
31. R. Rado, Studien zur Kombinatorik, *Math. Z.* **36** (1933), 424–480.
32. R. Rado, Some partition theorems, in *Combinatorial theory and its applications, III*, Proc. Colloq., Balatonfüred, 1969, pp. 929–936 (János Bolyai Mathematical Society, 1970).
33. A. Robertson and K. Myers, Some two color, four variable Rado numbers, *Adv. Appl. Math.* **41**(2) (2008), 214–226.
34. J. H. Sanders, *A generalization of Schur's theorem*. PhD thesis, Yale University, 1968. http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:6908427.
35. D. Saracino, The 2-color Rado number of $x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_k$, *Ars Combin.* **129** (2016), 315–321.
36. I. Schur, Über die Kongruenz $x^m + y^m \equiv z^m (\mathrm{mod}.p)$, *Jahresber. Dtsch. Math.-Ver.* **25** (1916), 114–117.
37. I. D. Shkredov, Fourier analysis in combinatorial number theory, *Russ. Math. Surv.* **65**(3) (2010), 513–567.
38. T. C. Tao, *Higher order Fourier analysis*, Graduate Studies in Mathematics, Volume 142 (American Mathematical Society, Providence, RI, 2012).
39. T. C. Tao and V. H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, Volume 105 (Cambridge University Press, Cambridge, 2006).
40. B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wiskd., II. Ser.* **15** (1927), 212–216.
41. A. Walker, Gowers norms control Diophantine inequalities, preprint (arXiv.org/abs/1703.00885), 2018.
42. J. Wolf, Finite field models in arithmetic combinatorics – ten years on, *Finite Fields Appl.* **32** (2015), 233–274.

https://doi.org/10.1017/S0013091520000048 Published online by Cambridge University Press