*Mathematika* 61 (2015) 14–48 © University College London
doi:10.1112/S0025579314000163

# MATRIX PROGRESSIONS IN MULTIDIMENSIONAL SETS OF INTEGERS

SEAN PRENDIVILLE

*Abstract.* We obtain density estimates for subsets of the $n$-dimensional integer lattice lacking four-term matrix progressions. As a consequence, we show that a subset of the grid $\{1, 2, \dots, N\}^2$ lacking four corners in a square has size at most $CN^2(\log\log N)^{-c}$. Our proofs involve the density increment method of Roth [*J. London Math. Soc.* **28** (1953), 104–109] and Gowers [*Geom. Funct. Anal.* **11**(3) (2001), 465–588], together with the $U^3$-inverse theorem of Green and Tao [*Proc. Edinb. Math. Soc.* (2) **51**(1) (2008), 73–153].

§1. *Introduction.* One consequence of Roth's celebrated work [**9**, **10**] on linear equations in dense sets of integers is that for any $t_1, t_2 \in \mathbb{Z}$ there exists an absolute constant $C = C(t_1, t_2)$ such that if $A \subset \{1, \dots, N\}$ lacks configurations of the form $\{x, x + t_1 d, x + t_2 d\}$ with $d \neq 0$, then we have the bound $|A| \leqslant CN(\log\log N)^{-1}$. Gowers [**3**] later extended this result to show that a set $A \subset \{1, \dots, N\}$ lacking four elements in arithmetic progression satisfies $|A| \leqslant CN(\log\log N)^{-c}$. The purpose of this paper is to obtain the following multidimensional analogues of these results.

THEOREM 1.1. *Let $T_1$, $T_2$ be non-singular $n \times n$ rational matrices and let $K$ denote a finite union of proper subspaces of $\mathbb{R}^n$. There exists an absolute constant $C(T_1, T_2, K)$ such that if $A \subset \{1, 2, \dots, N\}^n$ does not contain any configurations of the form $\{\mathbf{x}, \mathbf{x} + T_1\mathbf{d}, \mathbf{x} + T_2\mathbf{d}\}$ with $\mathbf{d} \in \mathbb{Z}^n \setminus K$, then $A$ satisfies*

$$|A| \leqslant \frac{CN^n}{\log\log N}. \tag{1.1}$$

THEOREM 1.2. *Suppose, in addition to the assumptions of Theorem 1.1, that $T_3$ is an $n \times n$ non-singular rational matrix with both $T_3 - T_1$ and $T_3 - T_2$ non-singular. Then there exists a constant $C = C(T_1, T_2, T_3, K)$ such that any subset $A \subset \{1, 2, \dots, N\}^n$ lacking configurations of the form $\{\mathbf{x}, \mathbf{x} + T_1\mathbf{d}, \mathbf{x} + T_2\mathbf{d}, \mathbf{x} + T_3\mathbf{d}\}$ with $\mathbf{d} \in \mathbb{Z}^n \setminus K$, satisfies the bound*

$$|A| \leqslant \frac{CN^n}{(\log\log N)^c}, \tag{1.2}$$

*where $c = 2^{-26}$.*

Received 13 June 2013, published online 13 August 2014.
MSC (2010): 11B30 (primary).

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

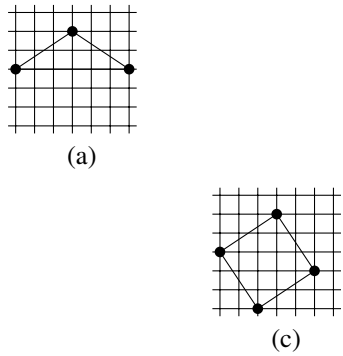



Figure 1: We obtain density bounds for sets lacking the following configurations: (a) axis-aligned isosceles triangles; (b) right-angled isosceles triangles; (c) squares.

Theorems 1.1 and 1.2 both give a sufficient cardinality above which a subset of $\{1, \dots, N\}^n$ contains non-trivial three or four-term matrix progressions. Examples of configurations we are able to guarantee using Theorem 1.1 include isosceles triangles with one side parallel to the $x$-axis (Figure 1(a)), obtained by taking

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad K = (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}),$$

as well as right-angled isosceles triangles (Figure 1(b)), obtained by taking

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad K = \{(0, 0)\}.$$

Similarly, in Theorem 1.2 we can set

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad T_3 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \qquad K = \{(0, 0)\}$$

and thereby deduce the following corollary.

COROLLARY 1.3. *If a set $A \subset \{1, \dots, N\}^2$ lacks four points which together form the corners of a square (Figure 1(c)) then $|A| \ll N^2(\log\log N)^{-c}$ with $c = 2^{-26}$.*

We note that Graham [**1**, Problem 3.7] has conjectured that if a set $A \subset \mathbb{Z}^2$ lacks four points which together form the corners of an *axis-aligned* square, then the sum $\sum_{(x,y)\in A} (1/(x^2 + y^2))$ converges. Such a result is very much out of reach of the methods employed in this paper.

One can deduce a weaker bound in Theorem 1.1 from the quantitative 'corners' theorem of Shkredov [**13**, **14**], even when $T_1$ and $T_2$ are *singular* matrices (see Appendix B for details). The best current bounds in Shkredov's theorem imply a bound of the form (1.1) with the denominator $\log\log N$ replaced

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

by $(\log\log N)^{1/22}$. We note that a simple adaptation of our method combined with the argument of Szemerédi and Heath-Brown [8, 15] yields a denominator of the form $(\log N)^c$. Moreover, it may be possible to synthesise our methods with the spectacular machinery of Sanders [11] to obtain a denominator of the form $(\log N)^{1-o(1)}$. Rather than pursue these objectives, we instead focus on four-point configurations (Theorem 1.2).

Since, at present, there are no quantitative bounds for subsets of $\mathbb{Z}^3$ lacking three-dimensional corners $\mathbf{x}, \mathbf{x}+(d,0,0), \mathbf{x}+(0,d,0), \mathbf{x}+(0,0,d)$, there is no current analogue of Shkredov's theorem from which to deduce Theorem 1.2. We can however obtain the qualitative bound $|A| = o(N^n)$ using the multidimensional Szemerédi theorem of Furstenberg and Katznelson [2].

The extra condition in Theorem 1.2 that all differences $T_3 - T_i$ $(i = 1, 2)$ are non-singular stems from our employment of the differencing method of Gowers [3]. In this method one extracts a quadratic non-uniformity estimate for the set $A$ from a linear non-uniformity estimate on the shifted sets $A \cap (A + \mathbf{k})$. However, one needs such an estimate for 'many' $\mathbf{k}$, in particular one needs a set of $\mathbf{k}$ dense in $\{1, \dots, N\}^n$. If one assumes some $T_3 - T_i$ is singular, then the set of shifts $\mathbf{k}$ for which we can obtain a non-uniformity estimate is necessarily contained in a subspace of dimension smaller than $n$. This set is too sparse for our methods to succeed.

To prove Theorem 1.2 we employ a multidimensional version of the 'weak' inverse theorem for the Gowers $U^3$-norm, analogous to the one-dimensional result obtained in [3]. Although one can obtain such a result by modifying the argument in [3], we opt for a shortcut, applying the strong $U^3(G)$-inverse theorem of Green and Tao [6], then 'localizing'. This argument is more succinct, but it is somewhat wasteful in that it ignores much of the extra information provided by the strong inverse theorem. We note that one could work much harder, emulating the methods of Green–Tao [7] to utilize this information, and thereby hope to improve the exponent of $\log\log N$ in Theorem 1.2.

Our arguments will be familiar to experts in the field and are very much modelled on Gowers's approach in [4]. Our main novelty, we believe, is in formulating the idea of suitably non-singular matrix progressions, and in showing that these are the natural configurations to which the methods of Gowers apply. Indeed, in the last section of [4], Gowers poses the challenge of adapting his methods to prove a quantitative version of the multidimensional Szemerédi theorem. The methods employed herein clearly demarcate multidimensional configurations into two types: those which should be amenable to a direct application of higher order Fourier analysis as in [4] (suitably non-singular matrix progressions), and those which require consideration of additional notions of uniformity (singular/sparse matrix progressions) as in Shkredov [13].

Even using higher order uniformity, our methods could not re-prove the quantitative corners result. Our approach requires that a set not containing the expected number of corners must be $k$-degree non-uniform for some $k$, as measured by the Gowers $U^k$-norm. However, as Shkredov observed (see for example the discussion in Green [5]), there are sets $A \subset \{1, \dots, N\}^2$ which

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

are uniform (and hence $k$-degree uniform for all $k \geqslant 2$), but which contain substantially more corners than expected. For instance, take $A = B^2$ where $B \subset \{1, \dots, N\}$ is linearly uniform as measured by the $U^2(\mathbb{Z})$-norm. In order to prove Shkredov's theorem, one must therefore make use of some other notion of uniformity than that measured by the $U^k$-norms (as Shkredov does). This obstacle becomes apparent in our approach in that the corner configuration

$$\{\mathbf{x}, \mathbf{x} + (d, 0), \mathbf{x} + (0, d)\} \quad (d \neq 0),$$

can only be defined as a subset of a matrix configuration

$$\{\mathbf{x}, \mathbf{x} + T_1\mathbf{d}, \dots, \mathbf{x} + T_k\mathbf{d}\} \quad (\mathbf{d} \in \mathbb{Z}^2 \setminus K),$$

if we allow more than one $T_i$ to be singular, or allow more than one difference $T_i - T_j$ $(i \neq j)$ to be singular. For instance, one could take

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad K = \{0\} \times \mathbb{R},$$

but now both $T_1$ and $T_2$ are singular. To overcome this, one might take

$$T_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad T_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad K = \{0\} \times \mathbb{R},$$

but now $T_3 - T_1$ and $T_3 - T_2$ are singular.

1.1. *Notation.* In order to prove Theorems 1.1 and 1.2 it is useful to work with more general subsets of $\mathbb{Z}^n$ than $\{1, 2, \dots, N\}^n$. Notice that this latter set is equal to the intersection $(0, N]^n \cap \mathbb{Z}^n$. We therefore define a *half-open cube* of side-length $L$ to be a subset of $\mathbb{R}^n$ of the form

$$Q = \mathbf{x} + (0, L]^n.$$

Given a half-open cube $Q$ we use $[Q]$ to denote the set of integer points in $Q$, namely

$$[Q] = Q \cap \mathbb{Z}^n.$$

Notice that if $Q$ has side-length $L \geqslant 1$ then the cardinality $|[Q]|$ satisfies

$$(L - 1)^n \leqslant |[Q]| < (L + 1)^n. \tag{1.3}$$

It will also be convenient to have the following abbreviations:

$$[L] = (0, L] \cap \mathbb{Z}, \qquad [\pm L] = [-L, L] \cap \mathbb{Z}, \qquad (\pm L] = (-L, L] \cap \mathbb{Z}.$$

We use $|\cdot|_\infty$ to denote the norm $|\mathbf{x}|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|$, and reserve $|\cdot|$ for the Euclidean norm on $\mathbb{R}^n$. So for example

$$[\pm L]^n = \{\mathbf{x} \in \mathbb{Z}^n : |\mathbf{x}|_\infty \leqslant L\}.$$

We also use $\mathbb{Z}_N$ to denote the additive group $\mathbb{Z}/N\mathbb{Z}$ of integers modulo $N$.

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

Let $\mathbb{T}^n$ denote the $n$-dimensional torus $\mathbb{R}^n/\mathbb{Z}^n$. Given a function $f : \mathbb{Z}^n \to \mathbb{C}$ with $\sum_{\mathbf{x}} |f(\mathbf{x})| < \infty$ we define its Fourier transform $\hat{f} : \mathbb{T}^n \to \mathbb{C}$ by

$$\hat{f}(\boldsymbol{\alpha}) = \sum_{\mathbf{x}} f(\mathbf{x}) \mathrm{e}(-\boldsymbol{\alpha} \cdot \mathbf{x}), \tag{1.4}$$

where $\mathrm{e}(\beta) = \mathrm{e}^{2\pi \mathrm{i}\beta}$. Let $\oint$ denote the integral over $\mathbb{T}^n$ with Lebesgue measure induced from $[0, 1)^n$. Given finite sets $A_1, \dots, A_s \subset \mathbb{Z}^n$, integer matrices $M_1, \dots, M_s$ and $\mathbf{c} \in \mathbb{Z}^n$, we have the orthogonality relations

$$\begin{aligned} &\#\{(\mathbf{x}_1, \dots, \mathbf{x}_s) \in A_1 \times \cdots \times A_s : M_1\mathbf{x}_1 + \cdots + M_s\mathbf{x}_s = \mathbf{c}\} \\ &\quad = \oint \hat{1}_{A_1}(\boldsymbol{\alpha} M_1) \cdots \hat{1}_{A_s}(\boldsymbol{\alpha} M_s) \mathrm{e}(-\boldsymbol{\alpha} \cdot \mathbf{c})\, d\boldsymbol{\alpha}. \end{aligned}$$

If $Q \subset \mathbb{R}^n$ is a half-open cube and $A \subset [Q]$ with $|A| = \delta|[Q]|$, then we define the balanced function of $A$ with respect to $Q$ to be the function

$$b_A = 1_A - \delta 1_{[Q]}.$$

Notice that $b_A$ has mean zero and is dependent on $Q$, although we suppress this dependence in our notation.

We use $C$ to denote a large absolute constant and $c$ to denote a small positive absolute constant. These constants will usually depend on other parameters, and we indicate this dependence in the statement of all results in which the constants occur. To differentiate between absolute constants within a proof we use subscripts, however the same $C_i$ may denote different absolute constants in different proofs.

Throughout this paper we have course to use simple facts from the geometry of numbers. Fortunately, we do not have to go very deep into the subject, no deeper than standard results concerning subgroups of $\mathbb{Z}^n$. We relegate the proof of these facts and all pertinent definitions to Appendix A.

§2. *Three-term matrix progressions.* The purpose of this section is to prove Theorem 1.1. It is notationally simpler to prove the following weaker version of the theorem, which is in fact equivalent.

THEOREM 2.1. *Let $t$ be a positive integer, $T$ a non-singular $n \times n$ integer matrix and $K$ a finite union of proper subspaces of $\mathbb{R}^n$. There exists an absolute constant $C = C(t, T, K)$ such that if $A \subset \{1, 2, \dots, N\}^n$ does not contain a configuration of the form $\{\mathbf{x}, \mathbf{x} + t\mathbf{d}, \mathbf{x} + T\mathbf{d}\}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$, then $A$ satisfies the bound*

$$|A| \leqslant \frac{CN^n}{\log\log N}. \tag{2.1}$$

*Proof that Theorem 2.1 implies Theorem 1.1.* Let $A \subset [N]^n$ be a set which does not contain any configurations of the form $\{\mathbf{x}, \mathbf{x} + T_1\mathbf{d}, \mathbf{x} + T_2\mathbf{d}\}$ with $\mathbf{d} \in$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

$\mathbb{Z}^n \setminus K$. We first show that $A$ has a large subset all of whose differences take the form $T_1\mathbf{d}$ with $\mathbf{d} \in \mathbb{Z}^n$. To this end, consider the $\mathbb{Z}^n$-subgroup

$$H = \mathbb{Z}^n \cap (T_1 \cdot \mathbb{Z}^n).$$

Since $T_1$ is non-singular, one sees that $H$ is a full-rank subgroup of $\mathbb{Z}^n$. Hence by Corollary A.3, the number of cosets of $H$ which intersect $[N]^n$ is a constant $C_H \ll_{T_1} 1$. By the pigeon-hole principle the set $A$ must have size $|A|/C_H$ on some such coset $\mathbf{x} + H$. Let $A' = A \cap (\mathbf{x} + H)$.

Clearly there exists a positive integer $t \ll_{T_1,T_2} 1$ such that the matrix $T = tT_2T_1^{-1}$ has only integer entries. Notice that this matrix is also non-singular. Let $K' = T_1 \cdot K$ and suppose that $A'$ contains $\mathbf{x}, \mathbf{x}+t\mathbf{d}', \mathbf{x}+T\mathbf{d}'$, for some $\mathbf{d}' \in \mathbb{R}^n \setminus K'$. By our construction of $A'$, there exists $\mathbf{d} \in \mathbb{Z}^n$ with $t\mathbf{d}' = T_1\mathbf{d}$. Notice that we must have $\mathbf{d} \notin K$, so $A$ contains $\mathbf{x}, \mathbf{x} + T_1\mathbf{d}, \mathbf{x} + T_2\mathbf{d}$ with $\mathbf{d} \in \mathbb{Z}^n \setminus K$. Since this cannot happen, Theorem 2.1 tells us that

$$|A|/C_H \leqslant |A'| \leqslant \frac{CN^n}{\log\log N}.$$

Theorem 1.1 follows. □

Until the proof of Theorem 2.1 is complete, let us fix $t \in \mathbb{Z} \setminus \{0\}$ and $T$ a non-singular $n \times n$ integer matrix. Define

$$S = T - tI,$$

where $I$ is the identity matrix. Given $A_1, A_2, A_3 \subset \mathbb{Z}^n$ define the counting function $R(A_1, A_2, A_3)$ to be the number of triples $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in A_1 \times A_2 \times A_3$ for which there exists $\mathbf{d} \in \mathbb{R}^n$ with $\mathbf{y} = \mathbf{x}+t\mathbf{d}$ and $\mathbf{z} = \mathbf{x}+T\mathbf{d}$. If $A_1 = A_2 = A_3$ we simply write

$$R(A) := R(A, A, A).$$

Notice that $R(A_1, A_2, A_3)$ is equal to the number of $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in A_1 \times A_2 \times A_3$ satisfying the matrix equation

$$S\mathbf{x} - T\mathbf{y} + t\mathbf{z} = 0. \tag{2.2}$$

Hence if the $A_i$ are finite, then by the orthogonality relations

$$R(A_1, A_2, A_3) = \oint \hat{1}_{A_1}(\boldsymbol{\alpha}S)\hat{1}_{A_2}(-\boldsymbol{\alpha}T)\hat{1}_{A_3}(t\boldsymbol{\alpha})\, d\boldsymbol{\alpha}.$$

We prove Theorem 2.1 using the following density increment result.

LEMMA 2.2. *There exist absolute constants $C = C(t, T)$ and $c = c(t, T) > 0$ such that for any $\delta > 0$ and any real*

$$L \geqslant \exp(C/\delta), \tag{2.3}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*if $Q \subset \mathbb{R}^n$ is a half-open cube of side-length at least $L$ and $A \subset [Q]$ satisfies*

$$|A| = \delta|[Q]| \quad \textit{and} \quad R(A) \leqslant c\delta|A|^2, \tag{2.4}$$

*then there exists a half-open cube $Q_1$ of side-length*

$$L_1 \geqslant L^{1/(n+2)}, \tag{2.5}$$

*along with $q \in \mathbb{N}$ and $\mathbf{r} \in \mathbb{Z}^n$ such that*

$$|A \cap (\mathbf{r} + q \cdot [Q_1])| \geqslant (\delta + c\delta^2)|[Q_1]|. \tag{2.6}$$

Theorem 2.1 follows readily from this.

*Proof that Lemma 2.2 implies Theorem 2.1.* Suppose that $A \subset [N]^n$ does not contain any configurations of the form $\mathbf{x}, \mathbf{x} + t\mathbf{d}, \mathbf{x} + T\mathbf{d}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$.

We aim to construct a sequence of quadruples $(Q_i, A_i, L_i, \delta_i)$ satisfying all of the following conditions:

(i) $Q_i$ is a half-open cube of side-length $L_i$;
(ii) $A_i \subset [Q_i]$ with $A_i = \delta_i|[Q_i]|$;
(iii) $A_i$ does not contain any configurations of the form $\mathbf{x}, \mathbf{x} + t\mathbf{d}, \mathbf{x} + T\mathbf{d}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$;
(iv) $L_{i+1} \geqslant L_i^{1/(n+2)}$;
(v) $\delta_{i+1} \geqslant \delta_i + c\delta_i^2$.

Taking $Q_0 = (0, N]^n$, $A_0 = A$, $L_0 = N + 1$ and $\delta_0 = |A_0|/|[Q_0]|$, we have our initial quadruple. Let us suppose we have constructed $(Q_j, A_j, X_j, \delta_j)$ for all $1 \leqslant j \leqslant i$. In order to apply Lemma 2.2, we must estimate $R(A_i)$. Notice that if $\mathbf{x}, \mathbf{x} + t\mathbf{d} \in A_i$, then $t\mathbf{d} \in [\pm L_i]^n \cap K$. Hence

$$R(A_i) \leqslant |A_i||K \cap [\pm L_i]^n|. \tag{2.7}$$

By Lemma A.1 and the fact that $K$ is a union of *proper* subspaces of $\mathbb{R}^n$, there exists a constant $C_K$ for which we have

$$|K \cap [\pm L_i]^n| \leqslant C_K L_i^{n-1}. \tag{2.8}$$

Let $c$ be as in Lemma 2.2. In order to satisfy (2.4), we need

$$C_K L_i^{n-1} \leqslant c\delta_i|A_i|. \tag{2.9}$$

Since $\delta_i L_i^n \ll |A_i|$, there exists an absolute constant $C_1$ such that if $L_i \geqslant C_1 C_K/(c\delta_0)$ then we do indeed have (2.9). Let $C_2 = \max\{C_1 C_K c^{-1}, C\}$, with $C$ as in Lemma 2.2, and let us suppose that

$$L_i \geqslant \exp(C_2/\delta_i). \tag{2.10}$$

Then combining this with equations (2.7) and (2.9), we see that

$$R(A_i) \leqslant c\delta_i|A_i|^2.$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

The bound (2.10) also gives us $L_i \geqslant \exp(C/\delta_i)$. Hence (2.3) and (2.4) are satisfied, so by Lemma 2.2, there exists a half-open cube $Q_{i+1}$ of side-length

$$L_{i+1} \geqslant L_i^{1/(n+2)},$$

along with $q_i \in \mathbb{N}$ and $\mathbf{r}_i \in \mathbb{Z}^n$ such that

$$|A_i \cap (\mathbf{r}_i + q_i \cdot [Q_{i+1}])| \geqslant (\delta_i + c\delta_i^2)|[Q_{i+1}]|.$$

Define $A_{i+1}$ to be the set of $\mathbf{y} \in [Q_{i+1}]$ for which $\mathbf{r}_i + q_i\mathbf{y} \in A_i$. Notice that $A_{i+1}$ does not contain any configurations of the form $\mathbf{x}, \mathbf{x} + t\mathbf{d}, \mathbf{x} + T\mathbf{d}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$. Setting $\delta_{i+1} = |A_{i+1}|/|[Q_{i+1}]|$ we have constructed another quadruple satisfying (i)–(v), subject to the assumption (2.10).

Iterating this construction $\lceil 1/(c\delta)\rceil$ times, we obtain a density $\delta_i \geqslant 2\delta$. Let $j = \lfloor 1/(\delta \log 2)\rfloor$, then after

$$I := \lceil 1/(c\delta)\rceil + \lceil 1/(2c\delta)\rceil + \cdots + \lceil 1/(2^j c\delta)\rceil$$

iterations we have density at least $2^{j+1}\delta > 1$, a contradiction. Hence for some $i \leqslant I$ the assumption (2.10) cannot hold. Notice that

$$\begin{aligned} i &\leqslant \lceil 1/(c\delta)\rceil + \lceil 1/(2c\delta)\rceil + \cdots + \lceil 1/(2^j c\delta)\rceil \\ &\leqslant 1/(\delta \log 2) + 1 + 2/c\delta \\ &\leqslant 3/c\delta. \end{aligned}$$

Iterating (iv) we have

$$L_i \geqslant N^{1/(n+2)^i},$$

and trivially $\delta_i \geqslant \delta$. Thus

$$N^{1/(n+2)^{3/c\delta}} \leqslant L_i < \exp(C_2/\delta).$$

Taking logarithms of either side of this inequality gives

$$\log N \leqslant C_2(n+2)^{3/c\delta}/\delta.$$

Taking another logarithm we have

$$\begin{aligned} \log\log N &\leqslant (3/c)\delta^{-1}\log(C_2(n+2)) + \log(1/\delta) \\ &\leqslant (3c^{-1}\log(C_2(n+2)) + 1)\delta^{-1}. \end{aligned}$$

The bound (2.1) now follows, thereby establishing Theorem 2.1. □

The remainder of this section is therefore taken up with a proof of Lemma 2.2. Given a half-open cube $Q$, we say that a subset $A \subset [Q]$ is *linearly uniform* if all the Fourier coefficients of its balanced function $b_A$ are small, so that $\|\hat{b}_A\|_\infty/|[Q]|$ is smaller than some fixed power of $\delta$. In this section we will

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

show that if $A$ contains much less than the expected number of configurations of the form $\{\mathbf{x}, \mathbf{x}+t\mathbf{d}, \mathbf{x}+T\mathbf{d}\}$, then $A$ cannot be linearly uniform. To reach this conclusion we compare $R(A)$ with $\delta R([Q], A, A)$, but we must first establish that $R([Q], A, A)$ is sufficiently large. To show this, one must deal with the case when $A$ is concentrated near the boundary of $[Q]$. This is done in our first lemma, which also proves useful in §3.

LEMMA 2.3. *Let $Q$ be a half-open cube of side-length $L \geqslant 30n^2$ and let $A \subset [Q]$ with $|A| = \delta|[Q]|$. Suppose that for every half-open cube $Q'$ of side-length at least $L/(12n)$, we have*

$$|A \cap Q'| < 2\delta|[Q']|. \tag{2.11}$$

*Then there exists a half-open subcube $Q_0 \subset Q$ of side-length $L/(12n)$ with $|A \cap Q_0| \geqslant \frac{1}{2}\delta|[Q_0]|$ and*

$$Q_0 + \left[-\frac{L}{12n}, \frac{L}{12n}\right]^n \subset Q. \tag{2.12}$$

*Proof.* Setting $m = 12n$, we can partition $Q$ into $m^n$ half-open subcubes each of side-length $L/m$. Let $\mathcal{C}$ denote the set of such subcubes. Define $\partial\mathcal{C}$ to be the set of $Q' \in \mathcal{C}$ whose boundary intersects the boundary of $Q$, and let $\mathcal{C}_1$ denote the set of $Q' \in \mathcal{C}$ satisfying

$$|A \cap Q'| \geqslant (\delta/2)|[Q']|.$$

Notice that

$$\begin{aligned}
\delta|[Q]| = |A| &= \sum_{Q' \in \mathcal{C}} |A \cap Q'| \\
&= \sum_{Q' \in \mathcal{C}_1} |A \cap Q'| + \sum_{Q' \in \mathcal{C} \setminus \mathcal{C}_1} |A \cap Q'| \\
&< 2\delta \sum_{Q' \in \mathcal{C}_1} |[Q']| + \frac{1}{2}\delta \sum_{Q' \in \mathcal{C} \setminus \mathcal{C}_1} |[Q']| \\
&\leqslant \frac{3\delta}{2}|\mathcal{C}_1|\left(\frac{L}{m}+1\right)^n + \frac{\delta}{2}|[Q]|.
\end{aligned} \tag{2.13}$$

We have

$$\frac{L}{m} + 1 = \frac{L-1}{m}\left(1 + \frac{m+1}{L-1}\right) \leqslant \frac{L-1}{m}\exp\left(\frac{m+1}{L-1}\right).$$

Since $L \geqslant 30n^2 \geqslant 2n(m+1)+1$, we therefore have

$$\left(\frac{L}{m}+1\right)^n \leqslant 2\left(\frac{L-1}{m}\right)^n \leqslant 2|[Q]|m^{-n}.$$

Combining this with (2.13), we see that $|\mathcal{C}_1| > m^n/6$. There are at most $2nm^{n-1}$ elements in $\partial\mathcal{C}$. Since $m = 12n$, we obtain $|\mathcal{C}_1| > |\partial\mathcal{C}|$. The lemma now follows. □

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

LEMMA 2.4. *Let $Q$ be a half-open cube of side-length $L \geqslant 30n^2$ and let $A \subset [Q]$ with $|A| = \delta|[Q]|$. There exists a constant $c = c(t, T)$ such that either*

$$R([Q], A, A) \geqslant c|A|^2, \tag{2.14}$$

*or there exists a half-open cube $Q'$ of side-length at least $L/(12n)$ such that*

$$|A \cap Q'| \geqslant 2\delta|[Q']|. \tag{2.15}$$

*Proof.* Suppose (2.15) does not hold and let $Q_0$ denote the half-open cube provided by Lemma 2.3. Let $r$ denote the rank of $S = T - tI$ and consider the $r$-dimensional vector space $U_1 = S \cdot \mathbb{Q}^n$. Setting $K = \ker S \leqslant \mathbb{Q}^n$, it follows from rank–nullity that there exists an $r$-dimensional subspace $U_0 \leqslant Q^n$ such that $\mathbb{Q}^n = K \oplus U_0$, and for which the restriction $S|_{U_0}$ is a linear isomorphism from $U_0$ to $U_1$. From the elementary theory of normed spaces, there must exist a constant $M = M(S) > 0$ such that for all $\mathbf{u} \in U_0$ we have

$$|\mathbf{u}|_\infty \leqslant M|S\mathbf{u}|_\infty.$$

Increasing $M$ if necessary, we may assume $M \in \mathbb{N}$. Let $H$ denote the $\mathbb{Z}^n$-subgroup $H = (S \cdot (U_0 \cap \mathbb{Z}^n)) \cap \mathbb{Z}^n$. Clearly $H$ also has rank $r$. Let us partition $Q_0$ into $(2M)^n$ half-open subcubes each of side-length $L/(24n|t|M)$. On some such subcube $Q_1$, the set $A$ must have density at least $\delta/2$, by which we mean $|A \cap Q_1| \geqslant (\delta/2)|[Q_1]|$. Define $A' = A \cap Q_1$. Notice that by (2.12) we have

$$A' + [\pm L/(12n)]^n \subset [Q].$$

Moreover if $\mathbf{y}, \mathbf{z} \in A'$ with $\mathbf{y} - \mathbf{z} \in H$, then $\mathbf{y} - \mathbf{z} = S \cdot \mathbf{u}$ for some $\mathbf{u} \in \mathbb{Z}^n \cap U_0$ and $|\mathbf{y} - \mathbf{z}| \leqslant L/(24n|t|M)$ so

$$|\mathbf{u}|_\infty \leqslant M|\mathbf{y} - \mathbf{z}|_\infty \leqslant L/(24|t|n).$$

Setting $\mathbf{x} := \mathbf{y} + t\mathbf{u}$ we deduce that $T\mathbf{y} - t\mathbf{z} = S\mathbf{y} + t(\mathbf{y} - \mathbf{z}) = S\mathbf{x}$. Moreover, by the triangle inequality $\mathbf{x}$ satisfies

$$\mathbf{x} + [\pm L/(24n)]^n \subset [Q]. \tag{2.16}$$

Given $\mathbf{x}$ such that (2.16) holds, we would like to estimate from below the number of $\mathbf{x}' \in [Q]$ such that $S\mathbf{x} = S\mathbf{x}'$, or equivalently we would like a lower bound for the size of the intersection $(\mathbf{x} + K) \cap [Q]$. Since

$$\mathbf{x} + (K \cap [\pm L/(24n)]^n) \subset (\mathbf{x} + K) \cap [Q],$$

it suffices to estimate $|K \cap [\pm L/(24n)]^n|$ from below. By Lemma A.1, there exists a constant $C_K > 0$ such that

$$|K \cap [\pm L/(24n)]^n| \geqslant L^{n-r}/C_K.$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

It follows that for every pair $\mathbf{y}, \mathbf{z} \in A'$ with $\mathbf{y} - \mathbf{z} \in H$ we have the lower bound

$$\#\{\mathbf{x} \in [Q] : S\mathbf{x} = T\mathbf{y} - t\mathbf{z}\} \geqslant C_K L^{n-r}.$$

Thus

$$\begin{aligned} R([Q], A, A) &\geqslant R([Q], A', A') \\ &\geqslant \sum_{\mathbf{y}-\mathbf{z}\in H} 1_{A'}(\mathbf{y}) 1_{A'}(\mathbf{z}) |\{\mathbf{x} \in [Q] : S\mathbf{x} = T\mathbf{y} - t\mathbf{z}\}| \\ &\geqslant C_K^{-1} L^{n-r} \sum_{\mathbf{y}-\mathbf{z}\in H} 1_{A'}(\mathbf{y}) 1_{A'}(\mathbf{z}). \end{aligned} \tag{2.17}$$

By Corollary A.4 we have

$$\sum_{\mathbf{y}-\mathbf{z}\in H} 1_{A'}(\mathbf{y}) 1_{A'}(\mathbf{z}) \geqslant |A'|^2 L^{r-n}/C_H.$$

Combining this with (2.17), we obtain

$$\begin{aligned} R([Q], A, A) &\geqslant \frac{1}{C_K C_H} |A'|^2 \\ &\gg_{t,T} |A|^2. \end{aligned}$$

□

Using Lemma 2.4 we are able to extract a non-uniformity estimate for $b_A$.

LEMMA 2.5. *There exists an absolute constant $c = c(t, T) > 0$ such that for any $\delta > 0$ and any real $L \geqslant 30n^2$, if $Q \subset \mathbb{R}^n$ is a half-open cube of side-length at least $L$ and $A \subset [Q]$ satisfies $|A| = \delta|[Q]|$ and*

$$R(A) \leqslant c\delta|A|^2, \tag{2.18}$$

*then either*

$$\|\hat{b}_A\|_\infty \geqslant c\delta|A|, \tag{2.19}$$

*or there exists a half-open cube $Q'$ of side-length at least $L/(12n)$ such that*

$$|A \cap Q'| \geqslant 2\delta|[Q']|. \tag{2.20}$$

*Proof.* Let us suppose that (2.20) does not hold for any such $Q'$. Combining this with $L \geqslant 30n^2$ and Lemma 2.4 we see that there exists a constant $c_1 > 0$, dependent only on $t$ and $T$, such that

$$R([Q], A, A) \geqslant c_1|A|^2.$$

Taking $c = c_1/2$ and assuming (2.18), we see that the difference between $\delta R([Q], A, A)$ and $R(A)$ satisfies

$$c\delta|A|^2 \leqslant |R(A) - \delta R([Q], A, A)|.$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

By the orthogonality relations, the invertibility of $T$ and Parseval's inequality, we therefore have

$$\begin{aligned} c\delta|A|^2 &\leqslant \left| \oint \hat{b}_A(\boldsymbol{\alpha} S)\hat{1}_A(-\boldsymbol{\alpha} T)\hat{1}_A(t\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \right| \\ &\leqslant \|\hat{b}_A\|_\infty \|\hat{1}_A\|_2^2 \\ &\leqslant \|\hat{b}_A\|_\infty |A|. \end{aligned}$$

The result follows. □

We are now able to prove Lemma 2.2.

*Proof of Lemma 2.2.* Let $c_1$ denote the constant obtained in Lemma 2.5 and set $C = \max\{1/(36n^2), 4\pi n/c_1\}$. Let us suppose that $L \geqslant \exp(C/\delta)$, that $Q \subset \mathbb{R}^n$ is a half-open cube of side-length at least $L$ and that $A \subset [Q]$ satisfies both $|A| = \delta|[Q]|$ and

$$R(A) \leqslant c_1\delta|A|^2. \tag{2.21}$$

The conditions of Lemma 2.5 are then satisfied. It follows that either

$$\|\hat{b}_A\|_\infty \geqslant c_1\delta|A|, \tag{2.22}$$

or there exists a half-open cube $Q'$ of side-length at least $L/(6n)$ such that

$$|A \cap Q'| \geqslant 2\delta|[Q']|. \tag{2.23}$$

Since $L^{1/2} \geqslant 1/(6n)$, we certainly have

$$L/(6n) \geqslant L^{1/(n+2)}.$$

Hence if (2.23) holds, then we can take $L_1 = L/(6n)$, $Q_1 = Q'$, $\mathbf{r} = 0$ and $q = 1$ to obtain (2.5) and (2.6). In this case we are done.

Let us therefore suppose that (2.22) holds, so that there exists $\boldsymbol{\alpha} \in \mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ satisfying

$$\left| \sum_{\mathbf{x}} b_A(\mathbf{x})\mathrm{e}(\boldsymbol{\alpha} \cdot \mathbf{x}) \right| \geqslant c_1\delta|A|.$$

By Kronecker's theorem on Diophantine approximation, there exists an integer $1 \leqslant q \leqslant L^{n/(n+1)}$ such that

$$\|q\alpha_i\| \leqslant L^{-1/(n+1)} \quad (1 \leqslant i \leqslant n).$$

Let us partition $[Q]$ into congruence classes modulo $q$. Each class takes the form

$$\mathbf{r} + q \cdot [Q(\mathbf{r})],$$

where $Q(\mathbf{r})$ is a half-open cube of side-length $L/q$. Let $m$ be a positive integer satisfying

$$\frac{L^{n/(n+1)}\log L}{q} \leqslant m \leqslant \frac{2L^{n/(n+1)}\log L}{q}. \tag{2.24}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

We can partition each $Q(\mathbf{r})$ into $m^n$ half-open cubes $Q_1(\mathbf{r}), Q_2(\mathbf{r}), \dots$ of side-length $L/(qm)$. Then for any $\mathbf{y}, \mathbf{z} \in [Q_i(\mathbf{r})]$ we have

$$\begin{aligned}|e(\boldsymbol{\alpha}\cdot(\mathbf{r}+q\mathbf{y})) - e(\boldsymbol{\alpha}\cdot(\mathbf{r}+q\mathbf{z}))| &\leqslant 2\pi\|q\boldsymbol{\alpha}\cdot(\mathbf{y}-\mathbf{z})\| \\ &\leqslant 2\pi n L^{n/(n+1)}/qm \\ &\leqslant 2\pi n/\log L.\end{aligned}$$

Thus

$$c_1\delta|A| \leqslant \sum_{\substack{\mathbf{r}\in[q]^n\\ 1\leqslant i\leqslant m^n}} \left|\sum_{\mathbf{x}\in\mathbf{r}+Q_i(\mathbf{r})} b_A(\mathbf{x})\right| + \frac{2\pi n|A|}{\log L}.$$

Since $L \geqslant \exp(4\pi n/(c_1\delta))$ we certainly have

$$\frac{1}{2}c_1\delta|A| \leqslant \sum_{\substack{\mathbf{r}\in[q]^n\\ 1\leqslant i\leqslant m^n}} \left|\sum_{\mathbf{x}\in\mathbf{r}+Q_i(\mathbf{r})} b_A(\mathbf{x})\right|. \tag{2.25}$$

Let $\mathcal{P}$ denote the set of pairs $(\mathbf{r}, i)$ for which

$$\sum_{\mathbf{x}\in\mathbf{r}+Q_i(\mathbf{r})} b_A(\mathbf{x}) \geqslant 0.$$

Since

$$\sum_{\substack{\mathbf{r}\in[q]^n\\ 1\leqslant i\leqslant m^n}} \sum_{\mathbf{x}\in\mathbf{r}+Q_i(\mathbf{r})} b_A(\mathbf{x}) = \sum_{\mathbf{x}} b_A(\mathbf{x}) = 0,$$

we can add this quantity to (2.25) and deduce that

$$2\sum_{(\mathbf{r},i)\in\mathcal{P}} \sum_{\mathbf{x}\in\mathbf{r}+Q_i(\mathbf{r})} b_A(\mathbf{x}) \geqslant \frac{1}{2}c_1\delta|A|.$$

It follows that there exists a half-open cube $Q_1$ of side-length $L_1 = L/(qm)$, along with $\mathbf{r} \in \mathbb{Z}^n$ and $q \in \mathbb{N}$ such that

$$\sum_{\mathbf{x}\in\mathbf{r}+q\cdot[Q_1]} b_A(\mathbf{x}) \geqslant \frac{c_1\delta|A|}{4(qm)^n}. \tag{2.26}$$

Increasing $C$ by a constant dependent only on $n$, one can ensure that $L \geqslant \exp(C/\delta)$ certainly guarantees that

$$\frac{|[Q]|}{(qm)^n} \geqslant \left(\frac{L-1}{qm}\right)^n \geqslant \frac{1}{2}\left(\frac{L^n}{(qm)^n}+1\right) \geqslant \frac{|[Q_1]|}{2}.$$

Using this in (2.26), we have

$$\sum_{\mathbf{x}\in\mathbf{r}+q\cdot[Q_1]} b_A(\mathbf{x}) \geqslant \frac{1}{8}c_1\delta^2|[Q_1]|.$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

Taking $c = c_1/8$ in the statement of Lemma 2.2, it only remains to show that

$$L_1 = L/(qm) \geqslant L^{1/(n+2)}. \tag{2.27}$$

By (2.24), we have $L/(qm) \geqslant L^{1/(n+1)}/(2\log L)$, so that we obtain (2.27) from $L \geqslant \exp(C/\delta)$, if necessary increasing $C$ by a factor dependent only on $n$. □

This completes our proof of Theorem 1.1.

§3. *Four-term matrix progressions.* In proving Theorem 1.2, we follow the large-scale structure of the previous section, with one significant difference. Instead of utilizing a lack of linear non-uniformity, we have to resort to methods from quadratic Fourier analysis, developed by Gowers [**3**] and Green–Tao [**6**], to show that a set with fewer than the expected number of four-point configurations has a density increment on some half-open cube.

We deduce Theorem 1.2 from the following seemingly weaker result. The proof of their equivalence is similar to the deduction of Theorem 1.1 from Theorem 2.1, and is left as an exercise for the reader.

THEOREM 3.1. *Let $T_1$, $T_2$ be $n \times n$ non-singular integer matrices and let $t \in \mathbb{N}$. Suppose that both $T_1 - tI$ and $T_2 - tI$ are non-singular. Let $K$ denote a finite union of proper subspaces of $\mathbb{R}^n$. There exist absolute constants $C = C(t, T_i, K)$ and $c = 2^{-26}$, such that any subset $A \subset \{1, 2, \ldots, N\}^n$ which does not contain a configuration of the form $\{\mathbf{x}, \mathbf{x}+t\mathbf{d}, \mathbf{x}+T_1\mathbf{d}, \mathbf{x}+T_2\mathbf{d}\}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$, satisfies the bound*

$$|A| \leqslant \frac{CN^n}{(\log\log N)^c}. \tag{3.1}$$

Until the proof of Theorem 2.1 is complete, let us fix a positive integer $t$ along with non-singular $n \times n$ integer matrices $T_1, T_2$ such that the differences

$$S_i = T_i - tI$$

are both non-singular.

Given $A_1, A_2, A_3, A_4 \subset \mathbb{Z}^n$ write $R(A_1, A_2, A_3, A_4)$ for the number of tuples

$$(\mathbf{x}, \mathbf{y}, \mathbf{z}_1, \mathbf{z}_2) \in A_1 \times A_2 \times A_3 \times A_4$$

for which there exists $\mathbf{d} \in \mathbb{R}^n$ with $\mathbf{y} = \mathbf{x} + t\mathbf{d}$, $\mathbf{z}_1 = \mathbf{x} + T_1\mathbf{d}$ and $\mathbf{z}_2 = \mathbf{x} + T_2\mathbf{d}$. Notice that $(\mathbf{x}, \mathbf{y}, \mathbf{z}_1, \mathbf{z}_2)$ is a tuple counted by $R(A_1, A_2, A_3, A_4)$ if and only if it satisfies the simultaneous matrix equations

$$S_i\mathbf{x} - T_i\mathbf{y} + t\mathbf{z}_i = 0 \quad (i = 1, 2). \tag{3.2}$$

As in §2, we use $R(A)$ to denote $R(A, A, A, A)$.

LEMMA 3.2. *There exist positive absolute constants $C = C(t, T_1, T_2)$ and $c = c(t, T_1, T_2)$ such that for any $\delta > 0$ and any real*

$$L \geqslant \exp(C\delta^{-C}), \tag{3.3}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*if $Q \subset \mathbb{R}^n$ is a half-open cube of side-length at least $L$ and $A \subset [Q]$ satisfies $|A| = \delta|[Q]|$ and*

$$R(A) \leqslant c\delta^2|A|^2, \tag{3.4}$$

*then there exists a half-open cube $Q_1$ of side-length*

$$L_1 \geqslant L^{c\delta^C}, \tag{3.5}$$

*along with $q \in \mathbb{N}$ and $\mathbf{r} \in \mathbb{Z}^n$ such that*

$$|A \cap (\mathbf{r} + q \cdot [Q_1])| \geqslant (\delta + c\delta^{2^{25}})|[Q_1]|. \tag{3.6}$$

Lemma 3.2 implies Theorem 3.1.

*Proof of Theorem 3.1.* Suppose $A \subset \{1, \dots, N\}^n$ does not contain any configurations of the form $\{\mathbf{x}, \mathbf{x} + t\mathbf{d}, \mathbf{x} + T_1\mathbf{d}, \mathbf{x} + T_2\mathbf{d}\}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$. Set $Q_0 = (0, N]^n$, $A_0 = A$, $L_0 = N + 1$ and $\delta_0 = |A_0|/|[Q_0]|$. Let $C_1$ and $c_1$ be as in the statement of Lemma 3.2 and suppose that for $1 \leqslant j \leqslant i$ we have found quadruples $(Q_j, A_j, L_j, \delta_j)$ satisfying the following:

(i) $Q_i$ is a half-open cube of side-length $L_i$;
(ii) $A_i \subset [Q_i]$ with $A_i = \delta_i|[Q_i]|$;
(iii) $A_i$ does not contain any configurations of the form $\mathbf{x}, \mathbf{x} + t\mathbf{d}, \mathbf{x} + T_1\mathbf{d}$, $\mathbf{x} + T_2\mathbf{d}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$;
(iv) $L_{i+1} \geqslant L_i^{c_1\delta_1^C}$;
(v) $\delta_{i+1} \geqslant \delta_i + c_1\delta_i^{2^{25}}$.

If $\mathbf{x}, \mathbf{x} + t\mathbf{d} \in A_i$, then $t\mathbf{d} \in [\pm L_i]^n \cap K$. Hence

$$R(A_i) \leqslant |A_i||K \cap [\pm L_i]|. \tag{3.7}$$

By Lemma A.1 there exists a constant $C_K$ such that

$$|K \cap [\pm L_i]^n| \leqslant C_K L_i^{n-1}. \tag{3.8}$$

Let $C_2$ be sufficiently large to ensure that the assumption

$$L_i \geqslant \exp(C_2\delta_i^{-C_2}) \tag{3.9}$$

implies that $C_K L_i^{n-1} \leqslant c_1\delta_i|A_i|$. Clearly we can guarantee that $C_1 \leqslant C_2 \ll_{t,T_1,T_2,K} 1$. Assuming (3.9) we therefore have

$$R(A_i) \leqslant c_1\delta_i^2|A_i|^2.$$

Employing Lemma 3.2, there exists a half-open cube $Q_{i+1}$ of side-length

$$L_{i+1} \geqslant L_i^{c_1\delta_i^{C_1}},$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

along with $q_i \in \mathbb{N}$ and $\mathbf{r}_i \in \mathbb{Z}^n$ such that

$$|A_i \cap (\mathbf{r}_i + q_i \cdot [Q_{i+1}])| \geqslant (\delta_i + c_1\delta_i^{2^{25}})|[Q_{i+1}]|.$$

Define $A_{i+1}$ to be the set of $\mathbf{y} \in [Q_{i+1}]$ for which $\mathbf{r}_i + q_i\mathbf{y} \in A_i$. Notice that $A_{i+1}$ does not contain any configurations of the form $\mathbf{x}, \mathbf{x}+t\mathbf{d}, \mathbf{x}+T_1\mathbf{d}, \mathbf{x}+T_2\mathbf{d}$ with $\mathbf{d} \in \mathbb{R}^n \setminus K$. Setting $\delta_{i+1} = |A_{i+1}|/|[Q_{i+1}]|$, we have obtained another such quadruple $(Q_{i+1}, A_{i+1}, L_{i+1}, \delta_{i+1})$.

We iterate this construction, subject to the assumption

$$L_i \geqslant \exp(C_2/\delta_i^{C_2}). \tag{3.10}$$

After $\lceil c_1^{-1}\delta^{-2^{25}} \rceil \leqslant 2c_1^{-1}\delta^{-2^{25}}$ iterations, we have $\delta_i > 1$, which is not possible. Hence for some $i \leqslant 2c_1^{-1}\delta^{-2^{25}}$ we have

$$L_i < \exp(C_2/\delta_i^{C_2}).$$

Using this together with (v) and the trivial bound $\delta_i \geqslant \delta$, we obtain

$$N^{(c_1\delta^{C_1})^i} \leqslant \exp(C_2/\delta^{C_2}).$$

Taking logarithms of either side of this inequality gives

$$(c_1\delta^{C_1})^i \log N \leqslant C_2/\delta^{C_2}.$$

Taking logarithms again, we have

$$\begin{aligned}\log\log N &\leqslant \log(C_2/\delta^{C_2}) + i\log(1/(c_1\delta^{C_1}))\\ &\leqslant \log(C_2/\delta^{C_2}) + 2c_1^{-1}\delta^{-2^{25}}\log(1/(c_1\delta^{C_1}))\\ &\ll_{c_1,C_1,C_2} \delta^{-2^{26}}.\end{aligned}$$

Since $c_1$, $C_1$ and $C_2$ are constants dependent only on $t$, $T_1$, $T_2$ and $K$, the bound (3.1) follows with $c = 2^{-26}$. □

We are left with the task of proving Lemma 3.2.

3.1. *A quadratic non-uniformity estimate.* We begin with a result analogous to Lemma 2.4, establishing that the number of $\mathbf{x}, \mathbf{d} \in \mathbb{Z}^n$ with $\mathbf{x}, \mathbf{x}+t\mathbf{d} \in [Q]$ and $\mathbf{x}+T_1\mathbf{d}, \mathbf{x}+T_2\mathbf{d} \in A$ is within a constant of the maximum.

LEMMA 3.3. *Let $Q$ be a half-open cube of side-length $L \geqslant 30n^2$ and let $A \subset [Q]$ with $|A| = \delta|[Q]|$. There exists a constant $c = c(t, T_1, T_2) > 0$ such that either*

$$R([Q], [Q], A, A) \geqslant c|A|^2, \tag{3.11}$$

*or there exists a half-open cube $Q'$ of side-length at least $L/(12n)$ such that*

$$|A \cap Q'| \geqslant 2\delta|[Q']|. \tag{3.12}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*Proof.* Let us assume (3.12) does not hold. By Lemma 2.3 there exists a half-open subcube $Q_0 \subset Q$ of side-length $L/(12n)$ with $|A \cap Q_0| \geqslant \frac{1}{2}\delta|[Q_0]|$ and

$$Q_0 + \left[-\frac{L}{12n}, \frac{L}{12n}\right]^n \subset Q.$$

Let $S$ denote the matrix $T_1 - T_2$ and let $r$ denote its rank. Consider the $\mathbb{Q}^n$-subspaces $U_1 = S \cdot \mathbb{Q}^n$ and $K = \ker S$. By rank–nullity, there exists an $r$-dimensional subspace $U_0 \leqslant \mathbb{Q}^n$ such that $\mathbb{Q}^n = K \oplus U_0$ and for which $S$ induces a linear isomorphism from $U_0$ to $U_1$. By the elementary theory of normed spaces, we can find an integer $M_1 \in \mathbb{N}$ such that for all $\mathbf{u} \in U_1$ we have the estimate

$$|S^{-1}\mathbf{u}|_\infty \leqslant M_1|\mathbf{u}|_\infty. \tag{3.13}$$

Similarly, there exists $M_2 \in \mathbb{N}$ such that for all $\mathbf{x} \in \mathbb{Q}^n$ we have

$$|T\mathbf{x}|_\infty \leqslant M_2|\mathbf{x}|_\infty. \tag{3.14}$$

Notice that $M = M_1M_2 \ll_{T_1,T_2} 1$. Let $L_0 = L/(12n)$, the side-length of $Q_0$. Let us partition $Q_0$ into half-open cubes, each of side-length $L_1 = L_0/(4M)$. By the pigeon-hole principle, some such cube $Q_1$ must have $|A \cap Q_1| \geqslant (\delta/2)|[Q_1]|$. Set $A_1 = A \cap Q_1$.

First we show there are many pairs $(\mathbf{z}_1, \mathbf{z}_2) \in A_1^2$ for which there exists $\mathbf{u} \in \mathbb{Z}^n$ with

$$\mathbf{z}_1 - \mathbf{z}_2 = S\mathbf{u} \quad \text{and} \quad T_1\mathbf{u} \in \mathbb{Z}^n.$$

To this end, consider the $\mathbb{Z}^n$-subgroup

$$H = \mathbb{Z}^n \cap S \cdot (\mathbb{Z}^n \cap U_0 \cap T_1^{-1}\mathbb{Z}^n).$$

One can check that this has rank $r$. By Corollary A.4 there are at least $|A_1|^2L^{r-n}/C_H$ pairs $(\mathbf{z}_1, \mathbf{z}_2) \in A_1^2$ with $\mathbf{z}_1 - \mathbf{z}_2 \in H$. For each such pair we show that there are many pairs $(\mathbf{x}, \mathbf{y}) \in [Q] \times [Q]$ satisfying

$$S_i\mathbf{x} - T_i\mathbf{y} + \mathbf{z}_i = 0 \quad (i = 1, 2).$$

We know there exists $\mathbf{u} \in \mathbb{Z}^n \cap U_0 \cap T_1^{-1}\mathbb{Z}^n$ satisfying

$$\mathbf{z}_1 - \mathbf{z}_2 = S\mathbf{u}.$$

By (3.13), we have

$$\begin{aligned} |\mathbf{u}|_\infty &= |S^{-1}S\mathbf{u}|_\infty \\ &\leqslant M_1|S\mathbf{u}|_\infty \\ &= M_1|\mathbf{z}_1 - \mathbf{z}_2|_\infty \leqslant \frac{L}{48nM_2}. \end{aligned}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

Let us define a second $\mathbb{Z}^n$-subgroup

$$H' = \mathbb{Z}^n \cap K \cap T_1^{-1}\mathbb{Z}^n.$$

For each $\mathbf{d} \in H' \cap [-L_1, L_1]^n$ set

$$\mathbf{x} = \mathbf{z}_1 - T_1(\mathbf{u} + \mathbf{d}) \quad \text{and} \quad \mathbf{y} = \mathbf{x} + \mathbf{u} + \mathbf{d}.$$

We claim that

$$\mathbf{x}, \mathbf{y} \in [Q] \quad \text{and} \quad S_i\mathbf{x} - T_i\mathbf{y} + \mathbf{z}_i = 0 \quad (i = 1, 2). \tag{3.15}$$

To see the equations are satisfied we make the substitution

$$\begin{aligned} S_1\mathbf{x} - T_1\mathbf{y} + \mathbf{z}_1 &= T_1\mathbf{x} - \mathbf{x} - T_1\mathbf{x} - T_1(\mathbf{u} + \mathbf{d}) + \mathbf{z}_1 \\ &= -\mathbf{x} - T_1(\mathbf{u} + \mathbf{d}) + \mathbf{z}_1 \\ &= 0. \end{aligned}$$

Similarly, since $\mathbf{z}_1 - \mathbf{z}_2 = T_1(\mathbf{u} + \mathbf{d}) - T_2(\mathbf{u} + \mathbf{d})$ we have

$$\begin{aligned} S_2\mathbf{x} - T_2\mathbf{y} + \mathbf{z}_2 &= T_2\mathbf{x} - \mathbf{x} - T_2\mathbf{x} - T_2(\mathbf{u} + \mathbf{d}) + \mathbf{z}_1 + T_2(\mathbf{u} + \mathbf{d}) - T_1(\mathbf{u} + \mathbf{d}) \\ &= -\mathbf{x} + \mathbf{z}_1 - T_1(\mathbf{u} + \mathbf{d}) \\ &= 0. \end{aligned}$$

We know that $\mathbf{x} \in \mathbb{Z}^n$, since $\mathbf{z}_1 \in \mathbb{Z}^n$ and $\mathbf{u}, \mathbf{d} \in T_1^{-1}\mathbb{Z}^n$. This also implies that $\mathbf{y} \in \mathbb{Z}^n$, since $\mathbf{u}, \mathbf{d} \in \mathbb{Z}^n$. To see that $\mathbf{x} \in Q$ note that $\mathbf{z}_1 \in Q_0$, that $Q_0 + [-L_0, L_0]^n \subset Q$ and that

$$\begin{aligned} |T_1(\mathbf{u} + \mathbf{d})|_\infty &\leqslant M_2(|\mathbf{u}|_\infty + |\mathbf{d}|_\infty) \\ &\leqslant L/(24n) = L_0/2. \end{aligned}$$

In fact, we see that

$$\mathbf{x} \in Q_0 + [-L_0/2, L_0/2]^n.$$

Combining this with the estimate $|\mathbf{u} + \mathbf{d}|_\infty \leqslant L_0/2$ we obtain $\mathbf{y} \in Q$. This establishes (3.15).

The number $R([Q], [Q], A, A)$ is thus at least the number of pairs $(\mathbf{z}_1, \mathbf{z}_2) \in A_1^2$ with $\mathbf{z}_1 - \mathbf{z}_2 \in H$, multiplied by the number of $\mathbf{d} \in H' \cap [-L_1, L_1]^n$. The number of pairs is at least $|A_1|L^{r-n}/C_H$, whilst the number of $\mathbf{d}$ is, by Lemma A.2, at least $L_1^{n-r}/C_{H'} \gg_{T_1,T_2} L^{n-r}$. Hence

$$R([Q], [Q], A, A) \gg_{T_1,T_2} |A^2|L^{r-n}L^{n-r} = |A|^2,$$

as required. □

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

Next we emulate Lemma 2.5 by showing that if $A \subset [Q]$ does not contain the expected number of four-point configurations, then $A$ has non-uniform balanced function. This time however, non-uniformity is measured in terms of the Gowers $U^3$-norm. To define this norm we introduce the forward difference operator $\Delta$, which for each element $\mathbf{h}$ of an abelian group $G$ maps a real function $f : G \to \mathbb{R}$ to the function

$$\Delta(f, \mathbf{h})(\mathbf{x}) = f(\mathbf{x})f(\mathbf{x}+\mathbf{h}) \quad (\mathbf{x} \in G).$$

The Gowers $U^k(G)$-norm of $f : G \to \mathbb{R}$ is then defined via the recursion

$$\|f\|_{U^1(G)} = \sum_{\mathbf{x}\in G} f(\mathbf{x}),$$

$$\|f\|_{U^{k+1}(G)} = \left(\sum_{\mathbf{h}\in G} \|\Delta(f, \mathbf{h})\|_{U^k(G)}^{2^k}\right)^{1/2^{k+1}}.$$

Clearly $\|\cdot\|_{U^1}$ is not positive. However for $k \geqslant 2$, one can show that $\|\cdot\|_{U^k}$ is indeed a norm, a fact whose justification affords us the opportunity to introduce our conventions for Fourier analysis on finite abelian groups. Notice that the $U^3$-norm is not defined for functions such as $b_A$, with domain equal to an infinite group. We must therefore find a finite abelian group $G$ to serve as a proxy for $\mathbb{Z}^n$. To this end, we use $G = \mathbb{Z}_p^n$, where $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ for some large prime $p$.

Recall the definition (1.4) of the Fourier transform of a function $f : \mathbb{Z}^n \to \mathbb{C}$. We emulate this definition for functions $f : \mathbb{Z}_p^n \to \mathbb{C}$ by setting

$$\hat{f}(\mathbf{a}) = \sum_{\mathbf{x}\in\mathbb{Z}_p^n} f(\mathbf{x})\mathrm{e}(-\mathbf{a}\cdot\mathbf{x}/p) \quad (\mathbf{a} \in \mathbb{Z}_p^n).$$

Here we are using the counting measure on $G = \mathbb{Z}_p^n$. Reversing the usual notational conventions of additive combinatorics, we use the uniform probability measure $\mathbb{P}$ on $\widehat{G} \cong \mathbb{Z}_p^n$. This makes the analogy with $G = \mathbb{Z}^n$ and $\widehat{G} \cong \mathbb{T}^n$ more transparent. Indeed, one can think of $\widehat{G} \cong \mathbb{Z}_p^n$ as a discrete approximation for $\mathbb{T}^n$ via the embedding $\mathbf{a} \mapsto \mathbf{a}/p$ (notice that we use the uniform probability measure on $\mathbb{T}^n$).

The integral of a function $g : \widehat{G} \to \mathbb{C}$ (one can think of $g = \hat{f}$ for some $f : \mathbb{Z}_p^n \to \mathbb{C}$) is then written using the expectation notation

$$\int_{\widehat{G}} g\, d\mathbb{P} = \mathop{\mathbb{E}}_{\mathbf{a}\in\widehat{G}} g(\mathbf{a}) = \frac{1}{p^n}\sum_{\mathbf{a}\in\mathbb{Z}_p^n} g(\mathbf{a}).$$

In particular, for $q \geqslant 1$ we have

$$\|g\|_q = \left(\mathop{\mathbb{E}}_{\mathbf{a}\in\widehat{G}} |g(\mathbf{a})|^q\right)^{1/q}.$$

One can then check the identity

$$\|f\|_{U^2} = \|\hat{f}\|_4. \tag{3.16}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

It follows from this and Fourier inversion that $\|\cdot\|_{U^2}$ is indeed a norm, and by the recursive definition, so is $\|\cdot\|_{U^k}$ for $k \geqslant 3$. The recursive definition also provides us with the useful identity

$$\|f\|_{U^3} = \left(\sum_{\mathbf{h}\in\mathbb{Z}_p^n} \|\widehat{\Delta(f,\mathbf{h})}\|_4^4\right)^{1/8}. \tag{3.17}$$

Given integer matrices $M_1, \dots, M_s$, sets $A_1, \dots, A_s \subset \mathbb{Z}^n$ and $\mathbf{c} \in \mathbb{Z}^n$ we have the orthogonality relations

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathbf{a}} \hat{1}_{A_1}(\mathbf{a}M_1)\cdots\hat{1}_{A_s}(\mathbf{a}M_s)\mathrm{e}(-\mathbf{a}\cdot\mathbf{c})\\ &\quad = \#\{(\mathbf{x}_1,\dots,\mathbf{x}_s)\in A_1\times\cdots\times A_s : M_1\mathbf{x}_1+\cdots+M_s\mathbf{x}_s \equiv \mathbf{c} \pmod p\}.\end{aligned} \tag{3.18}$$

With these notational conventions in hand, we are able to prove our analogue of the non-uniformity estimate in Lemma 2.5.

LEMMA 3.4. *There exist absolute constants $C = C(t, T_1, T_2)$ and $c = c(t, T_1, T_2)$ such that for any $\delta > 0$ and any real $L \geqslant 30n^2$, if $Q \subset \mathbb{R}^n$ is a half-open cube of side-length at least $L$ and $A \subset [Q]$ satisfies $|A| \geqslant \delta|[Q]|$ and*

$$R(A) \leqslant c\delta^2|A|^2, \tag{3.19}$$

*then either there exists a half-open cube $Q'$ of side-length at least $L/(12n)$ such that*

$$|A \cap Q'| \geqslant 2\delta|[Q']| \tag{3.20}$$

*or there exists a prime $p$ in the range $CL \leqslant p \leqslant 2CL$ such that*

$$\|b_A\|_{U^3(\mathbb{Z}_p^n)} \geqslant c\,\delta^{9/4}p^{n/2}. \tag{3.21}$$

*Here we regard $b_A$ as a function on $\mathbb{Z}_p^n$ via the natural pull-back.*

It makes little difference in applications, but the exponent $9/4$ in (3.21) can be improved to 2 when $T_1 - T_2$ is non-singular.

*Proof.* Let $m \in \mathbb{N}$ denote the largest absolute value of any entry in $T_1$ and $T_2$. Increasing $m$ if necessary, we may assume that

$$m > \max\{|\det T_1|, |\det T_2|, |\det S_1|, |\det S_2|\}. \tag{3.22}$$

Notice that $m \ll_{T_1,T_2} 1$. By Bertrand's postulate, there exists a prime $p$ in the range

$$(mn+t)L < p \leqslant 2(mn+t)L. \tag{3.23}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

It follows from (3.22) that $T_1$, $T_2$, $S_1$ and $S_2$ are all invertible over $\mathbb{F}_p$. The size bound (3.23) certainly implies that $t \in \mathbb{Z}_p \setminus \{0\}$, moreover it implies that for any $\mathbf{x}, \mathbf{y}, \mathbf{z}_1, \mathbf{z}_2 \in [Q]$ we have

$$S_i\mathbf{x} - T_i\mathbf{y} + t\mathbf{z}_i = 0,$$

if and only if

$$S_i\mathbf{x} - T_i\mathbf{y} + t\mathbf{z}_i \equiv 0 \mod p. \tag{3.24}$$

Given $A_i \subset \mathbb{Z}^n$ $(1 \leqslant i \leqslant 4)$, let $R_p(A_1, A_2, A_3, A_4)$ denote the number of quadruples $(\mathbf{x}, \mathbf{y}, \mathbf{z}_1, \mathbf{z}_2) \in A_1 \times A_2 \times A_3 \times A_4$ which satisfy (3.24) for $i = 1, 2$. Then for $A \subset [Q]$ we have

$$R_p([Q], [Q], A, A) = R([Q], [Q], A, A) \quad \text{and} \quad R_p(A) = R(A).$$

Let us assume that for every half-open cube $Q'$ of side-length at least $L/(12n)$ we have $|A \cap Q'| < 2\delta|[Q']|$ and let $c_0$ denote the constant in Lemma 3.3. Then we have

$$R_p([Q], [Q], A, A) \geqslant c_0|A|^2.$$

Let $c_1 := c_0/2$ and take $c \leqslant c_1$ in (3.19). Then we obtain the bound

$$|R_p(A) - \delta^2 R_p([Q], [Q], A, A)| \geqslant c_1\delta^2|A|^2. \tag{3.25}$$

From the orthogonality relations (3.18) one can decompose the difference $R_p(A) - \delta^2 R_p([Q], [Q], A, A)$ into the sum of two integrals

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathbf{a},\mathbf{b}} \hat{b}_A(\mathbf{a}S_1 + \mathbf{b}S_2)\hat{1}_A(-\mathbf{a}T_1 - \mathbf{b}T_2)\hat{1}_A(t\mathbf{a})\hat{1}_A(t\mathbf{b})\\ &\quad + \mathop{\mathbb{E}}_{\mathbf{a},\mathbf{b}} \delta\hat{1}_{[Q]}(\mathbf{a}S_1 + \mathbf{b}S_2)\hat{b}_A(-\mathbf{a}T_1 - \mathbf{b}T_2)\hat{1}_A(t\mathbf{a})\hat{1}_A(t\mathbf{b}).\end{aligned} \tag{3.26}$$

Using the identity $\hat{1}_A = b_A + \delta\hat{1}_{[Q]}$, one can decompose the first term in (3.26) further into the sum

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathbf{a},\mathbf{b}} \hat{b}_A(\mathbf{a}S_1 + \mathbf{b}S_2)\hat{b}_A(-\mathbf{a}T_1 - \mathbf{b}T_2)\hat{1}_A(t\mathbf{a})\hat{1}_A(t\mathbf{b})\\ &\quad + \mathop{\mathbb{E}}_{\mathbf{a},\mathbf{b}} \hat{b}_A(\mathbf{a}S_1 + \mathbf{b}S_2)\delta\hat{1}_{[Q]}(-\mathbf{a}T_1 - \mathbf{b}T_2)\hat{1}_A(t\mathbf{a})\hat{1}_A(t\mathbf{b}).\end{aligned}$$

It follows that there exist functions $f_1, f_2 \in \{b_A, \delta 1_{[Q]}\}$, at least one of which is equal to $b_A$, such that

$$\left|\mathop{\mathbb{E}}_{\mathbf{a},\mathbf{b}} \hat{f}_1(\mathbf{a}S_1 + \mathbf{b}S_2)\hat{f}_2(-\mathbf{a}T_1 - \mathbf{b}T_2)\hat{1}_A(\mathbf{a})\hat{1}_A(\mathbf{b})\right| \geqslant c_1\delta^2|A|^2/3. \tag{3.27}$$

Using the identity

$$\left(\mathop{\mathbb{E}}_{\mathbf{b}} |\hat{1}_A(t\mathbf{b})|^2\right)^{1/2} = |A|^{1/2},$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

together with the Cauchy–Schwarz inequality, one sees that the quantity $c_1^2\delta^4|A|^3/9$ is bounded above by

$$\mathop{\mathbb{E}}_{\mathbf{a},\mathbf{a}',\mathbf{b}} \hat{f}_1(\mathbf{a}S_1+\mathbf{b}S_2)\hat{f}_1(-\mathbf{a}'S_1-\mathbf{b}S_2)\hat{f}_2(-\mathbf{a}T_1-\mathbf{b}T_2) \times \hat{f}_2(\mathbf{a}'T_1+\mathbf{b}T_2)\hat{1}_A(t\mathbf{a})\hat{1}_A(-t\mathbf{a}'). \tag{3.28}$$

On considering the underlying equations in (3.28), one obtains the estimate

$$c_1^2\delta^4|A|^3/9 \leqslant \sum_{\substack{S_1\mathbf{x}-T_1\mathbf{y}+t\mathbf{z}=0\\ S_1\mathbf{x}'-T_1\mathbf{y}'+t\mathbf{z}'=0\\ S_2(\mathbf{x}'-\mathbf{x})=T_2(\mathbf{y}'-\mathbf{y})}} f_1(\mathbf{x})f_1(\mathbf{x}')f_2(\mathbf{y})f_2(\mathbf{y}')1_A(\mathbf{z})1_A(\mathbf{z}'). \tag{3.29}$$

(Here the identities over which we sum hold in $\mathbb{Z}_p^n$.)

Given $\mathbf{k}\in\mathbb{Z}_p^n$ let us define

$$\mathbf{k}' = T_2^{-1}S_2\mathbf{k} \quad \text{and} \quad \mathbf{k}'' = (T_2-T_1)T_2^{-1}\mathbf{k}.$$

Using the change of variable $\mathbf{x}'=\mathbf{x}+\mathbf{k}$, one can check that (3.29) is equivalent to

$$\begin{aligned} c_1^2\delta^4|A|^3/9 &\leqslant \sum_{\mathbf{k}}\sum_{S_1\mathbf{x}-T_1\mathbf{y}+t\mathbf{z}=0} \Delta(f_1,\mathbf{k})(\mathbf{x})\Delta(f_2,\mathbf{k}')(\mathbf{y})\Delta(1_A,\mathbf{k}'')(\mathbf{z}) \\ &= \sum_{\mathbf{k}}\mathop{\mathbb{E}}_{\mathbf{a}}\widehat{\Delta(f_1,\mathbf{k})}(\mathbf{a}S_1)\widehat{\Delta(f_2,\mathbf{k}')}(-\mathbf{a}T_1)\widehat{\Delta(1_A,\mathbf{k}'')}(t\mathbf{a}). \end{aligned} \tag{3.30}$$

By Hölder's inequality and the non-singularity of $S_1$ and $T_1$ over $\mathbb{F}_p$, the right-hand side of (3.30) is at most

$$\begin{aligned} &\sum_{\mathbf{k}}\|\widehat{\Delta(f_1,\mathbf{k})}\|_4\|\widehat{\Delta(f_2,\mathbf{k}')}\|_4\|\widehat{\Delta(1_A,\mathbf{k}'')}\|_2 \\ &\quad\leqslant \left(\sum_{\mathbf{k}}\|\widehat{\Delta(f_1,\mathbf{k})}\|_4^4\right)^{1/4}\left(\sum_{\mathbf{k}}\|\widehat{\Delta(f_2,\mathbf{k}')}\|_4^4\right)^{1/4}\left(\sum_{\mathbf{k}}\|\widehat{\Delta(1_A,\mathbf{k}'')}\|_2^2\right)^{1/2}. \end{aligned} \tag{3.31}$$

Using $\Delta(1_A,\mathbf{k}'')\leqslant 1_A$ we have the estimate†

$$\left(\sum_{\mathbf{k}}\|\widehat{\Delta(1_A,\mathbf{k}'')}\|_2^2\right)^{1/2} \leqslant p^{n/2}|A|^{1/2}. \tag{3.32}$$

Since $T_2^{-1}S_2$ is non-singular over $\mathbb{F}_p$, the map $\mathbf{k}\mapsto\mathbf{k}'$ is a permutation of $\mathbb{Z}_p^n$. Combining these facts with (3.31) we obtain

† The right-hand side of this inequality can be improved to $|A|$ if $T_1-T_2$ is assumed to be non-singular.

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

$$\frac{c_1\delta^2|A|^{5/4}}{3p^{n/4}} \leqslant \left(\sum_{\mathbf{k}} \|\widehat{\Delta(f_1,\mathbf{k})}\|_4^4\right)^{1/8}\left(\sum_{\mathbf{k}} \|\widehat{\Delta(f_2,\mathbf{k})}\|_4^4\right)^{1/8}$$
$$= \|f_1\|_{U^3}\|f_2\|_{U^3}.$$

Our choice of $p$ ensures that $|A| \gg_{t,T_1,T_2} \delta p^n$. Thus

$$\delta^{13/4}p^n \ll_{t,T_1,T_2} \|f_1\|_{U^3}\|f_2\|_{U^3}.$$

We know that $\{f_1, f_2\} = \{b_A, \delta 1_{[Q]}\}$ or $\{f_1, f_2\} = \{b_A\}$. Hence either

$$\delta^{9/4}|G|^{1/2} \ll_{t,T_1,T_2} \|b_A\|_{U^3} \quad \text{or} \quad \delta^{13/8}|G|^{1/2} \ll_{t,T_1,T_2} \|b_A\|_{U^3}.$$

Clearly the former lower bound is the inferior of the two, from which the lemma now follows on taking $c$ sufficiently small. □

3.2. *$U^3$-bias yields a density increment.* In §2 we established that if $b_A$ has a large inner product with a linear phase, then $A$ has a density increment on the affine image of a large cube. The linear correlation of $b_A$ was a direct consequence of the largeness of $\|\hat{b}_A\|_\infty$. Gowers [3] was the first to observe that if $\|f\|_{U^3}$ is large, then $f$ has large inner product with a quasi-quadratic phase. This result was generalized and strengthened considerably by Green and Tao [6], and we employ their result in this section. Given this quadratic correlation of $b_A$, we then run an argument similar to that found in §2 to show that $A$ has density increment on the affine image of a large cube.

The inverse theorem for the $U^3$-norm of Green and Tao is somewhat technical to state, and we begin by making the required definitions.

*Definition* 3.5 (Regular Bohr sets). Let $S \subset \mathbb{Z}_p^n$, $|S| = d$ and $\rho > 0$. The Bohr set $B(S, \rho)$ is defined by

$$B(S, \rho) = \{\mathbf{x} \in \mathbb{Z}_p^n : \|\mathbf{x} \cdot \mathbf{a}/p\| \leqslant \rho \text{ for all } \mathbf{a} \in S\}.$$

The quantities $d$ and $\rho$ are called, respectively, the *dimension* and *width* of $B(S, \rho)$. We say $B(S, \rho)$ is *regular* if for all $|\sigma| \leqslant 1/100d$ we have

$$\left|\frac{|B(S, (1+\sigma)\rho)|}{|B(S, \rho)|} - 1\right| \leqslant 100d|\sigma|.$$

*Definition* 3.6 (Local quadratic phase). Let $G$ be an abelian group and $B \subset G$ a non-empty subset of $G$. We say $\phi : B \to \mathbb{T}$ is a *local linear phase* on $B$ if it is a Freiman 2-homomorphism from $B$ to $\mathbb{T}$. In other words, for any $\mathbf{x}, \mathbf{x}', \mathbf{y}$, $\mathbf{y}' \in B$ with $\mathbf{x} - \mathbf{x}' = \mathbf{y} - \mathbf{y}'$ we have $\phi(\mathbf{x}) - \phi(\mathbf{x}') = \phi(\mathbf{y}) - \phi(\mathbf{y}')$. We say $\phi : B \to \mathbb{T}$ is a *local quadratic phase* on $B$ if for any $\mathbf{h} \in G$, the discrete derivative $\mathbf{x} \mapsto \phi(\mathbf{x} + \mathbf{h}) - \phi(\mathbf{x})$ is a local linear phase on $B \cap (B - \mathbf{h})$.

The following is a special case of the inverse theorem for the $U^3$-norm, as established by Green and Tao [6].

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

THEOREM 3.7 (Green and Tao [6]). *Let $p$ be an odd prime, let $f : \mathbb{Z}_p^n \to \mathbb{C}$ be a 1-bounded function ($\|f\|_\infty \leqslant 1$) and let $0 < \eta \leqslant 1$. Suppose that*

$$\|f\|_{U^3} \geqslant \eta p^{n/2}.$$

*Then there exists a regular Bohr set $B = B(S, \rho)$ with $|S| \leqslant (2/\eta)^C$, $\rho \geqslant (\eta/2)^C$ and a local quadratic phase $\phi : B \to \mathbb{T}$ such that*

$$\sum_{\mathbf{y} \in \mathbb{Z}_p^n} \left| \sum_{\mathbf{x} \in \mathbf{y}+B} f(\mathbf{x}) e(\phi(\mathbf{x})) \right| \geqslant (\eta/2)^C p^n |B|.$$

*Moreover, $C = 2^{24}$ is permissible in the above.*

Since we are content to aim for a denominator in (1.2) which is polynomial in $\log\log N$, we can ease our exposition by replacing the regular Bohr set obtained above by the affine image of a half-open cube. We note that one could work much harder, utilising the methods of [7], along with the Bohr set afforded by the above theorem, to improve the denominator in (1.2).

The following lemma establishes that a Bohr set contains the dilation of a large cube.

LEMMA 3.8. *Let $p$ be a prime and $B(S, \rho) \subset \mathbb{Z}_p^n$ a Bohr set of dimension $d = |S|$ and width $0 < \rho \leqslant 1/2$. Set $M = \lfloor p^{1/nd} \rfloor$. Then there exists $q \in \mathbb{Z}_p \setminus \{0\}$ such that we have the containment*

$$q \cdot [\pm \rho M/n]^n \subset B(S, \sigma). \tag{3.33}$$

*Moreover, all elements of $q \cdot (\pm \rho M/n]^n$ are distinct modulo $p$.*

*Proof.* Let $S = \{\mathbf{a}_1, \dots, \mathbf{a}_d\}$ and define

$$\boldsymbol{\alpha} = p^{-1}(\mathbf{a}_1, \dots, \mathbf{a}_d) \in \mathbb{T}^{nd}.$$

Consider the homomorphism $\mathbb{Z}_p \to \mathbb{T}^{nd}$ given by

$$q \mapsto q\boldsymbol{\alpha}.$$

Let us partition the torus $\mathbb{T}^{nd}$ into $M^{nd}$ half-open cubes of the form

$$\mathbf{m} + (0, 1/M]^{nd} \quad (\mathbf{m} \in \{0, 1, \dots, M-1\}^{nd}).$$

Since $p \geqslant M^{nd}$, either (a) every such cube contains $q\boldsymbol{\alpha}$ for some $q \in \mathbb{Z}_p$ or (b) some cube contains both $q_1\boldsymbol{\alpha}$ and $q_2\boldsymbol{\alpha}$ for distinct $q_1, q_2 \in \mathbb{Z}_p$. In case (a) we certainly have some $q \in \mathbb{Z}_p$ with $q\boldsymbol{\alpha} \in (0, 1/M]^{nd}$. We can only have $q = 0$ if $M = 1$, however if $M = 1$ then we also have $1.\boldsymbol{\alpha} \in (0, 1/M]^{nd}$. In case (b), letting $q = q_1 - q_2$ or $q = q_2 - q_1$, we also see that there exists $q \in \mathbb{Z}_p \setminus \{0\}$ with

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

$q\boldsymbol{\alpha} \in (0, 1/M]^{nd}$. Fix such a $q$. It follows that for any $\mathbf{x} \in \mathbb{Z}^n$ with $|\mathbf{x}|_\infty \leqslant \rho M/n$ we have the estimates

$$\|q\mathbf{x} \cdot \mathbf{a}_i/p\| \leqslant \rho \quad (1 \leqslant i \leqslant d).$$

Thus $q\mathbf{x} \in B(S, \rho)$ and (3.33) follows.

It remains to show each $q\mathbf{x}$ with $\mathbf{x} \in (\pm\rho M/n]^n$ is distinct modulo $p$. This follows from the fact that $q$ is non-zero modulo $p$, that the side-length of $Q = (-\rho M/n, \rho M/n]^n$ is $2\rho M/n \leqslant p$ and that the cube $Q$ is half-open. □

We also need information on the structure of local quadratic phases. Fortunately, these correspond to actual quadratic phases when restricted to affine images of cubes. This fact is a consequence of Lemma 10.6 of Green and Tao [6].

LEMMA 3.9 (Green and Tao [6]). *Let $\mathbf{y} \in \mathbb{Z}_p^n$, $q \in \mathbb{Z}_p \setminus \{0\}$ and $M \geqslant 1$. For every local quadratic phase $\phi : \mathbf{y} + q \cdot (\pm M]^n \to \mathbb{T}^n$ there exist*

$$c, \eta_i, \lambda_{ij} \in \mathbb{T} \quad (1 \leqslant i \leqslant j \leqslant n)$$

*satisfying*

$$\phi(\mathbf{y} + q \cdot \mathbf{x}) = \sum_{1 \leqslant i \leqslant j \leqslant n} \lambda_{ij} x_i x_j + \sum_{1 \leqslant i \leqslant n} \eta_i x_i + c \quad (\mathbf{x} \in (\pm M]^n).$$

We now combine the above results to give a suitably modified version of the $U^3$-inverse theorem.

THEOREM 3.10. *Let $p$ be an odd prime, $C = 2^{24}$, let $f : \mathbb{Z}_p^n \to \mathbb{C}$ be a 1-bounded function ($\|f\|_\infty \leqslant 1$) and let $0 < \eta \leqslant 1$. Suppose that*

$$\|f\|_{U^3} \geqslant \eta p^{n/2}.$$

*Then there exists $q \in \mathbb{Z}_p \setminus \{0\}$ and $L \gg_n \eta^{3C} p^{(\eta/2)^C/n}$ such that for each $\mathbf{y} \in \mathbb{Z}_p^n$ we can find*

$$\xi, \eta_i, \lambda_{ij} \in \mathbb{T} \quad (1 \leqslant i \leqslant j \leqslant n)$$

*so that on setting*

$$\phi_{\mathbf{y}}(\mathbf{x}) = \sum_{1 \leqslant i \leqslant j \leqslant n} \lambda_{ij} x_i x_j + \sum_{1 \leqslant i \leqslant n} \eta_i x_i + \xi \quad (\mathbf{x} \in (\pm L]^n),$$

*we have*

$$\sum_{\mathbf{y} \in \mathbb{Z}_p^n} \left| \sum_{\mathbf{x} \in (\pm L]^n} f(\mathbf{y} + q\mathbf{x}) e(\phi_{\mathbf{y}}(\mathbf{x})) \right| \gg (\eta/2)^C p^n |(\pm L]^n|.$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*Proof.* Assuming that $\|f\|_{U^3} \geqslant \eta p^{n/2}$, let $B = B(S, \rho)$ and $\phi : B \to \mathbb{T}$ denote the Bohr set and local quadratic phase guaranteed by Theorem 3.7. Let $d = |S|$, $M = \lfloor p^{1/nd} \rfloor$ and let $\varepsilon$ denote a small positive constant to be determined later. Clearly we may assume $\rho \leqslant 1/2$, so by Lemma 3.8 there exists $q \in \mathbb{Z}_p \setminus \{0\}$ such that $B(S, \varepsilon\rho)$ contains the dilate $P = q \cdot (\pm\varepsilon\rho M/n]^n \subset \mathbb{Z}_p^n$. Let $g$ denote the 1-bounded function $\mathbf{x} \mapsto f(\mathbf{x})\mathrm{e}(\phi(\mathbf{x}))$. Then for each $\mathbf{y} \in \mathbb{Z}_p^n$ we have

$$\begin{aligned}\left|\sum_{\mathbf{x}\in B+\mathbf{y}} g(\mathbf{x}) - \frac{1}{|P|}\sum_{\mathbf{x}\in B+\mathbf{y}}\sum_{\mathbf{z}\in P+\mathbf{x}} g(\mathbf{z})\right| &\leqslant \frac{1}{|P|}\sum_{\mathbf{z}\in P}\left|\sum_{\mathbf{x}\in B+\mathbf{y}} g(\mathbf{x}) - \sum_{\mathbf{x}\in B+\mathbf{y}+\mathbf{z}} g(\mathbf{x})\right| \\ &\leqslant \frac{1}{|P|}\sum_{\mathbf{z}\in P}|(B+\mathbf{y})\Delta(B+\mathbf{y}+\mathbf{z})|.\end{aligned}$$

For each $\mathbf{z} \in P \subset B(S, \varepsilon\rho)$ we have

$$B(S, (1-\varepsilon)\rho) \subset B + \mathbf{z} \subset B(S, (1+\varepsilon)\rho).$$

Ensuring that $\varepsilon \leqslant 1/(100d)$, regularity gives for any $\mathbf{y} \in \mathbb{Z}_p^n$ the estimate

$$\begin{aligned}|(B+\mathbf{y})\Delta(B+\mathbf{y}+\mathbf{z})| = |B\Delta(B+\mathbf{z})| &= |B \setminus (B+\mathbf{z})| + |(B+\mathbf{z}) \setminus B| \\ &\leqslant |B(S,\rho) \setminus B(S,(1-\varepsilon)\rho)| \\ &\quad + |B(S,(1+\varepsilon)\rho) \setminus B(S,\rho)| \\ &\leqslant 200d\varepsilon|B|.\end{aligned}$$

It follows that

$$\begin{aligned}\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\left|\sum_{\mathbf{x}\in B+\mathbf{y}} g(\mathbf{x})\right| &\leqslant |P|^{-1}\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\sum_{\mathbf{x}\in B+\mathbf{y}}\left|\sum_{\mathbf{z}\in P+\mathbf{x}} g(\mathbf{z})\right| + 200d\varepsilon p^n|B| \\ &= \frac{|B|}{|P|}\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\left|\sum_{\mathbf{z}\in P+\mathbf{y}} g(\mathbf{z})\right| + 200d\varepsilon p^n|B|. \qquad (3.34)\end{aligned}$$

Take

$$\varepsilon = \frac{(\eta/2)^C}{400d} \geqslant \frac{(\eta/2)^{2C}}{400}.$$

Since the left-hand side of (3.34) is at least $(\eta/2)^C p^n|B|$, and $200d\varepsilon = (\eta/2)^C/2$, we have

$$\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\left|\sum_{\mathbf{z}\in P+\mathbf{y}} g(\mathbf{z})\right| \geqslant \frac{1}{2}(\eta/2)^C p^n|P|. \qquad (3.35)$$

By Lemma 3.8, the map which takes $\mathbf{x} \in (\pm\varepsilon\rho M/n]^n$ to $q\mathbf{x} \in \mathbb{Z}_p^n$ is injective, hence (3.35) is equivalent to

$$\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\left|\sum_{\mathbf{x}\in(\pm\varepsilon\rho M/n]^n} f(\mathbf{y}+q\mathbf{x})\mathrm{e}(\phi(\mathbf{y}+q\mathbf{x}))\right| \geqslant \frac{1}{2}(\eta/2)^C p^n|(\pm\varepsilon\rho M/n]^n|.$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

That we can replace $\phi(\mathbf{y}+q\mathbf{x})$ by $\phi_{\mathbf{y}}(\mathbf{x})$ in each inner sum is a consequence of Lemma 3.9. Finally, set $L=\varepsilon\rho M/n$. Then $d\leqslant(\eta/2)^C$ implies that $M=\lfloor p^{1/nd}\rfloor\gg p^{(2/\eta)^C/n}$ and $\varepsilon\rho/n\gg_n(2/\eta)^{3C}$, from which $L\gg_n\eta^{-3C}p^{(2/\eta)^C/n}$ follows. □

The final ingredient required to convert Theorem 3.10 into a proof of Lemma 3.2 is a result of W. M. Schmidt on simultaneous Diophantine approximation.

THEOREM 3.11 (Schmidt [12]). *Let $\alpha_1,\dots,\alpha_d$ be reals. For any $X\geqslant 1$ and $\varepsilon>0$ there exists $1\leqslant q\leqslant X$ such that*

$$\|q^2\alpha_i\|\ll_{d,\varepsilon}X^{\varepsilon-1/(d^2+d)}\quad\text{for } i=1,\dots,d. \tag{3.36}$$

We are now in a position to prove the density increment lemma.

*Proof of Lemma 3.2.* Let us suppose $L\geqslant\exp(C\delta^{-C})$ and $R(A)\leqslant c\delta^2|A|^2$, with $c$ and $C$ to be determined later. We can certainly take $C$ large enough to ensure that $L\geqslant 30n^2$. Let $c_1$ denote the small constant in Lemma 3.4. Taking $c\leqslant c_1$, the conditions of Lemma 3.4 are satisfied. This lemma guarantees one of two possibilities. The first is that there exists a half-open cube $Q'$ of side-length $L'\geqslant L/(6n)$ such that

$$|A\cap Q'|\geqslant 2\delta|[Q']|. \tag{3.37}$$

Ensuring that $C$ is sufficiently large in terms of $n$, we easily attain $L/(6n)\geqslant L^{c\delta^C}$. Taking $L_1=L/(6n)$, $Q_1=Q'$, $\mathbf{r}=0$ and $q=1$ we obtain (3.5) and (3.6) and are done.

Let us therefore assume that the second possibility of Lemma 3.4 holds. Letting $C_1$ denote the large constant in Lemma 3.4, we see that there exists a prime $p$ in the range $C_1L<p\leqslant 2C_1L$ such that

$$\|b_A\|_{U^3(\mathbb{Z}_p^n)}\geqslant c_1\delta^{9/4}p^{n/2}. \tag{3.38}$$

Let $m=2^{24}$ and $\eta=c_1\delta^{9/4}$. Applying Theorem 3.10 and decreasing $c$ if necessary, we can find $q\in\mathbb{Z}_p\setminus\{0\}$ and $L_0$ satisfying

$$L_0\gg_n\eta^{3m}p^{(\eta/2)^m/n} \tag{3.39}$$

and for each $\mathbf{y}\in\mathbb{Z}_p^n$ we can find $\xi$, $\eta_i$ $\lambda_{ij}\in\mathbb{T}$ $(1\leqslant i\leqslant j\leqslant n)$ such that on setting

$$\phi_{\mathbf{y}}(\mathbf{x})=\sum_{1\leqslant i\leqslant j\leqslant n}\lambda_{ij}x_ix_j+\sum_{1\leqslant i\leqslant n}\eta_ix_i+\xi,$$

we have

$$\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\left|\sum_{\mathbf{x}\in(\pm L_0]^n}b_A(\mathbf{y}+q\mathbf{x})\mathrm{e}(\phi_{\mathbf{y}}(\mathbf{x}))\right|\gg(\eta/2)^mp^n|(\pm L_0]^n|. \tag{3.40}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

One can certainly decrease $c$ further in order to ensure that $(\eta/2)^m/n \geqslant 2c\delta^{9m/4}$. Thus using (3.39) we have

$$\begin{aligned} L_0 &\gg_n \eta^{3m} p^{(\eta/2)^m/n} \\ &\gg_{t,T_1,T_2} \delta^{27m/4} p^{2c\delta^{9m/4}} \\ &\gg_{t,T_1,T_2} \delta^{27m/4} L^{2c\delta^{9m/4}}. \end{aligned}$$

Let $c_2$ denote the product of all the implicit constants appearing in the above string of inequalities. Provided we choose $C$ large enough, the lower bound $L \geqslant \exp(C\delta^{-C})$ certainly gives us

$$L_0 \geqslant c_2\delta^{27m/4} L^{2c\delta^{9m/4}} \geqslant L^{c\delta^{9m/4}}.$$

Fix $\mathbf{y} \in \mathbb{Z}_p^n$ and let $\xi, \eta_i, \lambda_{ij}$ denote the corresponding coefficients of $\phi_{\mathbf{y}}$. Let $d$ denote the number of $\lambda_{ij}$ $(1 \leqslant i \leqslant j \leqslant n)$, so $d = n(n+1)/2$. One can check that $d^2 + d < 3n^4$. By Schmidt's theorem there exists $q_0 \in \mathbb{N}$ satisfying

$$q_0 \leqslant L_0^{6n^4/(6n^4+1)}, \quad \|q_0^2\lambda_{ij}\| \ll_n L_0^{-2/(6n^4+1)} \quad \text{for } 1 \leqslant i \leqslant j \leqslant n. \tag{3.41}$$

Set $Q_0 = (-L_0, L_0]^n$ and let us partition $[Q_0]$ into congruence classes modulo $q_0$. Each class takes the form

$$\mathbf{r} + q_0 \cdot [Q_0(\mathbf{r})],$$

where $Q_0(\mathbf{r})$ is a half-open cube of side-length $2L_0/q_0$. Let $m_0$ be an integer satisfying

$$\frac{L_0^{6n^4/(6n^4+1)} \log L_0}{q_0} \leqslant m_0 \leqslant \frac{2L_0^{6n^4/(6n^4+1)} \log L_0}{q_0}. \tag{3.42}$$

We can partition each $Q_0(\mathbf{r})$ into $m_0^n$ half-open cubes $Q_1(\mathbf{r}), Q_2(\mathbf{r}), \ldots$ of side-length $L_1 = 2L_0/(q_0m_0)$. For each $t$ and $\mathbf{r}$ let us pick a fixed element $\mathbf{a} = \mathbf{a}(t, \mathbf{r}) \in [Q_t(\mathbf{r})]$ and set

$$B_t(\mathbf{r}) = Q_t(\mathbf{r}) - \mathbf{a}.$$

Note that $B_t(\mathbf{r})$ is a half-open cube of side-length $L_1$ and that

$$B_t(\mathbf{r}) \subset [-L_1, L_1]^n.$$

For fixed $t$ and $\mathbf{r}$, let us define

$$\mu_i = \mu_i(t, \mathbf{r}) = q_0\Bigg(\eta_i + \sum_{j \geqslant i} \lambda_{ij}(r_j + q_0a_j) + \sum_{j \leqslant i} \lambda_{ji}(r_j + q_0a_j)\Bigg).$$

By Kronecker's theorem on Diophantine approximation, we can find $q_1 \in \mathbb{N}$ satisfying

$$q_1 \leqslant L_1^{n/(n+1)} \quad \text{and} \quad \|q_1\mu_i\| \leqslant L_1^{-1/(n+1)} \quad \text{for } i = 1, \ldots, n. \tag{3.43}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

Let us partition $B_t(\mathbf{r})$ into congruence classes modulo $q_1$, each taking the form

$$\mathbf{r}' + q_1 \cdot [B_t(\mathbf{r}, \mathbf{r}')],$$

for some half-open cube $B_t(\mathbf{r}, \mathbf{r}')$ of side-length $L_1/q_1$. Let $m_1 \in \mathbb{N}$ with

$$\frac{L_1^{n/(n+1)} \log L_1}{q_1} \leqslant m_1 \leqslant \frac{2L_1^{n/(n+1)} \log L_1}{q_1}.$$

Now let us further partition $B_t(\mathbf{r}, \mathbf{r}')$ into $m_1^n$ half-open cubes

$$Q_{t,1}(\mathbf{r}, \mathbf{r}'), Q_{t,2}(\mathbf{r}, \mathbf{r}'), \dots$$

each of side-length $L_2 = L_1/(q_1 m_1)$. Decreasing $c$ if necessary, we can bound $L_2$ from below by

$$\begin{aligned} L_2 = \frac{L_1}{q_1 m_1} &\geqslant \frac{L_1^{1/(n+1)}}{2 \log L_1} \\ &\gg \frac{L_0^{1/((n+1)(6n^4+1))}}{\log L_0 \log L_1} \gg L^{c\delta^{9m/4}}. \end{aligned} \tag{3.44}$$

Let $\mathbf{z}, \mathbf{z}'$ denote arbitrary elements of $[Q_{t,t'}(\mathbf{r}, \mathbf{r}')]$. Set

$$\begin{aligned} \mathbf{x} &= \mathbf{r} + q_0(\mathbf{a} + \mathbf{r}' + q_1 \mathbf{z}), \\ \mathbf{x}' &= \mathbf{r} + q_0(\mathbf{a} + \mathbf{r}' + q_1 \mathbf{z}'). \end{aligned}$$

Notice that $\mathbf{x} = \mathbf{c} + q_0 \mathbf{b}$, $\mathbf{x}' = \mathbf{c} + q_0 \mathbf{b}'$ where $\mathbf{b}, \mathbf{b}' \in B_t(\mathbf{r}) \subset [-L_1, L_1]^n$. Also

$$\begin{aligned} &(c_i + q_0 b_i)(c_j + q_0 b_j) - (c_i + q_0 b_i')(c_j + q_0 b_j') \\ &\quad = q_0^2 (b_i b_j - b_i' b_j') + q_0 c_i (b_j - b_j') + q_0 c_j (b_i - b_i'). \end{aligned}$$

Thus

$$\begin{aligned} \phi_{\mathbf{y}}(\mathbf{x}) - \phi_{\mathbf{y}}(\mathbf{x}') &= \sum_{1 \leqslant i \leqslant j \leqslant n} q_0^2 \lambda_{ij} (b_i b_j - b_i' b_j') + \sum_{1 \leqslant i \leqslant n} \mu_i (b_i - b_j) \\ &= \sum_{1 \leqslant i \leqslant j \leqslant n} q_0^2 \lambda_{ij} (b_i b_j - b_i' b_j') + \sum_{1 \leqslant i \leqslant n} q_1 \mu_i (z_i - z_j). \end{aligned}$$

Using the estimates (3.41) and (3.43), we therefore have

$$\begin{aligned} |\mathrm{e}(\phi_{\mathbf{y}}(\mathbf{x})) - \mathrm{e}(\phi_{\mathbf{y}}(\mathbf{x}'))| &\ll_n L_0^{-2/(6n^4+1)} L_1^2 + L_1^{-1/(n+1)} L_2 \\ &\leqslant \frac{1}{(\log L_0)^2} + \frac{1}{\log L_1} \\ &\ll_n \frac{1}{\log L_0}. \end{aligned}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

By (3.39) we have

$$\frac{1}{\log L_0} \ll_{t,T_1,T_2} \frac{\delta^{9m/4}}{\log L}.$$

Provided we take $C$ large enough, the estimate $L \geqslant \exp(C\delta^{-C})$ ensures that

$$|\mathrm{e}(\phi_{\mathbf{y}}(\mathbf{x})) - \mathrm{e}(\phi_{\mathbf{y}}(\mathbf{x}'))| \leqslant \frac{1}{(\log L)^{1/2}}.$$

Let $g_{\mathbf{y}}(t, \mathbf{r}, t', \mathbf{r}') = \mathbf{y} + q(\mathbf{r} + q_0(\mathbf{a} + \mathbf{r}'))$ and $q' = qq_1q_2$. Then for each $\mathbf{y} \in \mathbb{Z}_p^n$, we have

$$\left|\sum_{\mathbf{x}\in[Q_0]} b_A(\mathbf{y} + q\mathbf{x})\mathrm{e}(\phi_{\mathbf{y}}(\mathbf{x}))\right| = \sum_{\mathbf{r},t}\sum_{\mathbf{r}',t'}\left|\sum_{\mathbf{z}\in[Q_{t,t'}(\mathbf{r},\mathbf{r}')]} b_A(g_{\mathbf{y}}(t, \mathbf{r}, t', \mathbf{r}') + q'\mathbf{z})\right| + O_{t,T_1,T_2}(|[Q_0]|(\log L)^{-1/2}).$$

For each $\mathbf{y} \in \mathbb{Z}_p^n$ let $\mathcal{C}_{\mathbf{y}}$ denote the set cubes of the form

$$g_{\mathbf{y}}(t, \mathbf{r}, t', \mathbf{r}') + q' \cdot [Q_{t,t'}(\mathbf{r}, \mathbf{r}')].$$

Then $\mathcal{C}_{\mathbf{y}}$ partitions $[Q_0]$ into disjoint sets, each of which is the affine image of a half-open cube of side-length of order $\gg L^{c\delta^{9m/4}}$. We also know there exist absolute constants $C_4 \ll_{t,T_1,T_2} 1$ and $c_4(t, T_1, T_2) \gg_{t,T_1,T_2} 1$ such that

$$\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\sum_{P\in\mathcal{C}_{\mathbf{y}}}\left|\sum_{\mathbf{x}\in P} b_A(\mathbf{x})\right| + \frac{C_4 p^n|[Q_0]|}{(\log L)^{1/2}} \geqslant c_4\delta^{9m/4}p^n|[Q_0]|.$$

Decreasing $c$ if necessary and using $L \geqslant \exp(C/\delta^C)$ we have

$$\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\sum_{P\in\mathcal{C}_{\mathbf{y}}}\left|\sum_{\mathbf{x}\in P} b_A(\mathbf{x})\right| \gg_{t,T_1,T_2} \delta^{9m/4}p^n|[Q_0]|. \tag{3.45}$$

Let $\mathcal{C}_{\mathbf{y}}^+$ denote the set of $P \in \mathcal{C}_{\mathbf{y}}$ for which

$$\sum_{\mathbf{x}\in P} b_A(\mathbf{x}) \geqslant 0.$$

Since $b_A$ has mean zero on $\mathbb{Z}_p^n$, we can add $\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\sum_{P\in\mathcal{C}_{\mathbf{y}}}\sum_{\mathbf{x}\in P} b_A(\mathbf{x})$ to the left-hand side of (3.45) to obtain

$$\sum_{\mathbf{y}\in\mathbb{Z}_p^n}\sum_{P\in\mathcal{C}_{\mathbf{y}}^+}\sum_{\mathbf{x}\in P} b_A(\mathbf{x}) \gg_{t,T_1,T_2} \delta^{9m/4}p^n|[Q_0]|.$$

By the pigeon-hole principle, there exists $\mathbf{y}$ and $P \in \mathcal{C}_{\mathbf{y}}^+$ such that

$$\sum_{\mathbf{x}\in P} b_A(\mathbf{x}) \gg_{t,T_1,T_2} \delta^{9m/4}|P|.$$

From this (3.6) finally follows. □

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

A. *Appendix. Subgroups of* $\mathbb{Z}^n$

Given a subgroup $H$ of $\mathbb{Z}^n$ we call the dimension of the subspace of $\mathbb{R}^n$ generated by $H$ the *rank* of $H$.

LEMMA A.1. *If $H \leqslant \mathbb{Z}^n$ has rank $r$, then it is generated by $r$-linearly independent vectors.*

Given vectors $\mathbf{h}_1, \ldots, \mathbf{h}_r \in \mathbb{Q}^n$, let us write $(\mathbf{h}_1, \ldots, \mathbf{h}_r)$ for the matrix whose $i$th column is $\mathbf{h}_i$. Lemma A.1 tells us that if $H \leqslant \mathbb{Z}^n$ has rank $r$ then we can find linearly independent vectors $\mathbf{h}_1, \ldots, \mathbf{h}_r$ such that

$$(\mathbf{h}_1, \ldots, \mathbf{h}_r) \cdot \mathbb{Z}^r = \{m_1\mathbf{h}_1 + \cdots + m_r\mathbf{h}_r : \mathbf{m} \in \mathbb{Z}^r\} = H.$$

*Proof.* Since the subspace generated by $H$ has dimension $r$, the subgroup $H$ must itself contain $r$ linearly independent vectors $\mathbf{h}_1, \ldots, \mathbf{h}_r$. There must also exist $n - r$ standard basis vectors $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_{n-r}}$ which, taken together with $\mathbf{h}_1$, $\ldots$, $\mathbf{h}_r$, form a basis for $\mathbb{Q}^n$. Let $\mathbf{h}_{r+j} = \mathbf{e}_{i_j}$. Notice that $|\det(\mathbf{h}_1, \ldots, \mathbf{h}_n)|$ is a positive integer, in particular it is at least one. Every element in $H$ is a rational combination of $\mathbf{h}_1, \ldots, \mathbf{h}_r$. If every element of $H$ is an integer combination of $\mathbf{h}_1, \ldots, \mathbf{h}_r$, then we're done. Suppose otherwise, so we have some $\mathbf{u} \in H$ with

$$\mathbf{u} = \lambda_1\mathbf{h}_1 + \cdots + \lambda_r\mathbf{h}_r,$$

where not all $\lambda_i$ are integers. Relabelling if necessary, we may assume $0 < \{\lambda_1\} < 1$. Considering $-\mathbf{u}$ if necessary, we may assume $0 < \{\lambda_1\} \leqslant 1/2$. Subtracting integer multiples of $\mathbf{h}_1$, we may assume $\lambda_1 = \{\lambda_1\}$. Now $\mathbf{u}, \mathbf{h}_2, \ldots,$ $\mathbf{h}_r$ are linearly independent elements of $H$ with

$$(\mathbf{u}, \mathbf{h}_2, \ldots, \mathbf{h}_n)^T = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_r & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \cdot (\mathbf{h}_1, \ldots, \mathbf{h}_n)^T.$$

Therefore $|\det(\mathbf{u}, \mathbf{h}_2, \ldots, \mathbf{h}_n)| = \lambda_1|\det(\mathbf{h}_1, \ldots, \mathbf{h}_n)| \leqslant 2^{-1}|\det(\mathbf{h}_1, \ldots, \mathbf{h}_n)|$. If every element of $H$ is an integer combination of $\mathbf{u}, \mathbf{h}_2, \ldots, \mathbf{h}_n$ then we are done. Otherwise we iterate the above procedure. This process must terminate in at most $t$ iterations, where $t$ is any positive integer satisfying $2^{-t}|\det(\mathbf{h}_1, \ldots, \mathbf{h}_n)| < 1$. We are thus able to find linearly independent vectors $\mathbf{h}'_1, \ldots, \mathbf{h}'_r \in H$ whose integer span equals $H$. □

LEMMA A.2. *Let $H \leqslant \mathbb{Z}^n$ have rank $r$. There exists a constant $C_H > 0$ such that for all $N$ we have the estimates*

$$C_H^{-1}N^r \leqslant |H \cap [-N, N]^n| \leqslant C_H N^r. \tag{A.1}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*Proof.* By Lemma A.1 there exist linearly independent vectors $\mathbf{h}_1, \dots, \mathbf{h}_r$ such that $H = (\mathbf{h}_1, \dots, \mathbf{h}_r) \cdot \mathbb{Z}^r$. Consider the normed spaces $(U, \|\cdot\|_U)$, $(V, \|\cdot\|_V)$ where $U = \mathbb{R}^r$, $V = (\mathbf{h}_1, \dots, \mathbf{h}_r) \cdot \mathbb{R}^r$,

$$\|\mathbf{u}\|_U = \max_{1 \leqslant i \leqslant r} |u_i| = |\mathbf{u}|_\infty \quad \text{and} \quad \|\mathbf{v}\|_V = \max_{1 \leqslant i \leqslant n} |v_i| = |\mathbf{v}|_\infty.$$

Notice that the map $\mathbf{u} \mapsto (\mathbf{h}_1, \dots, \mathbf{h}_r) \cdot \mathbf{u}$ induces a bijective linear map from $U$ to $V$. It follows from the elementary theory of normed spaces that there exists a constant $C = C(\mathbf{h}_1, \dots, \mathbf{h}_r) > 0$ such that for any $\mathbf{u} \in \mathbb{R}^r$ we have the estimates

$$C^{-1}|\mathbf{u}|_\infty \leqslant \left| \sum_{i=1}^{r} u_i \mathbf{h}_i \right|_\infty \leqslant C|\mathbf{u}|_\infty. \tag{A.2}$$

Using this, we see that

$$\begin{aligned}(\mathbf{h}_1, \dots, \mathbf{h}_r) \cdot [\pm N/C]^r &\subset H \cap [-N, N]^n \\ &\subset (\mathbf{h}_1, \dots, \mathbf{h}_r) \cdot [\pm CN]^r.\end{aligned}$$

By (1.3), we therefore have

$$(2N/C - 1)^r \leqslant |H \cap [-N, N]^n| \leqslant (2CN + 1)^r. \qquad \square$$

COROLLARY A.3. *Let $H \leqslant \mathbb{Z}^n$ have rank $r$. There exists a constant $C_H$ such that for any $N$, if $t(N)$ denotes the number of cosets of $H$ which intersect $[\pm N]^n$, then*

$$t(N) \leqslant C_H N^{n-r}.$$

*Proof.* If $(\mathbf{x} + H) \cap [\pm N]^n \neq \emptyset$, then $\mathbf{x} + H = \mathbf{y} + H$ for some $\mathbf{y} \in [\pm N]^n$. Let $\mathbf{y}_1, \dots, \mathbf{y}_t \in [\pm N]^n$ be such that $\mathbf{y}_1 + H, \dots, \mathbf{y}_t + H$ are the distinct cosets of $H$ intersecting $[\pm N]^n$. Notice that

$$\begin{aligned}|(\mathbf{y}_i + H) \cap [\pm 2N]^n| &\geqslant |\mathbf{y}_i + (H \cap [\pm N]^n)| \\ &= |H \cap [\pm N]^n| \\ &\gg_H N^r,\end{aligned}$$

the final inequality following from Lemma A.2. Therefore

$$\begin{aligned}tN^r &\ll_H \sum_{i=1}^{t} |(\mathbf{y}_i + H) \cap [\pm 2N]^n| \\ &\leqslant |[\pm 2N]^n| \\ &= (4N + 1)^n.\end{aligned}$$

Thus $t \ll_H N^{n-r}$. $\square$

COROLLARY A.4. *Let $H \leqslant \mathbb{Z}^n$ have rank $r$ and let $Q \subset \mathbb{R}^n$ be a half-open cube of side-length $L$ with $A \subset [Q]$. Then there are at least $|A|^2 L^{r-n}/C_H$ pairs $(\mathbf{x}, \mathbf{y}) \in A^2$ with $\mathbf{x} - \mathbf{y} \in H$.*

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*Proof.* Let $\mathbf{g} \in \mathbb{Z}^n$. Then by the Cauchy–Schwarz inequality we have

$$\begin{aligned}\sum_{\mathbf{y}-\mathbf{z}\in H+\mathbf{g}} 1_A(\mathbf{y})1_A(\mathbf{z}) &= \sum_{\mathbf{u}} \sum_{\mathbf{y}\in H+\mathbf{g}+\mathbf{u}} 1_A(\mathbf{y}) \sum_{\mathbf{z}\in H+\mathbf{u}} 1_A(\mathbf{z}) \\ &\leqslant \Big(\sum_{\mathbf{u}} \Big|\sum_{\mathbf{y}\in H+\mathbf{g}+\mathbf{u}} 1_A(\mathbf{y})\Big|^2\Big)^{1/2} \Big(\sum_{\mathbf{u}} \Big|\sum_{\mathbf{z}\in H+\mathbf{u}} 1_A(\mathbf{z})\Big|^2\Big)^{1/2} \\ &= \sum_{\mathbf{u}} \Big(\sum_{\mathbf{y}\in H+\mathbf{u}} 1_A(\mathbf{y})\Big)^2 \\ &= \sum_{\mathbf{y}-\mathbf{z}\in H} 1_A(\mathbf{y})1_A(\mathbf{z}).\end{aligned}$$

Let $\mathbf{g}_1 + H, \dots, \mathbf{g}_t + H$ denote the set of cosets of $H$ for which there exist $\mathbf{y}, \mathbf{z} \in A$ with $\mathbf{y} - \mathbf{z} \in \mathbf{g}_i + H$. Notice that for each $i$ we have

$$(\mathbf{g}_i + H) \cap [-L, L]^n \neq \emptyset.$$

Hence by Corollary A.3, the number $t$ of such cosets is at most $C_H L^{n-r}$. It follows that

$$\begin{aligned}|A|^2 &= \sum_{i=1}^{t} \sum_{\mathbf{y}-\mathbf{z}\in\mathbf{g}_i+H} 1_A(\mathbf{y})1_A(\mathbf{z}) \\ &\leqslant C_H L^{n-r} \sum_{\mathbf{y}-\mathbf{z}\in H} 1_A(\mathbf{y})1_A(\mathbf{z}).\end{aligned}$$

□

B. *Appendix. Using Shkredov's theorem*

THEOREM B.1 (Shkredov [14]). *There exists an absolute constant $C$ such that if $A \subset \{1, \dots, N\}^2$ does not contain a configuration of the form*

$$\{\mathbf{x}, \mathbf{x} + (d, 0), \mathbf{x} + (0, d)\} \text{ with } d \neq 0,$$

*then $A$ satisfies the bound*

$$|A| \leqslant CN^2(\log\log N)^{-1/22}.$$

In this section we show how to use Theorem B.1 to obtain the following result.

THEOREM B.2. *Let $T_1$, $T_2$ be $n \times n$ rational matrices, not necessarily non-singular, and let $K$ denote a finite union of proper subspaces of $\mathbb{R}^n$. There exists an absolute constant $C$, dependent only on $T_1, T_2$ and $K$, such that for any integer $N \geqslant 3$ and $A \subset \{1, 2, \dots, N\}^n$, if $A$ does not contain any configurations of the form $\{\mathbf{x}, \mathbf{x} + T_1\mathbf{d}, \mathbf{x} + T_2\mathbf{d}\}$ with $\mathbf{d} \in \mathbb{Z}^n \setminus K$, then $A$ satisfies the bound*

$$|A| \leqslant \frac{CN^n}{(\log\log N)^{1/22}}. \tag{B.1}$$

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

*Proof.* Pick $\mathbf{d} \in \mathbb{Z}^n \setminus K$. Multiplying $\mathbf{d}$ by a sufficiently large positive integer, we may assume that the vectors $\mathbf{u} = T_1\mathbf{d}$ and $\mathbf{v} = T_2\mathbf{d}$ are both elements of the integer lattice $\mathbb{Z}^n$. We will show that if $A$ does not contain any configurations of the form $\{\mathbf{x}, \mathbf{x} + m\mathbf{u}, \mathbf{x} + m\mathbf{v}\}$ with $m \in \mathbb{Z} \setminus \{0\}$, then $A$ satisfies a bound of the form (B.1). Clearly this establishes the result.

Let $|A| = \delta N^n$. There are two cases to consider. The first is when $\mathbf{u}$ and $\mathbf{v}$ are colinear, say $t_1\mathbf{v} = t_2\mathbf{u}$ for some $t_1, t_2 \in \mathbb{Z}$. By the pigeon-hole principle and Lemma A.3, the set $A$ must have size $\gg_{n,\mathbf{u}} \delta N$ on some line of the form $\mathbf{x} + \mathbb{Z} \cdot \mathbf{u}$. Let $A'$ denote the set of $y \in \mathbb{Z}$ with $\mathbf{x} + y\mathbf{u} \in A$. Since $A'$ contains no configurations of the form $\{x, x + t_1 d, x + t_2 d\}$, we can use Roth's theorem [**10**] to conclude that $|A'| \ll_{t_1,t_2} N(\log\log N)^{-1}$. The bound (B.1) then follows with room to spare.

Next suppose that $\mathbf{u}$ and $\mathbf{v}$ are not colinear. Consider the set $A_{\mathbf{y}}$ of pairs $(a, b) \in \mathbb{Z}^2$ with $\mathbf{y} + a\mathbf{u} + b\mathbf{v} \in A$. Using results from Appendix A, one sees that some $A_{\mathbf{y}}$ must have size $\gg_{\mathbf{v},\mathbf{u}} \delta N^2$. Moreover, there exists an absolute constant $C = C(\mathbf{u}, \mathbf{v})$ such that $A_{\mathbf{y}}$ is contained in the cube $[-CN, CN]^2$. Since $A_{\mathbf{y}}$ does not contain any configurations of the form $\{\mathbf{x}, \mathbf{x} + (m, 0), \mathbf{x} + (0, m)\}$ with $m \in \mathbb{Z} \setminus 0$, we see from Shkredov's theorem (Theorem B.1) that $A_{\mathbf{y}}$ has size $\ll_{u,v} N^2(\log\log N)^{-1/22}$. The bound (B.1) then also follows in this case. □

*Acknowledgement.* During the completion of this work, the author was supported by an EPSRC doctoral training grant through the University of Bristol.

## *References*


1. E. S. Croot III and V. F. Lev, Open problems in additive combinatorics. In *Additive Combinatorics* (*CRM Proceedings and Lecture Notes* **43**), American Mathematical Society (Providence, RI, 2007), 207–233.
2. H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations. *J. Anal. Math.* **34** (1978), 275–291.
3. W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* **8**(3) (1998), 529–551.
4. W. T. Gowers, A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11**(3) (2001), 465–588.
5. B. J. Green, Finite field models in additive combinatorics. In *Surveys in Combinatorics 2005* (*London Mathematical Society Lecture Note Series* **327**), Cambridge University Press (Cambridge, 2005), 1–27.
6. B. J. Green and T. Tao, An inverse theorem for the Gowers $U^3(G)$ norm. *Proc. Edinb. Math. Soc.* (2) **51**(1) (2008), 73–153.
7. B. J. Green and T. Tao, New bounds for Szemerédi's theorem. II. A new bound for $r_4(N)$. In *Analytic Number Theory*, Cambridge University Press (Cambridge, 2009), 180–204.
8. D. R. Heath-Brown, Integer sets containing no arithmetic progressions. *J. Lond. Math. Soc.* (2) **35**(3) (1987), 385–394.
9. K. F. Roth, On certain sets of integers. *J. Lond. Math. Soc.* **28** (1953), 104–109.
10. K. F. Roth, On certain sets of integers. II. *J. Lond. Math. Soc.* **29** (1954), 20–26.
11. T. Sanders, On Roth's theorem on progressions. *Ann. of Math.* (2) **174**(1) (2011), 619–636.
12. W. M. Schmidt, *Small Fractional Parts of Polynomials (Regional Conference Series in Mathematics* **32***)*, American Mathematical Society (Providence, RI, 1977).
13. I. D. Shkredov, On a generalization of Szemerédi's theorem. *Proc. Lond. Math. Soc.* (3) **93**(3) (2006), 723–760.


https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press

**14.** I. D. Shkredov, On a two-dimensional analogue of Szemerédi's theorem in abelian groups. *Izv. Ross. Akad. Nauk Ser. Mat.* **73**(5) (2009), 181–224.
**15.** E. Szemerédi, Integer sets containing no arithmetic progressions. *Acta Math. Hungar.* **56**(1–2) (1990), 155–158.

Sean Prendiville,
Department of Mathematics and Statistics,
University of Reading,
Whiteknights, PO Box 220,
Reading, RG6 6AX,
U.K.
E-mail: s.m.prendiville@reading.ac.uk

https://doi.org/10.1112/S0025579314000163 Published online by Cambridge University Press